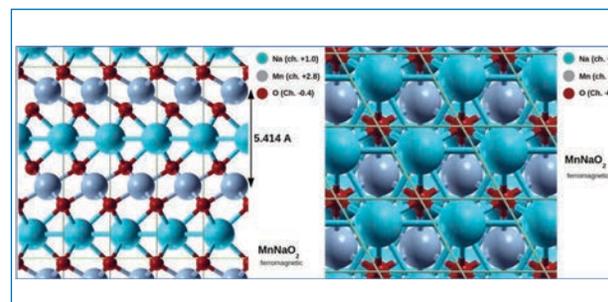
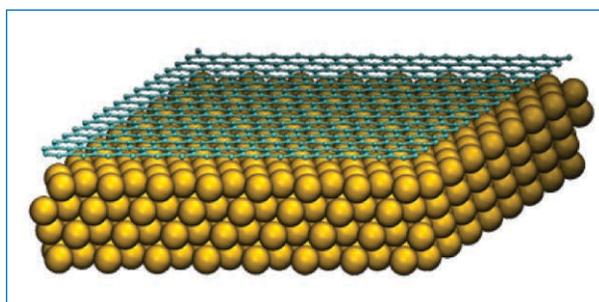
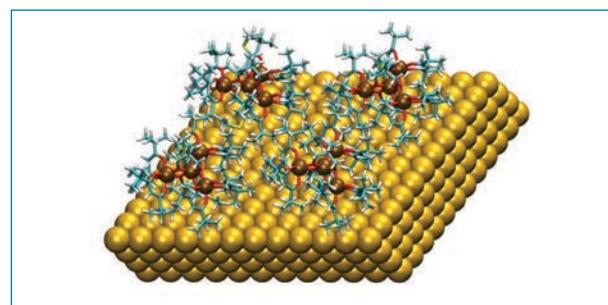
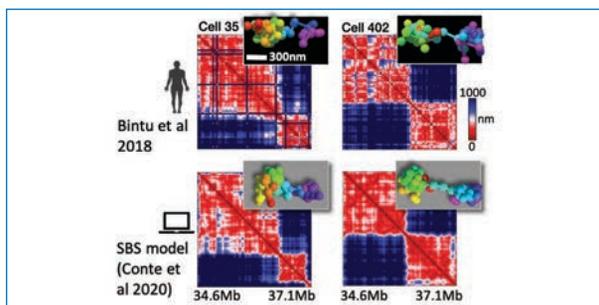
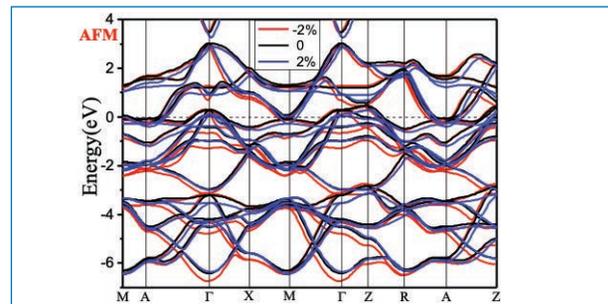
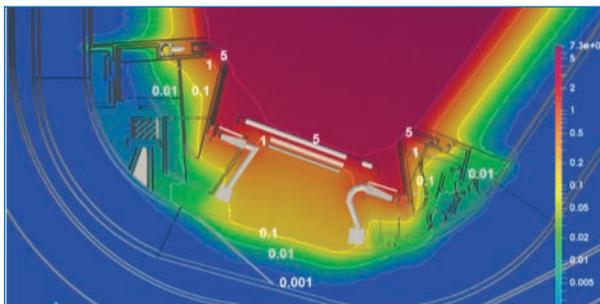


# High Performance Computing on CRESCO infrastructure: research activities and results 2019



HIGH PERFORMANCE COMPUTING ON  
CRESCO INFRASTRUCTURE:  
RESEARCH ACTIVITY AND RESULTS 2019

December 2020

Contributions provided by a selection of users of the CRESCO infrastructure

Scientific Editor:

*Guido Guarnieri* and *Giuseppe Santomauro*,  
ENEA, TERIN-ICT-HPC, CR Portici

Acknowledgements:

*Gianclaudio Ferro*,  
ENEA, TERIN-ICT-HPC, CR Bologna

Cover:

*Amedeo Trolese*,  
ENEA, TERIN-ICT, CR Frascati

ISBN: 978-88-8286-403-3

# Contents

<i>1. Preface</i>	5
<i>2. High Performance I/O Benchmarks for Energy Oriented Center of Excellence Project F. Iannone, S. Migliori, G. Bracco, M. Celino, S. Pecoraro, F. Ambrosino, A. Funel, G. Formisano, G. Guarnieri, F. Palombi, G. Ponti, G. Santomauro, F. Simoni, M. De Rosa and TERIN-ICT team</i>	7
<i>3. Biaxial Strain Effects on the Electron-Phonon Coupling in Iron-Selenide: A First Principles Study R. Zhang and G. De Marzi</i>	21
<i>4. Neural Networks for Statistical Process Control of Multiple Stream Processes A. Lepore, F. Ambrosino, G. Sposito and B. Palumbo</i>	29
<i>5. Behavior of the Trapped Solvent at the Interface between Block Copolymers Films and Random Copolymer Brush G. Munao, G. Donati, A. De Nicola, A. Pizzirusso, S. Caputo and G. Milano</i>	34
<i>6. Calculation of Lattice Thermal Expansion for a Class of Organic Crystals A. Ienco, P. Rossi and P. Paoli</i>	38
<i>7. Kinetic Analysis of Accelerator Driven System with MCNP 6.2 and JEFF 3.3 Nuclear Dataset on the CRESCO Computational Facility: The KUCA Reactor model N. Burgio and A. Santagata</i>	42
<i>8. Strategy Diversity Promotes Time Dependence Coexistence of Cooperation and Defection in Evolutionary Games I. Sendina-Nadal, I. Leyva, J.A. Almendral and S. Boccaletti</i>	46
<i>9. Computational Determination of Critical Currents in Granular Superconductors G. Mancini</i>	51
<i>10. Synchronization in Heterogeneous Ensembles of Oscillators: Explosive Transitions and Dynamical Control R. Gutierrez</i>	58
<i>11. The ENEA-REG System, A Multi-Component Regional Earth System Model A. Anav, A. Carillo, M. Palma, M.V. Struglia, U.U. Turuncoglu and G. Sannino</i>	62

<i>12. A Simple Procedure to Simulate a Small Angle X Ray Scattering Experiment</i> <i>E. Burresti and L. Tapfer</i>	<b>67</b>
<i>13. Molecular Dynamics Calculations for the Determination of PuO<sub>2</sub> Heat Capacity</i> <i>R. Calabrese</i>	<b>71</b>
<i>14. Hybrid Particle Field Model of Lipid A</i> <i>A. De Nicola, A. Pizzirusso, G. Donati and S. Caputo</i>	<b>75</b>
<i>15. Dose Calculation in DEMO PHTS</i> <i>P. Chiovaro and A. Del Nevo</i>	<b>80</b>
<i>16. Preliminary Computational Tests for Extracting Atmospheric Transfer Matrices in GAINS-TALY</i> <i>G. Briganti, A. Cappelletti, L. Ciancarella, I. D'Elia, M. D'Isidoro, A. Piersanti, M. Adani and G. Guarnieri</i>	<b>87</b>
<i>17. From Prototype to Production: Monte Carlo Studies for Updating the ENEA Neutron Active Interrogation System</i> <i>G. Gandolfo, L. Lepore, N. Cherubini and G.A. Marzo</i>	<b>92</b>
<i>18. Applications of hybrid simulation to fast frequency chirping of reversed shear Alfvén eigenmode</i> <i>T. Wang</i>	<b>96</b>
<i>19. Monte Carlo Calculation of the k<sub>Q</sub>, Q<sub>0</sub> Quality Correction Factors for a Micro Ionization Chamber for Radiotherapy Dosimetry</i> <i>M. Alonzo and M. Pimpinella</i>	<b>101</b>
<i>20. ABF Simulations for the Reconstruction of Free Energy Profiles Describing the Switching Mechanism of a pH-dependent DNA Nanodevice</i> <i>F. Iacovelli, A. Romeo and M. Falconi</i>	<b>105</b>
<i>21. The Molecular Way to Quantum Bits and Spintronics: an ab Initio study</i> <i>F. Totti, F. Buonocore, M. Celino, M. Briganti and A. Albino</i>	<b>110</b>
<i>22. Monte Carlo Simulations Supporting Studies of PWR's and of Experimental N_TOF Facility at CENR</i> <i>P. Console Camprini</i>	<b>117</b>
<i>23. Computation of Solar Sail Thrust by Using Light Scattering Models</i> <i>D. Zola and S. Scaglione</i>	<b>122</b>

<b>24. <i>Ab Initio Calculations Approach to Study a-SI:H/c-SI Interfaces</i></b>	
<i>F. Buonocore, P.L. Garcia-Muller, R. Mayo-Garcia, S. Giusepponi and M. Celino</i>	<b>126</b>
<b>25. <i>First Principles Molecular Dynamics Simulations of Amorphous GeO<sub>2</sub> Under Pressure in the range 0-30GPa</i></b>	
<i>G. Mancini and M. Celino</i>	<b>133</b>
<b>26. <i>Large Eddy Simulation Of H<sub>2</sub>/O<sub>2</sub> Combustion With O<sub>2</sub> Transcritical Injection</i></b>	
<i>E. Giacomazzi, D. Cecere, N. Arcidiacono and F.R. Picchia</i>	<b>137</b>
<b>27. <i>Integrative Polymer Modelling and Numerical Simulations to Predict the Complex 3D Spatial Organization of Chromosomes</i></b>	
<i>M. Conte, A. Esposito, L. Fiorillo, A.M. Chiariello, A. Abraham and S. Bianco</i>	<b>143</b>
<b>28. <i>Ab Initio Study of NA Insertion at TiO<sub>2</sub> Anatase Surfaces as Negative Electrode for Sodium Batteries</i></b>	
<i>A. Massaro, A.B. Munoz Garcia and M. Pavone</i>	<b>151</b>
<b>29. <i>Knowledge Extraction from Social Media Web Sources: Elements Affecting Web Crawling and Data Analytics Tasks in ENEAGRID</i></b>	
<i>D. Alderuccio and S. Migliori</i>	<b>155</b>
<b>30. <i>Neutronics Analysis of DEMO Divertor</i></b>	
<i>R. Villari, D. Flammini, G. Mariano, G. Mazzone, F. Moro and S. Noce</i>	<b>159</b>
<b>31. <i>Monte Carlo Simulation of Temporary Storage Facilities for Radioactive Waste: Exposure Evaluations</i></b>	
<i>L. Lepore, G. Gandolfo, N. Cherubini and G.A. Marzo</i>	<b>163</b>
<b>32. <i>Study of Advanced Configurations on Tokamak EAST with the EMC3 EIRENE</i></b>	
<i>C. Meineri</i>	<b>167</b>
<b>33. <i>Simulation of Electrical and Thermal Properties of Carbon Nanotubes at Nanoscale</i></b>	
<i>Y. Zhao, G. Donati, A. De Nicola, A. Pizzirusso, G. Munao, S. Caputo and G. Milano</i>	<b>171</b>
<b>34. <i>AB-Initio Calculations of Cathode Materials for Rechargeable Batteries</i></b>	
<i>A. Mosca Conte, F. Buonocore, P. Gori, M. Celino, P.P. Prosini and O. Pulci</i>	<b>177</b>
<b>35. <i>Exact Invariant Approach to the Investigation of Nonlinear Dynamics of Chirping Alfvén Modes</i></b>	
<i>Sergio Briguglio, X. Wang, C. Di Troia, G. Fogaccia , V. Fusco and G. Vlad</i>	<b>183</b>

<b>36. Peptide Bond Recognition via Vibration Assisted Electron Tunneling in Graphene Nanodevice</b>	
<b>G. Zollo</b>	<b>187</b>
<b>37. Neutronic Analyses of Plasma Facing Components of Demo Divertor</b>	
<b>S. Noce, D. Flammini, G. Mariano, G. Mazzone, F. Moro, F. Romanelli, R. Villari and J. You</b>	<b>201</b>
<b>38. Neutronic Analyses in Support of the Conceptual Design of the DTT Tokamak Radial Neutron Camera</b>	
<b>B. Caiffi, M. Angelone, A. Colangeli, D. Flammini, N. Fonnesu, G. Mariano, F. Moro and R. Villari</b>	<b>205</b>
<b>39. Thermal Enhancement of Data Center with Artful Workload Assignment</b>	
<b>D. De Chiara, M. Chinnici, A. Kor, E. Rondeau, J. Georges, A. Grishina and G. Guarnieri</b>	<b>210</b>
<b>40. Benchmark of the LAMMPS Code on CRESCO and SUMMIT HPC System</b>	
<b>M. Lupo Pasini, S. Giusepponi, F. Buonocore, M. Celino, A. Frattolillo and S. Migliori</b>	<b>225</b>
<b>41. Validation and Development of the FNG Source Routine on FLUKA</b>	
<b>A. Calamida, A. Colangeli, M. Pillon and S. Fiore</b>	<b>231</b>
<b>42. Nonequilibrium Alchemy in Sampl7 Challenge</b>	
<b>P. Procacci and G. Guarnieri</b>	<b>237</b>
<b>43. Lattice Atomistic Kinetic Monte Carlo simulation with NVIDIA GPUs</b>	
<b>M. D. Mazzeo and S. Pagnutti</b>	<b>243</b>
<b>44. Author Index</b>	<b>248</b>

## Foreword

During the year 2019, the CRESCO high performance computing clusters have provided 112 millions hours of “core” computing time, at a high availability rate, to 127 users, supporting ENEA research and development activities in many relevant scientific and technological domains. In the framework of joint programs with ENEA researchers and technologists, computational services have been provided also to academic and industrial communities. This report, the tenth of a series started in 2008, is a collection of 42 papers illustrating the main results obtained during 2019 using the CRESCO/ENEAGRID HPC facilities. The significant number of contributions proves the importance of the HPC facilities in ENEA for the research community. The topics cover various fields of research, such as materials science, efficient combustion, climate research, nuclear technology, plasma physics, biotechnology, aerospace, complex systems physics, geophysical flow, renewable energies, environmental issues, HPC technology. The report shows the wide spectrum of applications of high performance computing, which has become an all-round enabling technology for science and engineering.

Since 2008, the main ENEA computational resources is located near Naples, in Portici Research Centre. This is a result of the CRESCO Project (Computational Centre for Research on Complex Systems), co-funded, in the framework of the 2001-2006 PON (European Regional Development Funds Program), by the Italian Ministry of Education, University and Research (MIUR).

The CRESCO Project provided the financial resources to set up the first HPC x86\_64 Linux cluster in ENEA; a major computing installation for both the Italian and the International context: it ranked 126 in the HPC Top 500 June 2008 world list, with 17.1 Tflops and 2504 cpu cores. It was later decided to keep CRESCO as the name for all the Linux clusters in the ENEAGRID infrastructure, which integrates all ENEA scientific computing systems, and is currently distributed in six Italian sites. CRESCO computing resources were later upgraded in the framework of PON 2007-2013 with the project TEDAT and the cluster CRESCO4, 100 Tflops computing power.

In 2015 ENEA and CINECA, the main HPC institution in Italy, signed a collaboration agreement to promote joint activities and projects in HPC. In this framework, CINECA and ENEA participated successfully to a selection launched by EUROfusion, the European Consortium for the Development of Fusion Energy, for the procurement of a several PFlops HPC system, beating the competition of 7 other institutions. The new system MARCONI-FUSION started operation in July 2016 and it has been extended since 2023 with a power peak of 8 PFlops of conventional processors and 2 PFlops of accelerated co-processors.

The ENEA-CINECA agreement is a promising basis for the future development of ENEA HPC resources. A new CRESCO6 cluster of 1.4 Pflops has been installed in 2018 ranked 420th in November 2018 Top500 list representing the TIER 1 facility in the HPC ecosystem of the national scientific community.

In 2019 the ENEAGRID computational resources consist of more than 25000 computing cores (in production) and a raw data storage of about 5.2 PB in High Performance Storage RAID systems and 2.5 PB of tapes library.

The distributed architecture of ENEAGRID has allowed to implement a high availability layout for mission critical applications, such as numerical forecast models, deploying the CRESCO4f cluster at ENEA CR Frascati.

The success and the quality of the results produced by CRESCO stress the role that HPC facilities can play in supporting science and technology for all ENEA activities, national and international collaborations, and the ongoing renewal of the infrastructure provides the basis for an upkeep of this role in the forthcoming years.

Dipartimento Tecnologie Energetiche e Fonti Rinnovabili  
Divisione per lo Sviluppo Sistemi per l'Informatica e l'ICT  
CRESCO Team

# HIGH PERFORMANCE I/O BENCHMARKS FOR ENERGY ORIENTED CENTER OF EXCELLENCE PROJECT

F.Iannone\*, S.Migliori, G.Bracco, M. Celino, S.Pecoraro, F.Ambrosino, A.Funel,  
G.Formisano, G.Guarnieri, F.Palombi, G.Ponti, G.Santomauro, F.Simoni, M. De Rosa and  
TERIN-ICT team

*ENEA, TERIN-ICT, via Lungotevere Thaon de Ravel, Rome, Italy*

**ABSTRACT.** ENEA is a partner in Energy Oriented Center of Excellence (EoCoE-II) project aiming to develop a sustainable structure able to exploitation the usage of HPC in the energy domain to a wide user community comprising both academic and industrial users. The paper describes the development and test activities carried out by ENEA within the WP4: I/O & Data Flow of EOCoe II.

## 1 Introduction

Storage has become a key component in HPC systems, and the challenges for the Exascale era are huge. I/O performance is becoming a bottleneck in many cases of large simulations in HPC supercomputers requiring fault-tolerance techniques such as checkpointing/restart. Usually these techniques require a massive data movement in order to save the state of a running application. The data writing and reading time can consume a significant part of the overall application runtime and should be minimized using two approaches: *i*) improving the IO performances sizing high performance parallel filesystem in term of IO server nodes and high performance storage systems; *ii*) reducing the IO dataflow without losing necessary information via in situ and in transit processing as well as moving postprocessing elements directly into the frame of the running application. Both approaches require a middleware SW to separate the IO functions of the user applications from the IO libraries currently available in the HPC ecosystem such as: HDF5 parallel, MPI-IO, ADIOS, SionLib.

As the volume of data and application reads and writes increases, it is important to assess the scalability of I/O operations as a key contributor to overall application performance. Optimizing I/O performance presents unique challenges for application developers and performance measurement and analysis tools, as it often involves an integration of multiple techniques to observe performance and create integrated I/O performance views from multiple I/O layers.

## 2 EoCoE-II WP4: I/O and Data Flow

The paper describes the development and test activities carried out by ENEA within the WP4: I/O & Data Flow of EOCoe II. In particular ENEA is involved in the task 4.2.2 focusing on the overall I/O runtime and its improvement within Energy by Fusion area.

The I/O runtime is being deployed by means a Parallel Data Interface (PDI) development based on a middleware layer able to decouple fusion gyrokinetic codes, like GYSELA, from I/O functions. The PDI supports a plugin system to integrate existing I/O libraries, such as: Posix, NetCFD, HDF5, SionLib, embedded in many simulation code, therefore it is a good solution to optimize the Gysela code having as main output files in HDF5 format obtained by means functions library without MPI/Parallel support, making a bottleneck for the massive IO checkpoint jobs. A flexible way to improve the portability and maintenance of

Gysela code, making it independent from I/O libraries, is to develop a plugin for the PDI once it is integrated into the code.

Several plugins for PDI are under development in EOCO tasks of WP4 as well as the PDI Gysela code integration and waiting for a deployment of a stable release, ENEA activity in WP4 has been focused for benchmarking the overhead of PDI middleware using the standard I/O performance tool: IOR.

## 2.1 PDI: Parallel Data Interface

Parallel Data Interface (PDI) [1] offers a declarative API for simulation codes to expose information required by the implementation of I/O strategies. The I/O strategies are encapsulated inside plugins that access the exposed information. A weak coupling mechanism enables to connect both sides through a configuration file. In order to decouple this I/O implementation code both from PDI and from the simulation code, it is defined in dedicated object files that can either be loaded statically or dynamically (a plugin system). This means that PDI does not depend on any I/O library, only its plugins do. This also simplifies changing strategy from one execution to the other as the plugins to load are specified in the configuration file. PDI is freely and publicly available under a BSD license. It is written in C and offers a C API with Fortran bindings to the simulation code. This covers uses from C, C++ and Fortran, the three most widespread languages in the HPC community.

The PDI API contain functions to initialize and finalize the library, change the error handling behavior, emit events and expose buffers as presented in *Listing 1*. The initialization function takes the library configuration (a reference to the content of a YAML file) and the world MPI communicator that it can modify to exclude ranks underlying libraries reserve for I/O purpose. The error handling function enables to replace the callback invoked when an error occurs. The event function takes a character string as parameter that identifies the event to emit.

<pre>enum PDI_inout_t{PDI_IN=1,PDI_OUT=2,PDI_INOUT=3}; PDI_status_t PDI_init(PC t ree t conf , MPI Comm world ) ; PDI_status_t PDI_finalize ( ) ; PDI_errhandler_t PDI_errhandler(PDI_errhandler_t handler); PDI_status_t PDI_event(const char_event); PDI_status_t PDI_share(const char_name ,void data,PDI_inout_t access ); PDI_status_t PDI_access(const char_name,void_data,PDI_inout_t access ); PDI_status_t PDI_release(const char_name); PDI_status_t PDI_reclaim(const char_name);</pre>	<pre>PDI_status_t PDI_export(const charname,void_data); PDI_status_t PDI_expose(const char name,void_data ); PDI_status_t PDI_import(const char name,void_data ); PDI_status_t PDI_exchange(const char name,void_data ); PDI_status_t PDI_transaction_begin(const char_name ); PDI_status_t PDI_transaction_end( );</pre>
<p><i>Listing 1: The PDI public API</i></p>	<p><i>Listing 2: Simplified PDI API for buffer exposing</i></p>

The most interesting functions of this API are however the buffer sharing functions. They support sharing a buffer with PDI identified by a name character string and with a specified access direction specifying that information flows either to PDI (*PDI\_OUT*, read-only share), from PDI (*PDI\_IN*, write-only share) or in both directions (*PDI\_INOUT*). The *PDI\_share* and *PDI\_access* functions start a buffer sharing section while the *PDI\_release* or *PDI\_reclaim* function end it. *PDI\_share* is used for a buffer whose memory was previously owned by the user code while *PDI\_access* is used to access a buffer previously unknown to the user code. Reciprocally, *PDI\_reclaim* returns the memory responsibility to the user code

while *PDI\_release* releases it to PDI. In a typical code, the buffers are however typically shared for a brief period of time between two access by the code. The previously introduced API requires two lines of code to do that. The API presented in *Listing 2* simplifies this case. Its four first functions define a buffer sharing section that lasts during the function execution only. The functions differ in terms of access mode for the shared buffer: *share(OUT) + release* for *PDI\_export*; *share(OUT) + reclaim* for *PDI\_expose*; *share(IN) + reclaim* for *PDI\_import*; and, *share(INOUT) + release* for *PDI\_exchange*.

This API has the disadvantage that it does not enable to access multiple buffers at a time in plugins. Each buffer sharing section ends before the next one starts. The two transaction functions solve this. All sharing sections enclosed between calls to these functions have their end delayed until the transaction ends. This effectively supports sharing of multiple buffers together. The transaction functions also emit a named event after all buffers have been shared and before their sharing section ends.

At the heart of PDI is a list of currently shared buffers. Each shared buffer has a memory address, a name, an access and memory mode and a content data type. The access mode specifies whether the buffer is accessible for reading or writing and the memory mode specifies whose responsibility it is to deallocate the buffer memory. The content data type is specified using a type system very similar to that of MPI and is extracted from the YAML configuration file.

The data section of the configuration file (example in *Listing 3*) contains an entry for each buffer, specifying its type. The type can be a scalar, array or record type. Scalar types include all the native integer and floating point of Fortran and C (including boolean or character types.) Array types are specified by a content type, a number of dimensions and a size for each dimension. They support the situation where the array is embedded in a larger buffer with the buffer size and shift specified for each dimension. Record types are specified by a list of typed and named fields with specific memory displacement based on the record address.

The types can be fully described in the YAML file, but this makes them completely static and prevents the size of arrays to change at execution for example. Any value in a type specification can therefore also be extracted from the content of an exposed buffer using a dollar syntax similar to that of bash for example. The syntax supports array indexing and record field access. For the content of a buffer to be accessible this way, it does however needs to be specified in the metadata section of the YAML file instead of its data section. When a metadata buffer is exposed, its content is cached by PDI to ensure that it can be accessed at any time including outside its sharing section. The plugins to load are specified in the plugins section of the configuration file. Each plugin is loaded statically if linked with the application and dynamically otherwise. A plugin defines five function: an initialization function, a finalization function and three event handling functions. The event handling functions are called whenever one of the three types of PDI event occurs, just after a buffer becomes available, just before it becomes unavailable and when a named event is emitted.

<pre> <b>data</b> : my array : { sizes : [ \$N, \$N ] , type : double } <b>metadata</b> :   N: int # data i d and type   it: int <b>plugins</b> :   declh5 : # plug-in name     outputs :       # data to write , dataset and filename       my_array : { var : array2D , file : example\$it.h5 ,       # condition to write       s e l e c t : ( \$it &gt; 0 ) &amp;&amp; ( it % 10 ) } </pre>	<pre> main comm = MPI_COMM_WORLD call PDI_init ( PDI subtree , main comm ) call PDI_transaction_begin ( " checkpt " ) ptr_int =&gt; N; call PDI_expose ( "N" , ptr_int ) ptr_int=&gt;iter; call PDI_expose( " i t " , ptr_int ) call PDI_expose ( "my array" , ptr_A ) call PDI_transaction_end ( ) call PDI_finalize ( ) </pre>
Listing 3: Example of PDI configuration file	Listing 4: Example of PDI API usage

The *plugins* can access the configuration content and the buffer repository. Configuration specific to a given plugin is typically specified under this plugin in the plugins section of the YAML file. The YAML file can however also contain configuration used by plugins in any section. It can for example contain additional information in a buffer description.

Several *plugins* have been developed in the project, such as: a HDF5 *plugin* to read/write files In HDF5 format, the *decl'H5 plugin* interfaces a declarative interface built on top of HDF5 and the *usercode plugin* supports user written code as developed for the integration of PDI in IOR using POSIX as I/O library.

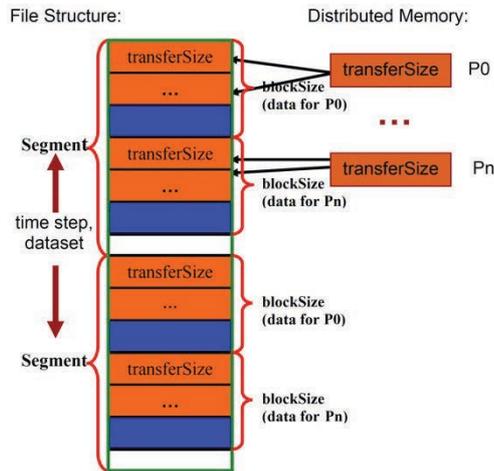
Let us now present an example to show how PDI usage works in practice. *Listing 4* shows the use of the Fortran API to expose to PDI two integers: *N* and *it*, and an array of dimension  $N \times N$ , *my\_array*. The configuration file for this example is the one presented in *Listing 3*. When the *PDI\_init* function is called, the configuration file is parsed and the *decl'H5 plugin* is loaded. This plugin initialization function is called and analyzes its part of the configuration to identify the events to which it should react. No plugin modifies the provided MPI communicator that is therefore returned unchanged. A transaction is then started in which three buffers are exposed: *N*, *it* and *my\_array*. The *decl'H5 plugin* is notified of each of these events but reacts to none. The transaction is then closed that triggers a named event to which the *decl'H5 plugin* does not react as well as three end of sharing section events, one for each buffer. The *decl'H5* reacts to the end of the *my\_array* sharing since this buffer is identified in the configuration file. It evaluates the value of the select clause and if nonzero writes the buffer content in a dataset whose name is provided by the *var* value ("array2D") to a HDF5 file whose name is provided by the *file* value ("*example\$it.h5*").

## 2.2 IOR

IOR [2] can be used for testing performance of parallel file systems using various interfaces (MPIIO, HDF5, PnetCDF, POSIX) and different access patterns. IOR uses MPI for process synchronization. An important feature of IOR is that it can simulate two basic parallel I/O strategies: shared file and one-file per process. In the shared file case all processes read/write to a single common file, while in the one-file per process each process reads/writes its own file. Depending on the selected interface, interface specific configuration options are

available. The fig.3 illustrates the relationship between the file structure and the processors when writing to a shared file.

**Fig.3 The design of the IOR benchmark for shared file type.  
Blocks are stored in separated files for the 1-file-per-processor mode of operation.**



It is organized as a sequence of “segments” that represent the application data for either one simulated time step of a single data variable (eg. pressure for timestep1, 2, 3, etc..) or a sequence of data variables (eg. pressure, temperature, velocity). For high-level file formats such as HDF5 and NetCDF, each segment directly corresponds to a “dataset” object in the nomenclature of these respective file formats. Each segment is divided evenly among the processors who share this data file into units called “blocks” to represent the array re-assembly performed by the parallel I/O layer. The process with rank 0 gets the first block and the process with rank 1 gets the second block and so on. The physical file layout corresponds to the application data resident in the distributed memories. Each block is further divided into many transfer units called *TransferSize*, in order to emulate the strided/stanza-like access patterns required to undo multidimensional domain decompositions, such as reassembling a bunch of 3D subdomains that reside on each processor into a single 3D logical array on disk. The *TransferSize* chunks directly correspond to the I/O transaction size, which is the amount of data transferred from the processor’s memory to file for each I/O function call (eg. the buffer size for a POSIX I/O call). For the one-file-per-processor case, the file structure is nearly identical to the diagram in Figure 3, but except that each process will write/read data to/from its own file (eg. each “block” is packed contiguously in separate files). The following parameters of IOR are important to our study: *API*, *SegmentCount*, *BlockSize*, *FilePer-Proc*, *ReadFile*, *WriteFile*, *TransferSize*, *NumTasks*. The *API* describes which I/O API to use. Currently IOR supports POSIX, MPI-IO, HDF5, and NetCDF APIs. The *ReadFile* and *WriteFile* indicate whether the read operation or write operation will be measured. The *SegmentCount* decides the number of datasets in the file. The *BlockSize* represents the size of the subdomain of the dataset stored on each processor. The *TransferSize* is the I/O transaction size used to transfer data from memory to the data file, which may require multiple

transfers per segment to copy the entire “*BlockSize*” to the data file. The *NumTasks* is the number of processors participated in the I/O operations.

## 2.3 JUBE

The JUBE benchmarking environment [3] provides a script based framework to easily create benchmark sets, run those sets on different computer systems and evaluate the results. It is actively developed by the *Jülich Supercomputing Centre of Forschungszentrum Jülich*, Germany. In order to alleviate these problems JUBE helps performing and analyzing benchmarks in a systematic way. It allows custom work flows to be able to adapt to new platforms. The platform file for ENEA CRESCO is XML file oriented to specific benchmark platform, taking into account the command set to submit a bunch of conditioned jobs with different parameters.

## 3 PDI in IOR development

The integration of PDI in IOR has been developed by means a PDI plugin, taking into account POSIX as backend interface with the GPFS support. A POSIX *usercode* plugin has been developed in order to decouple the posix system functions: *read()/write()*, from IOR function calls, replacing them with PDI API. The POSIX plugin developed: *xfer\_pdi()*, is depicted in the Listing 1 and implements the *read* and *write* functionalities of IOR PDI interface using POSIX I/O system calls. The PDI interface has been developed in the source code of IOR *tarball* v.3.30: *aiori-PDI.c*, with the function calls defined by the following C-language prototypes:

- `void *PDI_Create (char *testFileName, IOR_param_t * param);`
- `int PDI_Mknod (char *testFileName);`
- `void *PDI_Open (char *testFileName, IOR_param_t * param);`
- `static void PDI_Fsync (void *fd, IOR_param_t * param);`
- `static IOR_offset_t PDI_Xfer (int, void *, IOR_size_t *, IOR_offset_t, IOR_param_t *);`
- `IOR_offset_t PDI_GetFileSize (IOR_param_t * test, MPI_Comm testComm, char *testFileName);`
- `PDI_Delete (char *testFileName, IOR_param_t * param);`
- `void PDI_Close (void *fd, IOR_param_t * param);`
- `option_help * PDI_options (void ** init_backend_options, void * init_values)`

```

void xfer_pdi()
{
    // arguments of the function
    int* access; PDI_access("access", (void**)&access, PDI_IN);
    long* length; PDI_access("ts", (void**)&length, PDI_IN);
    long* offset; PDI_access("offset", (void**)&offset, PDI_IN);
    int* fd; PDI_access("file", (void**)&fd, PDI_IN);
    long l=*length;
    char* pptr;
    if (*access == READ) { //read
        PDI_access("value", (void**)&pptr, PDI_OUT);
    }
}

```

```

} else { // write
    PDI_access("value", (void*)&pptr, PDI_IN);
}
long long rc;
long long remaining = (long long)*length;
// function body
if (*access == READ) {
    if (verbose >= VERBOSE_4) {
        fprintf(stdout, "task %d reading from offset %lld\n", rank, *offset + length-remaining);
    }
    rc = read(*fd, pptr, *length);
} else {
    if (verbose >= VERBOSE_4) {
        fprintf(stdout, "task %d writing to offset %lld\n", rank, *offset + length - remaining);
    }
    rc = write(*fd, pptr, *length);
}
// release all metadata
PDI_release("access");
PDI_release("file");
PDI_release("ts");
PDI_release("offset");
PDI_release("value");
// update return value
PDI_expose("return", &rc, PDI_OUT);
}

```

Listing 1: PDI plugin for Posix interface

The PDI APIs use a *Yaml* file to handle data structures and plugins. The *Yaml* file of PDI in IOR is listed in the Listing 2. The plugin *xfer\_pdi* needs of some metadata in order to read and write a *buffer* with size *ts* in the *offset* of a file handled by a pointer to integer. The environment variable `YAML_PATH_FILENAME` set the absolute path and the name of the *Yaml* configuration file.

```

metadata:
  ts: int64
  access: int
  file: int
  offset: int64
  return: int64
data:
  buffer: {type: array, subtype: char, size: $ts}
plugins:
  user_code:
    on_data:
      buffer:
        xfer_pdi: value: $buffer

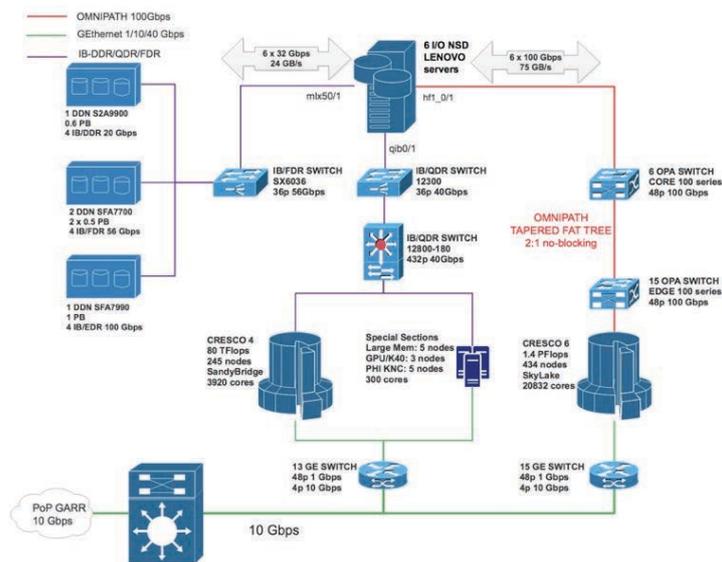
```

Listing 2: *Yaml* file configuration for PDI with Posix plugin.

## 4 CRESCO Data Center

A full environment of development and test of the activity has been installed in ENEA HPC CRESCO6. It is a high performance computing (HPC) system based on Linux Centos 7.3. The interconnect is based on Intel Omni-Path 100 Gbps. The main challenge, to deploy CRESCO6 HPC cluster into the CRESCO data center infrastructure, has been to design and implement a layout on which: the old InfiniBand QDR (40 Gbps) fabric of the CRESCO 4/5 clusters, including high performance storages DDN equipped with InfiniBand FDR (56Gbps), the new Omni-Path (100 Gbps) fabric of CRESCO6 cluster and the high performance filesystem IBM Spectrum Scale (GPFS), work all together into a single infrastructure. The main hardware components and the layout of the multi-fabric solution are depicted in the fig.2:

**Fig.2: CRESCO Data Center Infrastructure layout.**

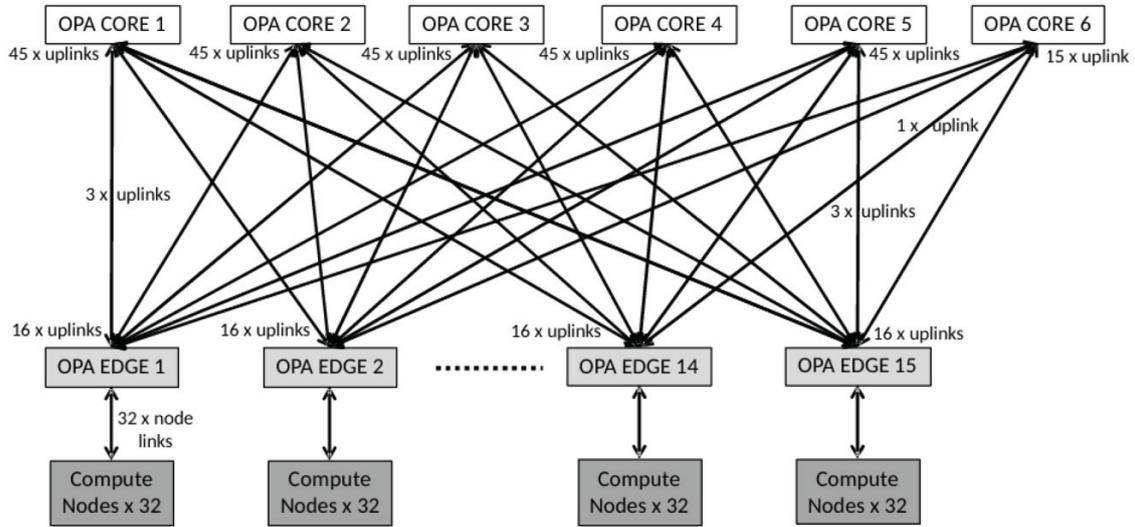


CRESCO6 consists of 434 nodes for a total of 20832 cores. It is based on Lenovo ThinkSystem SD530 platform, an ultra-dense and economical two-socket server in a 0.5 U rack form factor inserted in a 2U four-mode enclosure. Each node is equipped with:

- 2 Intel Xeon Platinum 8160 CPUs, each with 24 cores with a clock frequency of 2.1 GHz;
- A RAM of 192 GB, corresponding to 4 GB/core;
- A low-latency Intel Omni-Path 100 Series Single-port PCIe 3.0 x16 HFA network interface.

The nodes are interconnected by an Intel Omni-Path network with 21 Intel Edge switches 100 series of 48 ports each, bandwidth equal to 100 Gbps, latency equal to 100 ns. The connections between the nodes have 2 tier 2:1 no-blocking tapered fat-tree topology (Fig.3). The consumption of electrical power during massive computing workloads amounts to 190 kW.

**Fig.3: CRESCO Data Center Infrastructure layout.**



The storage systems are based on DDN solutions with a network based on InfiniBand. The technical specifications are as follow:

- 1 x DDN S2A9900, size 600 TB raw, IB/QDR, I/O bandwidth:6 GB/s;
- 2 x DDN SFA7700, size 1 PB raw, IB/FDR, I/O bandwidth: 11 GB/s;
- 1 x DDN SFA7990, size 1 PB raw, IB/EDR, I/O bandwidth:20 GB/s.

The high performance file system is based on GPFS Spectrun Scale with 6 NSD nodes able to provide parallel I/O. The hardware details of the NSD servers and the fabric switch which connects to the storage systems are:

- 6 x NSD servers for GPFS based on 2U Lenovo ThinkSystem SR650, 96 GB RAM, 2 IntelXeon Gold 5518 CPU @2.3 GHz 12 cores, 1 QLogic IBA7322 QDR InfiniBand HCA (rev02), 1 Mellanox MCX454A-FCAT DUAL FDR InfiniBand HCA, 1 Intel Omni-Path HFISilicon 100 Series;
- 1 x 1U switch Mellanox MSX6036F - 36 ports FDR InfiniBand for linking the NSD servers tothe storage DDN systems;
- 1 x 1U switch Qlogic 12300 – 36 ports QDR InfiniBand for linking the NSD servers to the CRESCO 4/5 clusters.

The main software stack of the CRESCO multi-fabric systems includes: operating systems, InfiniBand and Omni-Path fabric, RDMA, SRP, Multipath packages and finally IBM Spectrum Scale GPFS.

## SOFTWARE PLATFORM

The base software suite available in CRESCO Data Center and used in the development and test activity of EOCoE II are:

*Operating System* : Linux Centos 7.3 – kernel v.3.10.0-514.26.2.el7.x86\_64

*Distributed Filesystem*: Andrew File System – openAFS v.1.6.22

*High Performance Filesystem*: IBM Spectrum Scale GPFS – v.4.2.3

*Resources Management System*: IBM Spectrum LFS – v.9.1

*Compilers*: GNU GCC – v.7.3.0

*MPI library*: OpenMPI – v.3.1.2

*IO Library*: HDF5 – v.1.8.21 built with GNU GCC v.7.3.0

*Build tool*: Cmake – v.3.15

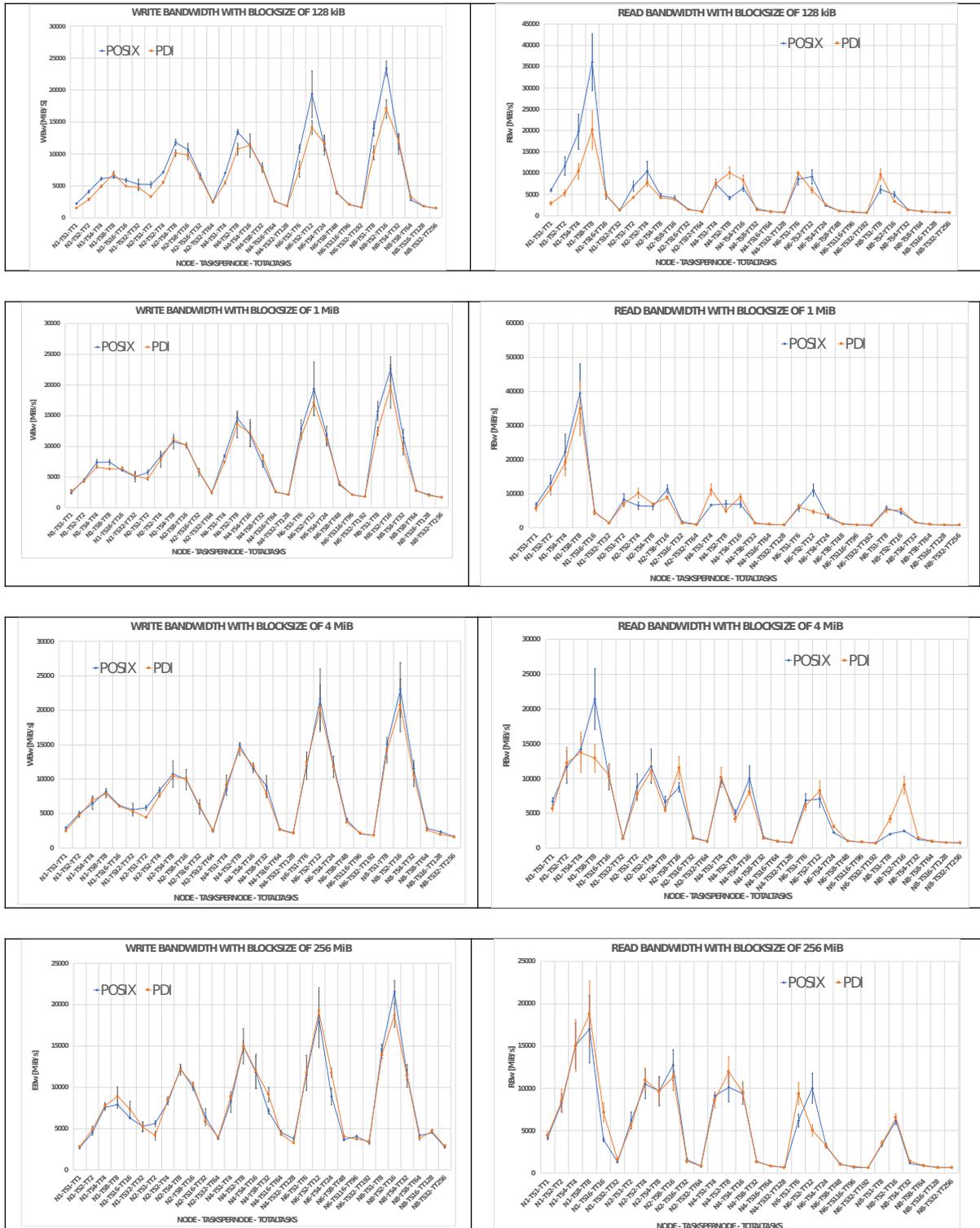
## **5 Benchmarks**

In order to run IOR benchmarks on the CRESCO6 system, a JUBE configuration of the IOR benchmarks has been deployed. JUBE has been configured to run IOR on distributed mode, with sequential jobs scheduled within the LSF platform resource manager.

The results of the IOR benchmark integrated with the PDI library for the POSIX interface have been compared with the case without the PDI layer to measure not only performances but also the overhead introduced by PDI. The experiments have been conducted for the following configurations:

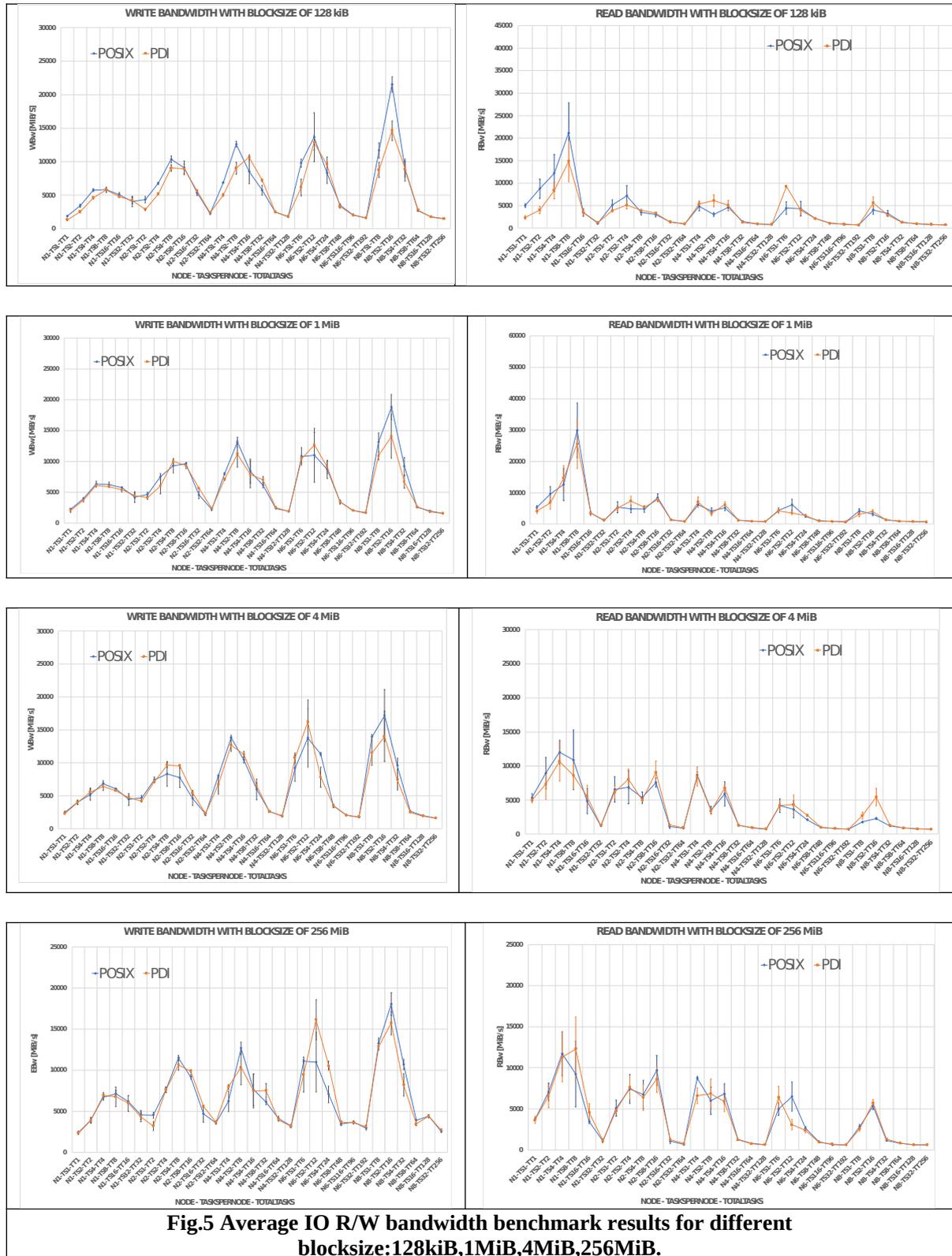
- number of nodes (N) = 1, 2, 4, 6, 8
- task per node (TS) = 1, 2, 4, 8, 16, 32

The total number of tasks (TT) for each combination is thus  $(TT) = (N) \cdot (TS)$ . For each combination (TT) the block size has been set to the four distinct values: 128 kiB, 1 MiB, 4MiB, 256 MiB to explore the effect of increasing I/O buffers. IOR strategy is the one-file per task, it means each task reads/writes its own file. Each task reads/writes an aggregate file size of 512 MiB and the four distinct value of block size correspond to the four numbers of segments: 4096, 512, 128, 2. The JUBE *xml* file: *iorpdi.xml* that provides to run the benchmark sessions over ENEA HPC system CRESCO6 is shown in the Annex V. Four benchmark sessions were carried out: three of them during the stop of production of ENEA HPC CRESCO6 cluster, when the IO bandwidth of the infrastructure is noiseless caused by users jobs. As described before the benchmarks have been performed with JUBE submitting 960 serial jobs for a total of about 1858 core-hours. The plots of the R/W bandwidth measured during the benchmarks are depicted in the fig.4. The better results of the four benchmark sessions are plotted as MAX values, whilst the error bars are the variance of the four benchmark results.



**Fig.4 Max IO R/W bandwidth benchmark results for different blocksize:128kiB,1MiB,4MiB,256MiB.**

In order to have a confirm on the statistic of the results, the plots of the average values measured in the four benchmark sessions are depicted in the fig.5.



An other data view of the benchmark results carried out in the four sessions, allows to make an evidence the overhead produced by PDI. In the following figs.6-7, the R/W bandwidth measured during the benchmark sessions as better performance of IO in both: maximum and average value.

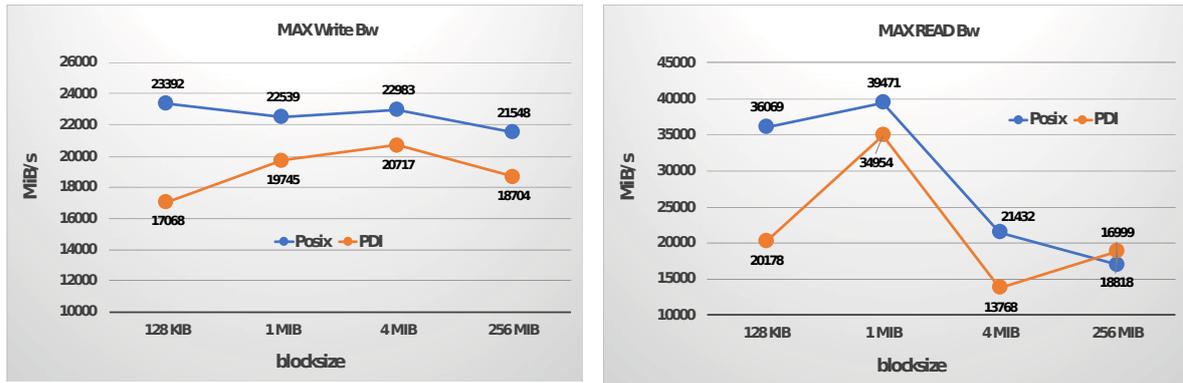


Fig.6 Better IO R/W bandwidth benchmark results as Max value vs. blocksize.

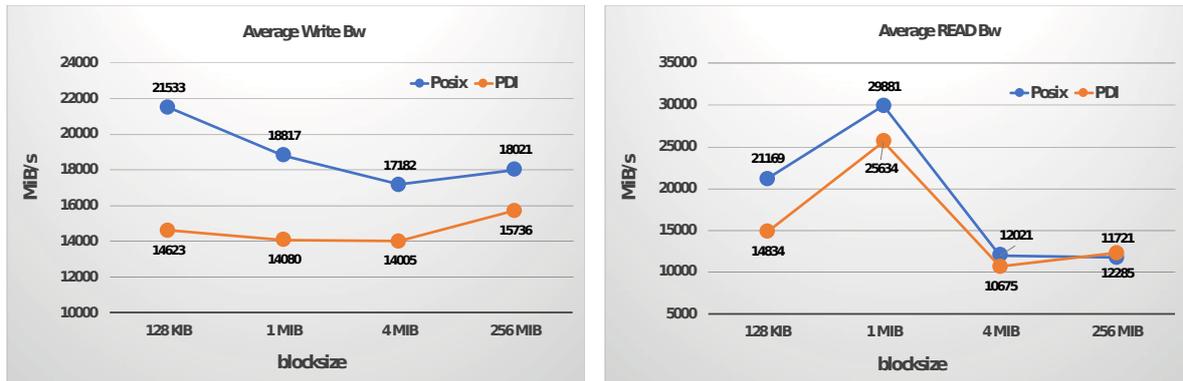


Fig.7 Better IO R/W bandwidth benchmark results as Average value vs. blocksize.

## Conclusions

The report of ENEA activity in EOCoE II WP4 has been provided. The development and test of the PDI integrated in IOR has been performed as request of the project team in order to evaluate the performance of PDI used in the IO of simulation codes.

The benchmark results measured during the sessions have been analysed and they have shown a significant overhead produced by PDI API. The overhead is measurable when many calls to PDI function: *PDI\_expose()* and *PDI\_access()* are done in the for-cycle performed in the case of small blocksize (128 kiB) in both read/write mode. Obviously the overhead is smaller when large data buffers are involved in the case of large blocksize (4 and 256 MiB).

## References

- [1] C.Roussel, K.Keller, M.Gaalich, L.B.Gomez, J.Bigot. *PDI, an approach to decouple I/O concerns from high performance simulation codes*. (2017) from <https://hal.archives-ouvertes.fr/hal-01587075/>
- [2] H.Shan, J.Shalf. *Using IOR to Analyze the I/O performance for HPC Platforms*. (2007) Cray User Group Conference (CUG'07)
- [3] Galonska, A.; Frings, W.; Gibbon, P.; Borodin, D.; Kirschner, A. JuBE-based Automatic Testing and Performance Measurement System for Fusion Codes. E.H. D'Hollander, G.R. Joubert, David Padua, Frans Peters, Mark Sawyer, IOS Press, 2012, Advances in Parallel Computing, Vol. 22. - 978-1-61499-040-6. - S. 465 – 472.

# BIAXIAL STRAIN EFFECTS ON THE ELECTRON-PHONON COUPLING IN IRON-SELENIDE: A FIRST PRINCIPLES STUDY

Rui Zhang<sup>1</sup> and Gianluca De Marzi<sup>2\*</sup>

<sup>1</sup>*College of Civil Engineering and Mechanics, Lanzhou University,  
222 South Tianshui Road, 730000-Lanzhou, China*

<sup>2</sup>*ENEA, FSN-COND, Centro Ricerche Frascati  
Via Enrico Fermi 45, 00044-Frascati, Italy*

**ABSTRACT.** We report on the enhancement of superconductivity in iron-selenide FeSe induced by in-plane biaxial compressive strain. It is found that, due to compressive strain, the superconducting transition temperature is increased significantly when the Fe spins are polarized in the antiferromagnetic checkerboard configuration. Furthermore, the strain dependence of the electronic and vibrational is discussed in detail, as well as the behaviour of the electron-phonon coupling constant at various levels of compressive strain. The research here presented provides a way to modulate superconductivity in layered iron compounds by means of biaxial strain, and can be regarded as a sound basis for further theoretical studies on different iron-based compounds.

## 1 Introduction

The binary compound FeSe (iron selenide) has the simplest crystal structure among other iron-based superconductor families, sharing with them the same iron pnictide layer structure, but without the presence of separating layers. In FeSe, the superconducting critical temperature,  $T_c$ , can be increased from 8.5 K to 36.7 K under an applied pressure of 8.9 GPa [1]. Furthermore, several experimental works [1-4] have unambiguously unveiled the existence of a universal correspondence between the parameters of the crystal structure - such as the anion height and the bond angle - and the value of the critical temperature.

The effect of tensile strain on the superconductivity in FeSe films has been experimentally investigated by Nie *et al.* [5]. Superconducting transitions have been observed in unstrained FeSe films with  $T_c$  close to the bulk value. However, no sign of superconductivity was observed in FeSe films under tensile strain, thus clearly demonstrating that tensile strain suppresses superconductivity in FeSe films. The influence of strain on the structural and electronic properties of superconducting FeSe has been studied *ab initio* by Winiarski *et al.* The authors found that the  $c$ -axis-strained FeSe inhibits the Fermi surface nesting, which in turn enhances spin fluctuations and, hence, superconductivity. In the  $ab$ -plane, the nesting of the Fermi surface is slightly weakened by compressive strain; conversely, it is completely destroyed by the application of a tensile strain [6].

The magnetic ground state of FeSe is the block-checkerboard, which turns into the collinear stripe phase under in-plane tensile strain [7]. By angle-resolved photoemission spectroscopy on tensile- or compressive strained and strain-free FeSe, Phan *et al.* have experimentally shown that the in-plane strain causes a marked change in the Fermi-surface volume, eventually leading to a change in  $T_c$  [8]. Other works confirm the correlation between in-plane strain and  $T_c$ : in epitaxial FeSe thin films with various magnitude of lattice strain, from tensile to compressive, it was confirmed that  $T_c$  is strain-sensitive [9]; in particular,  $T_c$  reaches a value of 12 K for the films with compressive strain, whereas

---

\* Corresponding author. E-mail: gianluca.demarzi@enea.it.

the superconductivity disappears for the films with large tensile strain. Although some authors [10] believe that the modest electron-phonon coupling (EPC) in these materials cannot explain the superconductivity in bulk iron-based superconductors, other studies raise concerns about whether the role of phonon can be completely ignored [11]. In particular, it is found that strained FeSe/SrTiO<sub>3</sub> interfaces [12] and/or heavy electron doping [13] can lead to a significant enhancement of the EPC, and hence of  $T_c$ . The enhancement of superconductivity induced by in-plane biaxial compressive strain has been experimentally observed in Fe-Se-Te thin crystals [14-17]. Due to compressive strain,  $T_c$  is increased by  $\approx 14\%$  in FeTe<sub>0.5</sub>Se<sub>0.5</sub> [15] and by  $\approx 30\%$ – $40\%$  in FeSe [16].

In this paper, in order to clarify the microscopic origin of the strain dependence of the superconducting properties in iron-based systems, we investigated the electronic and phonon properties of the FeSe binary compound from an *ab-initio* perspective.

The electronic band structures, phonon dispersion curves and EPC are calculated at different biaxial strain values, in both the non-magnetic (NM) and checkerboard antiferromagnetic (AF) phases.

To our knowledge, this is the first theoretical study on the influence of biaxial strain on the superconducting properties of FeSe.

## 2 Model and method

Computations have been carried out by way of the ENEA-CRESCO computational facility.

The Density Functional Theory (DFT) scheme here employed adopts a Local Density Approximation (LDA) of the exchange-correlation potential [18]. We used the pseudopotential method, based on plane-waves and Projector-Augmented Waves (PAW), as implemented in the QUANTUM-ESPRESSO (QE) package [19]. The electron-ion interactions have been modelled with ultra-soft pseudopotentials (US-PP) in the context of a plane wave expansion basis set. US-PPs have allowed the usage of an energy cut-off of 60 Ry for the wave functions and 600 Ry for the electron density.

For the electronic structure calculation, the Brillouin-zone (BZ) integrations are performed over a  $12 \times 12 \times 8$  wavevector  $k$ -space grid. The lattice parameters and atomic positions of the unstrained FeSe have been optimized starting from the experimental values, by means of the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm. Then, for each given biaxial strain, and with the in-plane lattice parameter fixed, all the atomic positions have been relaxed in the  $z$ -direction by optimization of the total energy for both the NM and checkerboard AF states, respectively.

Lattice dynamical calculations have been performed within the framework of the self-consistent density functional perturbation theory (DFPT) [20], in which the dynamical matrices are calculated by sampling the BZ with a  $2 \times 2 \times 2$   $q$ -space grid. A denser grid of  $k$ -points ( $24 \times 24 \times 24$ ) has been used in order to determine the EPC parameter  $\lambda$ , which is computed through a BZ average of the mode-resolved coupling strengths  $\lambda_{qj}$  ( $j$ -th mode with momentum  $q$ ) [21-23]:

$$\lambda = \frac{1}{N_q} \sum_{qj} \lambda_{qj} = 2 \int \frac{\alpha^2 F(\omega)}{\omega} d\omega \quad (1)$$

where  $N_q$  represents the total number of  $q$  points in the fine phonon mesh, whereas  $\alpha^2 F(\omega)$  is the Eliashberg spectral function [24].

FeSe is a multigap superconductor with two distinct superconducting gaps  $\Delta_1 \approx 3.5$  meV and  $\Delta_2 \approx 2.5$  meV [25]. Remarkably, the Fermi energies are comparable to the superconducting gaps;  $\Delta/\varepsilon_F$  is  $\sim 0.3$  and  $\sim 1$  for hole and electron bands, respectively [25]. These large  $\Delta/\varepsilon_F$  values indicate that FeSe is in the BCS-BEC crossover regime. Therefore, bearing in mind the limits of the application of the McMillan formalism to the iron-based systems, we still utilize the McMillan-Allen-Dynes formula to determine the superconducting  $T_c$  [22-24]:

$$T_c = \frac{\omega_{\ln}}{1.2} \exp \left[ \frac{-1.04(1+\lambda)}{\lambda - \mu^*(1+0.62\lambda)} \right] \quad (2)$$

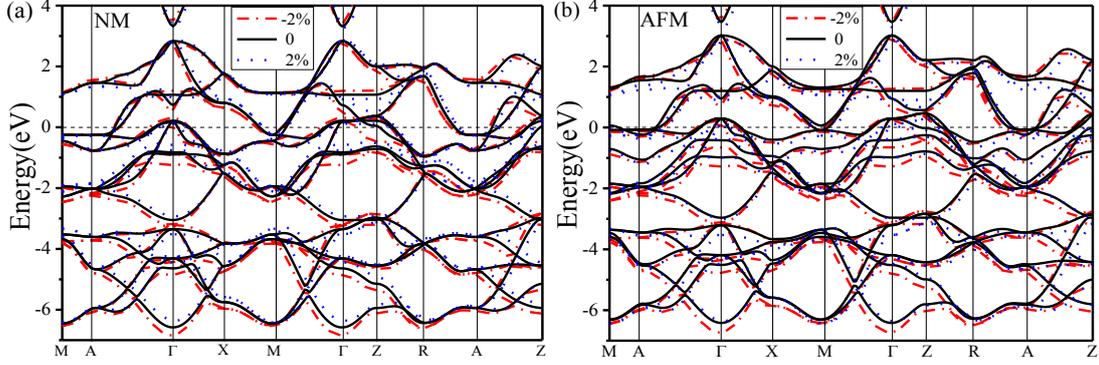
where  $\omega_{\ln}$  is the logarithmic average of the phonon frequency, which can be evaluated through [25]:

$$\omega_{\ln} = \exp \left[ \frac{2}{\lambda} \int \frac{d\omega}{\omega} \alpha^2 F(\omega) \ln \omega \right] \quad (3)$$

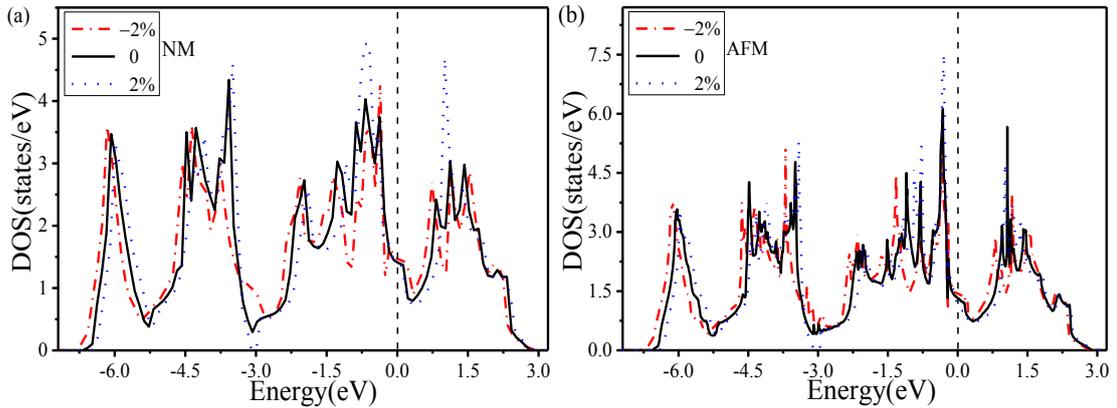
whereas  $\mu^*$  is the effective Coulomb repulsion, which is related to  $N(\varepsilon_F)$  through the Bennemann and Garland empirical expression  $\mu^* = 0.26 \cdot N(\varepsilon_F) / [1 + N(\varepsilon_F)]$  [26].

### 3 Results and discussion

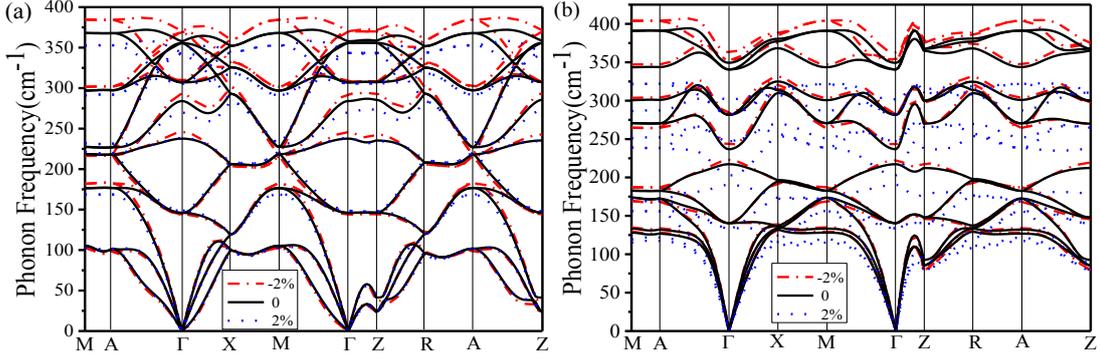
#### 3.1 Electronic structures and phonon dispersion curves



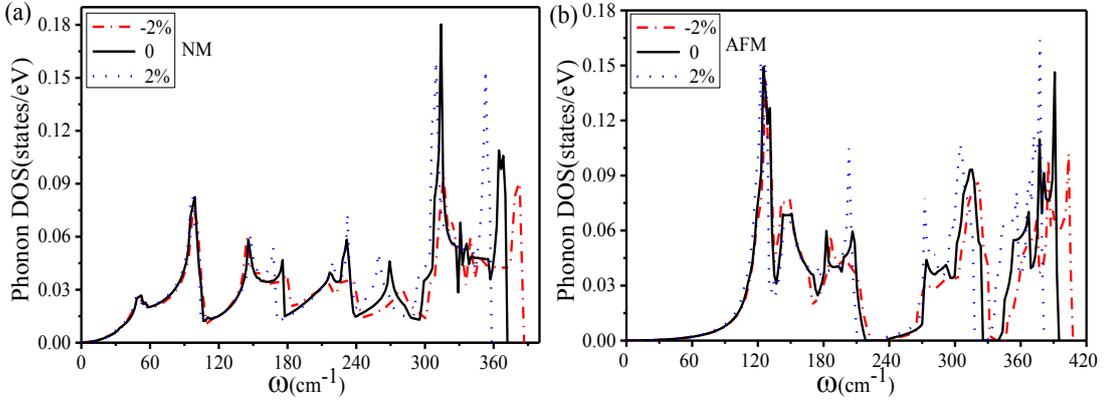
**Fig. 1:** Electronic band structures of bulk FeSe calculated at different biaxial strain values of -2% (red dash-dotted line), 0% (black line), and 2% (blue dotted line): (a) non-magnetic phase (NM); (b) checkerboard antiferromagnetic phase (AFM). The Fermi energy is set to zero.



**Fig. 2:** Electronic density of states (eDOS) of bulk FeSe calculated at different biaxial strain values of -2% (red dash-dotted line), 0% (black line), and 2% (blue dotted line): (a) non-magnetic phase (NM); (b) checkerboard antiferromagnetic phase (AFM). The Fermi energy is set to zero.



**Fig. 3:** Phonon dispersion curves of bulk FeSe calculated at different biaxial strain values of -2% (red dash-dotted line), 0% (black line), and 2% (blue dotted line): (a) non-magnetic phase (NM); (b) checkerboard antiferromagnetic phase (AFM).



**Fig. 4:** Phonon Density of States (pDOS) of bulk FeSe calculated at different biaxial strain values of -2% (red dash-dotted line), 0% (black line), and 2% (blue dotted line): (a) non-magnetic phase (NM); (b) checkerboard antiferromagnetic phase (AFM).

Using the optimized atomic coordinates, we firstly calculate the electronic band structures of FeSe for three different values of biaxial strain (Fig. 1). For the strain-free FeSe, there are five bands crossing the Fermi level. Across the Fermi level, those bands exhibit a downward shift as the strain is increased from -2% to 2% for both NM and AFM phases. The opposite effect occurs for the bands at lower energies around -7 eV.

As shown in Figure 2, The electronic density of states (eDOS) at the Fermi level is slightly reduced by the increasing strain, due to the downward shift of weakly dispersive energy band along  $\Gamma$ -Z near the Fermi level. Moreover, several peaks in the eDOS are clearly enhanced as the strain is increased from -2% to 2%.

Secondly, the lattice dynamics and EPC of FeSe under biaxial strain have been explored. In Figure 3, the phonon dispersion curves are plotted along selected high-symmetry directions in the BZ, for three different representative strain values. Notably, there are no negative phonon frequencies in any of the phonon dispersion curves, thus indicating that the strained FeSe structure is dynamically stable.

As the biaxial strain is increased from -2% to 2%, the high-frequency phonon modes soften, whereas the low-frequency modes remain unaffected. This behavior is also highlighted by the evolution of the phonon DOS (pDOS) with strain, plotted in Figure 4. It is also worth noticing that

the effect of the magnetic ordering hardens the higher frequency phonons, which reach up to 400  $\text{cm}^{-1}$ .

### 3.2 Electron-phonon coupling

In this last section we discuss the behavior of the EPC constant as a function of the biaxial strain, together with its implications on the superconducting properties of FeSe.

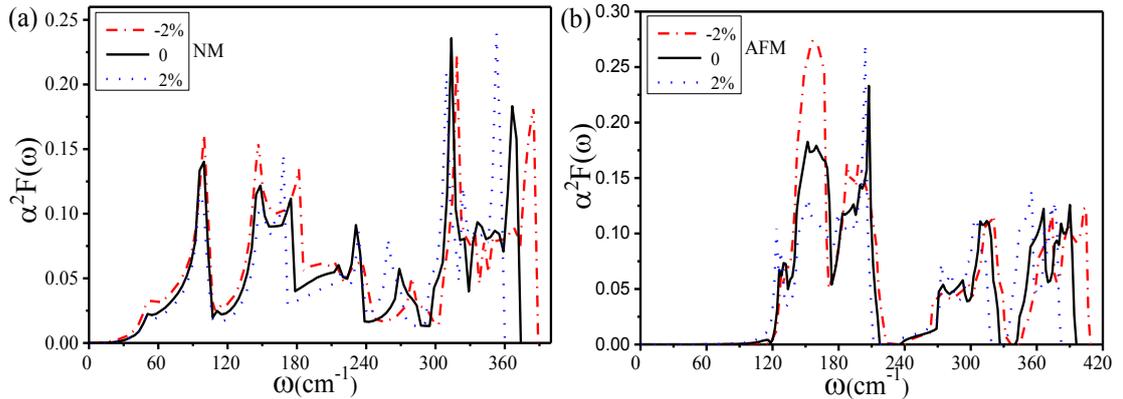
In Figure 5, the Eliashberg function  $\alpha^2F(\omega)$  is plotted as a function of phonon frequency, for three different strain values. As it can be easily inferred, although the peak positions of  $\alpha^2F(\omega)$  coincide with those of pDOS, their peak strengths are highly affected by strain. This implies that the low-frequency EPC matrix elements strongly depend on biaxial strain.

By proper integration of  $\alpha^2F(\omega)$ , several superconducting related quantities (namely: the EPC constant  $\lambda$ ,  $\omega_m$ , and  $T_c$ ) can be evaluated. The effective electron-electron repulsion  $\mu^*$  is evaluated by means of the Bennemann and Garland expression [26]. All these quantities are listed in Table I.

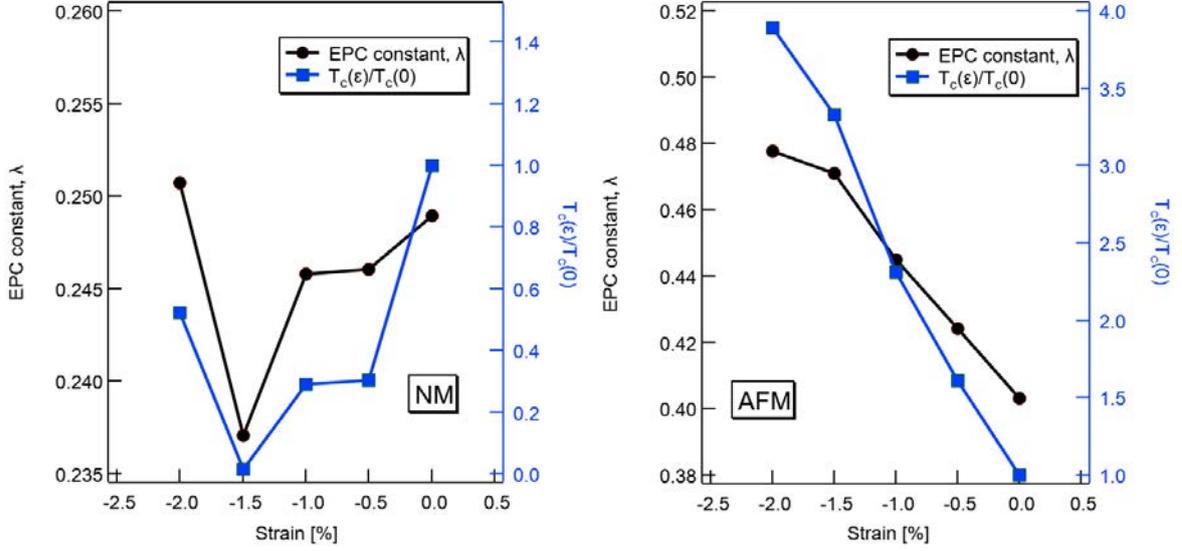
In unstrained FeSe, we have found that the EPC constant  $\lambda$  increases from  $\approx 0.25$  for non-spin-resolved FeSe (NM phase) to  $\approx 0.40$  for the checkerboard spin-resolved configuration (AFM phase), in agreement with previously published *ab-initio* results [27]. It should be pointed out that the value of  $\lambda$  is slightly higher than that calculated within the Generalized Gradient Approximation (GGA) ( $\lambda = 0.14\text{-}0.18$  [12, 13]). This is expected: in fact, LDA tends to underestimate the bond length, thus resulting in larger deformation potentials. The increase of the electron-phonon coupling due to magnetic ordering leads to a  $T_c$  of 0.17 K when iron moments are included in the calculation, and for practically zero values in the opposite case. The critical temperature based on the McMillan equation are not high enough to achieve consistency with the experimentally measured superconducting transition of 8 K for FeSe, even for the checkerboard spin-resolved case. This suggests that the electron-phonon mechanism alone cannot explain superconductivity in Fe-Se systems.

With the application of a compressive biaxial strain, a rise of  $\lambda$  can be clearly seen in the AFM phase (Fig. 6b). At -2% strain,  $\lambda$  is enhanced by about 18%, in comparison with the strain free case. However, even at this strain level FeSe can only superconduct at a rather low temperature of  $\approx 0.7$  K, which is far from the experimentally observed  $T_c$ .

Nevertheless, the behavior of the estimated  $T_c$  is in qualitative agreement with the experimental findings [8], thus implying that EPC might play a crucial role in biaxially strained iron-based superconductors.



**Fig. 5:** Eliashberg spectral function  $\alpha^2F(\omega)$  bulk FeSe calculated at different biaxial strain values of -2% (red dash-dotted line), 0% (black line), and 2% (blue dotted line): (a) non-magnetic phase (NM); (b) checkerboard antiferromagnetic phase (AFM).



**Fig. 6:** The EPC constant  $\lambda$  and  $T_c$  of FeSe as a function of the biaxial strain for non-spin-resolved (left) and spin-resolved configurations (right).

**Table I:** Summary of main calculated results for FeSe under biaxial strain in the non-spin-resolved (NM phase) and spin-resolved configurations (AFM phase).

Strain [%]	$N(\epsilon_F)$ [states/eV]	$\lambda$	$\omega_{ln}$ [K]	$\mu^*$	$T_c$ [K]
NM					
-2	1.55	0.25	197	0.158	$8.67 \times 10^{-7}$
0	1.47	0.25	206	0.155	$1.66 \times 10^{-6}$
2	1.38	0.23	150	0.151	$1.08 \times 10^{-7}$
AFM					
-2	1.55	0.48	221	0.159	0.66
0	1.51	0.40	229	0.156	0.17
2	1.51	0.31	209	0.156	$2.78 \times 10^{-3}$

## 4 Summary

In summary, we investigated the enhancement of superconductivity in the spin-resolved FeSe superconductor subjected to biaxial compressive strain. The electronic band structures, phonon dispersions curves and electron-phonon coupling (EPC) constant were computed at different values of biaxial strain. The critical temperature,  $T_c$ , has been evaluated using the Allen-Dynes modification of the McMillan formula, whereas the effective Coulomb repulsion was estimated using the

Bennemann-Garland expression. By applying a biaxial compression of -2%, we found that the EPC constant can increase up to about 390% with respect to its strain-free value, as the Fe spins align in the magnetic checkerboard configuration. Our calculations clarify that biaxial compressive strain can indeed remarkably enhance the EPC in FeSe. However, such enhancement is still too weak to fully explain the high  $T_c$  values observed in bulk FeSe, thus implying that other mechanisms - along with the electron-phonon coupling - needs to be considered in iron-based systems.

## Acknowledgments

This work was supported by the National Education Fund Management Committee (China Scholarship Council Grant No. 201806180042) and by Professor Xingzhe Wang (College of Civil Engineering and Mechanics, Lanzhou University). Computational resources have been provided by the ENEA Computational Research Centre on Complex systems.

## References

- [1] S. Medvedev, T. M. McQueen, I. A. Troyan, T. Palasyuk, M. I. Erements, R. J. Cava, S. Naghavi, F. Casper, V. Ksenofontov, G. Wortmann and C. Felser. Electronic and magnetic phase diagram of  $\beta$ -Fe<sub>1.01</sub>Se with superconductivity at 36.7 K under pressure. *Nat. Mater.* **8**, pp. 630-633, (2009).
- [2] Y. Mizuguchi, Y. Hara, K. Deguchi, S. Tsuda, T. Yamaguchi, K. Takeda, H. Kotegawa, H. Tou and Y. Takano. Anion height dependence of  $T_c$  for the Fe-based superconductor. *Supercond. Sci. Technol.* **23**, 054013 (5 pp.), (2010).
- [3] C. H. Lee, K. Kihou, A. Iyo, H. Kito, P. M. Shirage and H. Eisaki. Relationship between crystal structure and superconductivity in iron-based superconductors. *Solid State Commun.* **152**, pp. 644-648, (2012).
- [4] X. Chen, P. Dai, D. Feng, T. Xiang and Fu-Chun Zhang. Iron-based high transition temperature superconductors. *Natl. Sci. Rev.* **3**, pp. 371-395, (2014).
- [5] Y. F. Nie, E. Brahim, J. I. Budnick, W. A. Hines, M. Jain, and B. O. Wells Y. Nie. Suppression of superconductivity in FeSe films under tensile strain. *Appl. Phys. Lett.* **94**, 242505 (3 pp.), (2009).
- [6] M. J. Winiarski, M. Samsel-Czekala and A. Ciechan. Strain effects on the electronic structure of the iron selenide superconductor. *EPL* **100** 47005 (5 pp.), (2012).
- [7] C. Tresca, F. Ricci and G. Profeta. Strain effects in monolayer iron-chalcogenide superconductors. *2D Mater.* **2**, 015001 (6 pp.), (2015).
- [8] G. N. Phan, K. Nakayama, K. Sugawara, T. Sato, T. Urata, Y. Tanabe, K. Tanigaki, F. Nabeshima, Y. Imai, A. Maeda and T. Takahashi. Effects of strain on the electronic structure, superconductivity, and nematicity in FeSe studied by angle-resolved photoemission spectroscopy. *Phys. Rev. B* **95**, 224507 (6 pp.), (2017).
- [9] M. Kawai, F. Nabeshima and A. Maeda. Transport properties of FeSe epitaxial thin films under in-plane strain, *J. Phys.: Conf. Series* **1054**, 012023 (7 pp.), (2018).
- [10] A. Subedi, L. Zhang, D. J. Singh and M. H. Du. Density functional study of FeS, FeSe, and FeTe: Electronic structure, magnetism, phonons, and superconductivity. *Phys. Rev. B* **78**, 134514 (6 pp.), (2008).
- [11] Y.-Y. Xiang, F. Wang, D. Wang, Q.-H. Wang and D.-H. Lee. High-temperature superconductivity at the FeSe/SrTiO<sub>3</sub> interface. *Phys. Rev. B* **86**, 134508 (9 pp.), (2012).
- [12] B. Li, Z. W. Xing, G. Q. Huang and D. Y. Xing. Electron-phonon coupling enhanced by the FeSe/SrTiO<sub>3</sub> interface. *J. Appl. Phys.* **115**, 193907 (6 pp.), (2014).
- [13] D.-D. Zheng, M. Gao and X.-W. Yan. Electron-phonon coupling in heavily electron-doped bulk FeSe: a first-principles investigation. *Appl. Phys. Express* **12**, 013003 (4 pp.), (2019).
- [14] E. Bellingeri, I. Pallecchi, R. Buzio, A. Gerbi, D. Marrè, M. R. Cimberle, M. Tropeano, M. Putti, A. Palenzona, S. Kaciulis and C. Ferdeghini. Critical temperature enhancement by biaxial compressive strain in FeSe<sub>0.5</sub>Te<sub>0.5</sub> Thin Films. *J. Supercond. Nov. Magn.* **24**, pp. 35-41, (2011).

- [15] W. Wang *et al.* Scotch tape induced strains for enhancing superconductivity of FeSe<sub>0.5</sub>Te<sub>0.5</sub> single crystals. *Appl. Phys. Lett.* **105**, 232602 (4 pp.), (2014).
- [16] W. Wang, X. Wang, L. Zhang, J. Yang, X. Chen, Z. Zhang, M. Tian, Z. Yang and Y. Zhang. Scotch tape induced strains for structural variation of FeTe<sub>0.5</sub>Se<sub>0.5</sub> and Fe<sub>1.05</sub>Te single crystals. *AIP Advances* **6**, 025207 (5 pp.), (2016).
- [17] X. F. Wang, Z. T. Zhang, W.K. Wang, Y. H. Zhou, X. C. Kan, X. L. Chen, C. C. Gu, L. Zhang, L. Pi, Z. R. Yang and Y. H. Zhang. Enhancement of superconductivity in FeSe thin crystals induced by biaxial compressive strain. *Physica C supercond.* **537**, pp. 1-4, (2017).
- [18] J. P. Perdew and Y. Wang. *Phys Rev B* **45**, pp. 13244-13249, (1992).
- [19] P. Giannozzi *et al.* QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials. *J. Phys.: Cond. Mat.* **21**, 395502 (19 pp.), (2009).
- [20] S. Baroni, S. de Gironcoli, A. Dal Corso and P. Giannozzi. Phonons and related crystal properties from density-functional perturbation theory. *Rev. Mod. Phys.* **73**, pp. 515-562, (2001).
- [21] W. L. McMillan. Transition temperature of strong-coupled superconductors. *Phys. Rev.* **167**, pp. 331-344, (1968).
- [22] P. B. Allen. Neutron Spectroscopy of Superconductors. *Phys. Rev. B* **6**, pp. 2577-2579, (1972).
- [23] P. B. Allen and R. C. Dynes. Transition temperature of strong-coupled superconductors reanalyzed. *Phys. Rev. B* **12**, pp. 905-922, (1975).
- [24] G. M. Éliashberg. Interactions between electrons and lattice vibrations in a superconductor. *Sov. Phys. JETP* **11**, pp. 696-702, (1960).
- [25] S. Kasahara *et al.* Giant superconducting fluctuations in the compensated semimetal FeSe at the BCS–BEC crossover. *Nat Commun* **7**, 12843 (7 pp.), (2016).
- [26] K. H. Bennemann and J. W. Garland. Theory of superconductivity in d-bands metals. *AIP Conf. Proc.* **4**, pp. 103-137 (1972).
- [27] T. Bazhiron and M. L. Cohen. Spin-resolved electron-phonon coupling in FeSe and KFe<sub>2</sub>Se<sub>2</sub>. *Phys. Rev B* **86**, 134517 (8 pp.), (2012).

# NEURAL NETWORKS FOR STATISTICAL PROCESS CONTROL OF MULTIPLE STREAM PROCESSES

Antonio Lepore<sup>1\*</sup>, Fiorenzo Ambrosino<sup>2</sup>, Gianluca Sposito<sup>1</sup>, and Biagio Palumbo<sup>1</sup>

<sup>1</sup>*Università di Napoli Federico II, Department of Industrial Engineering, P.le V. Tecchio 80, 80125 Naples, Italy*

<sup>2</sup>*ENEA, Italian National Agency for New Technologies Energy and Sustainable Economic Development, Energy Technologies Department, P.le E. Fermi 1, 80055 Portici, Italy*

**ABSTRACT.** For the rail industry, the coach temperature regulation is becoming a crucial task to improve passenger comfort. Its importance is due to government regulations that require rail industries to implement methods for controlling the Heating, Ventilation and Air Conditioning (HVAC) on-board systems of modern trains. Modern automated systems create a new dimension in the diagnostics and maintenance of complex systems in a new cost-effective and efficient way, and calls for new methods to deal with high-dimensional, high-correlated and heterogeneous operational data collected on board. A novel method based on the application of neural networks (NN) to pattern recognition in statistical quality control is introduced by means of real industrial multiple-stream process data in order to support prognosis of faults and to recover interpretability, which is typically overlooked by NN approaches.

## 1 Objectives

In the industry 4.0 environment, statistical process control (SPC), fault diagnosis and fault detection play a key role to avoid unplanned down-times and to reduce maintenance costs of modern automated systems. The ongoing digitalization creates a new dimension in the diagnosis, maintenance and operation of these systems in a new cost-effective and efficient manner [6]. The challenge is to turn high dimensional, high correlated and heterogeneous operational data affected by noise and environmental fluctuations into value. Machine learning methods (ML) are naturally prone to automatically identify patterns at the price of massive training reference datasets. In particular, supervised ML methods [5] are able to identify normal and out-of-normal behaviours but require a sufficiently large number of labeled training data. Ideally, the same number of training data should be available for each behaviour class. However, this is not feasible in practice because systems usually work under normal conditions for the most part of their life and faults are generally very rare so making real datasets usually unbalanced. On the other hand, unsupervised ML approaches do not require labeled training data and additional information on the input data, but have the drawback of showing lower performances in fault detection and classification. Hence, in what follows, semi-supervised method, based on autoencoders [4], will be explored to mitigate drawbacks of both supervised and unsupervised methods and will be embedded into a SPC scheme [7] [10].

---

\*Corresponding author. E-mail: antonio.lepore@unina.it.

## 2 Methods

As is known, autoencoders are typically useful for rare-event classification. The semi-supervised approach consists in training the model based on normal behaviour data, only. Then it can be likely exploited to alarm for any change in the correlation structure and to provide signaling of assignable causes. We define a time-point-wise reconstruction error between the input data and its reconstructed output, defined as the error sum of squares (SSE) over all the sensors. Then, according to the empirical distribution of the reconstruction error of the training set, we define an upper threshold, referred to as upper control limit (UCL), to alarm for possible anomalies when the observed point exceeds the UCL. The UCL is determined by using the results for the validation set. Formal definition of training, validation and test set can be found in [5]. This is based on the intuition that the more a point differs from the training set, the less the autoencoder is able to restore it [8].

Autoencoders are a type of artificial neural network (NN) used to learn from a set of data. Its goal is to copy the input to the output, getting the significant aspects of the data to be copied. Autoencoder has two main parts: an encoder, which maps the input  $x$  into the so-called *code* or *latent* representation  $h$ , represented by the function  $f = h(x)$  and a decoder, whose goal is to obtain a representation  $x'$  very similar to its original input. It is described by the function  $x' = g(f)$ . Actually, an autoencoder may be regarded as a traditional feed-forward, non-recurrent NN. In particular, the decoder is symmetric to the encoder in terms of layer structure and the output layer must have the same number of neurons as the input layer [4]. Since our goal is to capture into the code layer the main properties of the input for the fault detection analysis, we design a code layer with smaller dimension than the input's. An autoencoder whose code dimension is less than the input dimension is called undercomplete. To minimize the reconstruction error, autoencoders are trained by minimizing a *loss function*  $L(x, g(f(x))) = L(x, x') = \|x - x'\|^2$ , where  $L$  is the loss function penalizing  $g(f(x))$  for being dissimilar from the input vector  $x$ .  $L$  is assumed to be the mean squared error. In the general case, which considers more than one hidden layer, we have to compute all these values for any layer of the NN.

## 3 Real-case Study

In the rail industry the efficient temperature regulation is becoming a necessity in the face of strong competition and overcrowded carriages. Over the past few years, new European standards were developed for the thermal comfort, taking into account the different operating requirements of rail vehicles. In order to meet this standard and to improve the performance of HVAC systems, railway companies have been collecting and continuously storing multiple-stream data for each train coach. We focus attention on four temperature measurements available for each train coach: 1) interior temperature, 2) exterior temperature, 3) theoretical temperature to be achieved by the room air and 4) temperature supplied by the air conditioning. The average coach temperature has to recover within  $\pm 2$  degrees of the required interior car temperature, as established by the European standards [3]. When a fault occurs, this requirement is no longer satisfied. We monitor all the 24 signals available (four per each train's coach) to detect if one or more coaches/streams show unusual behaviours with respect to other coaches/streams or from the normal operating conditions, i.e., in which domain expert guarantees the system is under usual behaviour, usually referred to in SPC literature as *Phase I*. We consider a training set corresponding to the Phase I and a test set corresponding to a period that contains normal and fault data, which is accordingly referred to as *Phase II*. The parameters of the model are learned by back propagation and Adam optimization algorithm based on normal data [4]. We need to define an error threshold, an upper limit to distinguish between normal and anomalous behaviour. We define it as the

Table 1: Autoencoder hyperparameters. In bold, the hyperparameters, with their ranges, chosen through grid search tuning procedure.

Hyperparameter	Explored values	Chosen value
<b>Learning rate</b>	{0.1,0.01,0.001,0.0001}	0.01
Number of input layer nodes		24
<b>Number of hidden layer nodes</b>	{1,4,10,12,14,16,18}	12
Number of output layer nodes		24
<b>Activation function for hidden layer</b>	{relu, elu, tanh, sigmoid}	elu
<b>Activation function for output layer</b>	{relu, elu, tanh, sigmoid}	elu
Epochs		100
Batch size		128
<b>n-th percentile</b>	{95th, 97th, 99th}	99th-percentile

$n$ -th percentile of the error distribution of the training data as UCL with  $n$  controlling the false alarm rate. After setting this value, if a point in the test set has a reconstruction error larger than UCL, it is signalled as possible anomaly. Since it directly affects the model fault performance, we consider it as a hyperparameter. The choice of the autoencoder hyperparameters is the result of a grid numerical search through Monte Carlo simulation investigation. For each hyperparameter, a finite range of values is explored (Table 1). The grid search algorithm trains a model for every combination of hyperparameter values. We train a single hidden layer autoencoder. In Table 1, the activation function is to be considered as different for each layer. Within the grid search algorithm, a total of 810 NNs are trained, each corresponding to a different combination of the selected hyperparameters. All the NNs were trained by means of the open source JUBE benchmarking environment [11]. Each network has needed a training time of about one minute on a single core of an Intel Xeon Platinum 8160 node of the ENEA CRESCO6 system (2.10GHz, 192 GB RAM, no GPU). The implementation of complete technique and analysis performed through the open source software environment Python 3.0 [12]. The experiments are implemented with Keras 2.2.2 [2] with TensorFlow 1.14.0 [1] as backend. The test set used to validate the model is unbalanced. Thus, precision, recall, and  $F_1$ -score (i.e. accuracy) [5] are used as indicators for evaluating the fault detection performance. The number of epochs used in the training phase is 100 and the batch size is 128. The optimal hyperparameter, resulting from the tuning process, are summarized in Table 1. This model reaches a very high level of fault prediction accuracy ( $F_1$ -score is 98%) and precision (99%), which accounts for the false positive rate. The reconstruction error referred to the test set is plotted in Figure 1. The  $x$ -axis and  $y$ -axis are, respectively, the time and SSE, over all features, between the input and reconstructed output. The horizontal line represents the 99-th percentile, while the dot vertical line is the time instant when the system manifests a fault. As an example in Figure 1, we observe that the reconstruction error is very large. In Figure 2, we plot the accuracy of our autoencoder model as a function of the number of hidden units and the learning rate. Note that the accuracy, in terms of  $F_1$ -score, increases with the number of neurons in the hidden layer, until a local maximum is reached, after which the representational capacity of the model starts decreasing because of the over-fitting problem. For the data set at hand, we select 12 hidden layer nodes, corresponding to the larger accuracy. The learning rate accounts for the speed at which a machine learning model *learns* [13]. As we can see in Figure 2, if we consider a learning rate smaller than 0.01, the accuracy decreases and indicates that the model needs more training epochs to

correctly update the weights. On the contrary, if we consider a larger learning rate, the model quickly converges to a (sub)-optimal solution. The  $F_1$ -score is practically zero, because the model is not able to recognize the faulty condition and wrongly classify all of them as normal.

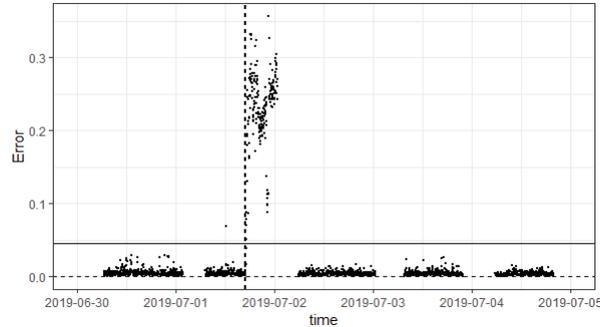


Figure 1: Reconstruction error vs time in the test set.

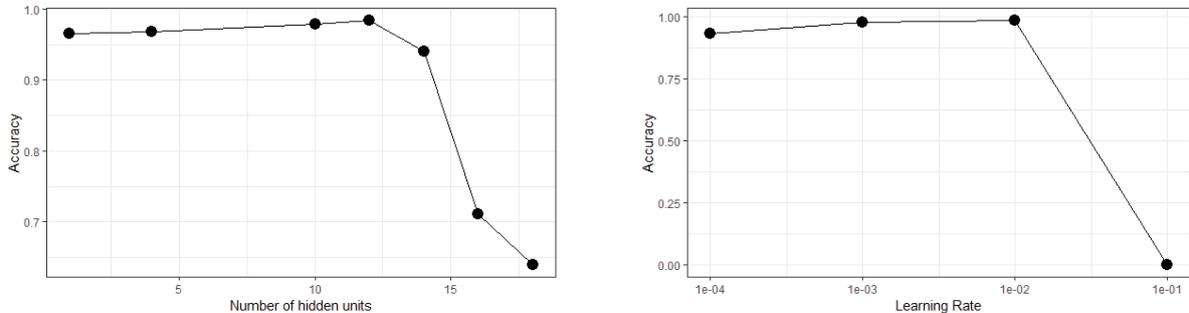


Figure 2: Test set accuracy, in terms of  $F_1$ -score, at different hidden layer nodes (left) and at different learning rates (right).

## 4 Conclusion

The proposed method for fault detection and isolation for HVAC systems embeds autoencoders in a semi-supervised fashion for a SPC scheme. The number of nodes in the hidden layer has been properly selected to avoid overfitting for the problem at hand. Further, the choice of the learning rate plays a key role. The high values of the evaluation performance indices prove that the proposed autoencoder is able to adequately learn the fault characteristics of the system. No data preprocessing is needed and the domain knowledge is limited to the selection of the normal-condition data to be used in the training phase.

## 5 Acknowledgement

The computing resources and the related technical support used for this work have been provided by CRESCO/ENEAGRID High Performance Computing infrastructure and its staff [9]. CRESCO/ENEAGRID High Performance Computing infrastructure is funded by ENEA, the Italian National Agency for New Technologies, Energy and Sustainable Economic Development and by Italian and European research programmes, see <http://www.cresco.enea.it/english> for information.

## References

- [1] Martín Abadi et al. TensorFlow: Large-scale machine learning on heterogeneous systems, 2015. Software available from tensorflow.org.
- [2] François Chollet. Keras. <https://keras.io>, 2015.
- [3] EN. Railway applications - Air conditioning for urban and suburban rolling stock - Part 1: Comfort parameters. UNI EN 14750-1:2006, European Committee for Standardization, 2006.
- [4] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep learning*. MIT press, 2016.
- [5] Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani. *An introduction to statistical learning*, volume 112. Springer, 2013.
- [6] Nasser Jazdi. Cyber physical systems in the context of industry 4.0. In *2014 IEEE international conference on automation, quality and testing, robotics*, pages 1–4. IEEE, 2014.
- [7] Douglas C Montgomery. *Statistical quality control*, volume 7. Wiley New York, 2009.
- [8] Dong Oh and Il Yun. Residual error based anomaly detection using auto-encoder in smd machine sound. *Sensors*, 18(5):1308, 2018.
- [9] Giovanni Ponti et al. The role of medium size facilities in the hpc ecosystem: the case of the new cresco4 cluster integrated in the eneagrid infrastructure. In *High Performance Computing & Simulation (HPCS), 2014 International Conference on*, pages 1030–1033. IEEE, 2014.
- [10] Stelios Psarakis. The use of neural networks in statistical process control charts. *Quality and Reliability Engineering International*, 27(5):641–650, 2011.
- [11] JUBE Developer Team. Jube benchmarking environment, 2016. Software available from [www.fz-juelich.de](http://www.fz-juelich.de).
- [12] Guido Van Rossum and Fred L Drake Jr. *Python reference manual*. Centrum voor Wiskunde en Informatica Amsterdam, 1995.
- [13] Hafidz Zulkifli. Understanding learning rates and how it improves performance in deep learning. *Software testing fundamentals*, 2018.

# BEHAVIOR OF THE TRAPPED SOLVENT AT THE INTERFACE BETWEEN BLOCK COPOLYMERS FILMS AND RANDOM COPOLYMER BRUSH.

Gianmarco Munaò<sup>1\*</sup>, Greta Donati<sup>1</sup>, Antonio De Nicola<sup>2</sup>, Antonio Pizzirusso<sup>1</sup>, Stefano Caputo<sup>1</sup> and Giuseppe Milano<sup>1,2</sup>

<sup>1</sup>*Dipartimento di Chimica e Biologia, Università di Salerno, Via Giovanni Paolo II, 132, I-84084, Fisciano (SA), Italy.*

<sup>2</sup>*Department of Organic Materials Science, Yamagata University, 4-3-16 Jonan Yonezawa, Yamagata-ken 992-8510, Japan.*

**ABSTRACT.** In this contribution we present a simulation study of the role played by the trapped solvent at the interface between block copolymers (BCP) films and random copolymer (RCP) brush. The study is performed by using the hybrid particle-field molecular dynamics method. This approach is applied to systems consisting of poly(styrene)-*b*-poly(methyl methacrylate) thin films deposited on poly(styrene-*r*-methyl methacrylate) brush layers. By calculating the density profile of BCP, RCP and solvent, we observe an increase in the amount of trapped solvent as the BCP film thickness increases. At the same time, a significant decrease of the interpenetration length between the BCP and RCP is found, thus suggesting that the interpenetration between block copolymer chains and brush chains is hampered by the solvent.

## 1 Introduction

Block copolymer (BCP)-based techniques represent nowadays useful protocols to design a wide range of nanostructures without the limitations due to conventional photolithographic methods [1]. Indeed BCPs are relatively inexpensive materials, and their self-assembly, driven by microphase separation, gives rise to periodic features with size, spacing, and shapes depending on the chain length, molecular weight, and composition [2]. In particular, the orientation of the BCP in thin films, typically deposited on a silicon substrate, can be controlled by several strategies, including the use of random copolymers (RCP) grafted to the substrate. Within this mechanism, the BCP ordering is the bottleneck of the process. The infiltration of solvent into the BCP film during the phase-separation process can increase the chain mobility, lowering the free energy barriers to the ordering process. A proper investigation of the role played by the solvent at the BCP-RCP interface is therefore crucial. In the present work we study the behavior of the trapped solvent, constituted by toluene, at the interface between poly(styrene)-*b*-poly(methyl methacrylate) thin films and poly(styrene-*r*-methyl methacrylate) brush layers. The study is performed by means of the hybrid particle-field molecular dynamics representation (MD-SCF) [3], which allows to obtain a fast relaxation of polymer chains [4]. In order to obtain accessible simulation times for a proper calculation of density profile and interpenetration length of BCP, RCP and solvent we thus make use of CRESCO supercomputing resources.

---

\*Corresponding author. E-mail: gmunao@unisa.it.

## 2 Simulation method

The simulation approach adopted in the present study is based on a combination of a standard molecular dynamics approach and a self-consistent field theory [5] for the calculation of non-bonded potentials. The resulting scheme is known as hybrid particle-field model [3]: according to such an approach, the hamiltonian of a system comprised by  $M$  molecules can be split as:

$$\hat{H}(\Gamma) = \hat{H}_0(\Gamma) + \hat{W}(\Gamma) \quad (1)$$

where  $\Gamma$  represents a point in the phase space and the symbol  $\hat{\phantom{x}}$  indicates that a given quantity is a function of the microscopic state corresponding to  $\Gamma$ . In Eq. 1,  $\hat{H}_0(\Gamma)$  is the Hamiltonian of a system with molecules experiencing only intramolecular interactions, whereas  $\hat{W}(\Gamma)$  is the contribution due to the other non-bonded interactions. The latter is calculated as an external potential  $V(r)$  on single particles, due to the density field. The details of the derivation of  $V(r)$  can be found elsewhere [3]. The mean field solution for the potential acting on a particle of type  $K$  at position  $r$ ,  $V(r)$  is:

$$V_K(\mathbf{r}) = k_B T \sum_{K'} \chi_{KK'} \Phi_{K'}(\mathbf{r}) + \frac{1}{\kappa} \left( \sum_K \Phi_K(\mathbf{r}) - 1 \right) \quad (2)$$

where  $k_B$  is the Boltzmann constant,  $T$  is the temperature,  $\chi_{KK'}$  are the mean field parameters for the interaction of a particle of type  $K$  with the density field due to particles of type  $K'$  and the second term on the right-hand side of Eq. 2 is the incompressibility condition,  $\kappa$  being the compressibility. Also,  $\Phi_K(\mathbf{r})$  and  $\Phi_{K'}(\mathbf{r})$  are the density functions of the CG beads of type  $K$  and  $K'$ , respectively.

In the present work proper coarse-grained (CG) models have been introduced in order to simulate the behavior of substrate, polymer and solvent. In the silica substrate 18720 beads (named B, each one representing five  $\text{SiO}_2$  groups) are arranged in a hexagonal configuration with a lattice distance of 0.7 nm. The substrate is constituted by eight connected layers with a total thickness of 4.9 nm. A single RCP chain is constituted by a random sequence of two different beads, named M and S, each one representing five PMMA and six PS repeating units, respectively. A variable number of RCP chains is grafted onto the surface of the silica substrate, according to the target grafting density to simulate. The percentages of M and S bead types are 35 and 65%, respectively. In the BCP chains, the M and S beads are organized in two different blocks with the S block containing 70% of the beads of the whole chain. In the CG model of the solvent, a single bead (named T), represents five toluene molecules. Finally, in order to simulate the interface between the BCP and air, we have also introduced vacuum beads (named V), according to a prescription provided in the literature [6], with the same molecular weight of the T bead type.

## 3 Results

The effect of the solvent at the RCP/BCP interface can be investigated by plotting the density profiles of toluene calculated along the  $z$  axis which is the direction perpendicular to the interface. In the top panels of Fig. 1 the density profiles for all the solvent concentrations and for the BCP film thickness of 30 nm are reported. We observe that the solvent profile shows a double-peak structure, with the first peak placed in proximity of the RCP/BCP interface. The second peak, located at the BCP/air interface, is due to a spurious effect, since in the present simulations the toluene concentration is constant and cannot leave the simulated system. As a consequence, a significant accumulation of toluene at the interface is observed. The amount of solvent at the RCP/BCP interface decreases as the toluene concentration decreases, even though the double-peak structure is still maintained. In the same conditions,

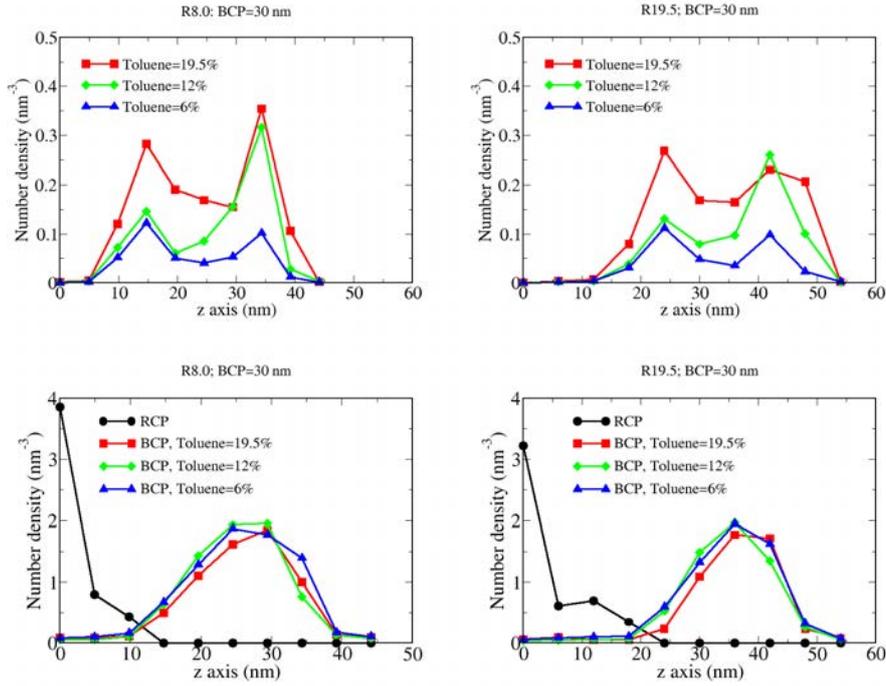


Figure 1: Top panels: density profiles of toluene along the z axis for R8.0 (left) and R19.5 (right). Bottom panels: density profiles of BCP and brush layer along the z axis for R8.0 (left) and R19.5 (right). The thickness of the BCP film is 30 nm.

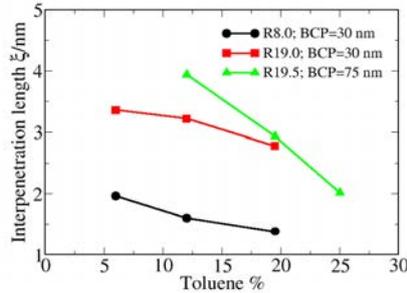


Figure 2: Interpenetration length  $\xi$  as a function of toluene for RCP with different molecular weights and BCP film thicknesses.

we also calculated the density profiles of the BCP along the z axis (bottom panels of Fig. 1). The BCP distributions are quite symmetrical, reaching a maximum value at approximately 2/3 of the simulation box height.

Finally, in order to complete the investigation of the behavior of the solvent at the BCP/RCP interface, we calculated the interpenetration length  $\xi$  between the BCP and RCP brush upon performing the overlap integral by using Eq. 3:

$$\xi = \int dz \rho_1(z) \rho_2(z) \quad (3)$$

where  $\rho_1(z)$  and  $\rho_2(z)$  are the density profiles of RCP and BCP, normalized by their bulk values, respectively. The overall behavior of  $\xi$  as a function of solvent concentration, molecular weight of the RCP, and thickness of the BCP is reported in Fig. 2.  $\xi$  decreases upon increasing the toluene concentration in agreement with experimental results. In addition, it is worth noting that  $\xi$  increases upon increasing the molecular weight of RCP while keeping constant the toluene concentration. Overall, the results shown in the present contribution highlight the strong connection between the existence of residual solvent in the RCP/BCP stack and the thickness of the interface between the brush layer and the BCP film. In particular, these results clearly demonstrate that RCP/BCP interpenetration is significantly hampered by increasing solvent content in the polymer stack.

The investigation of the role played by the solvent at the RCP/BCP interface with such a degree of accuracy has been possible thanks to the CRESCO Supercomputing resources, and for this purpose we have performed extensive molecular simulations on the cresco queues `cresco6_h144`, `cresco4_h144` and `small_h144`. A combined experimental and computational work concerning the topics addressed in the present study has been recently published on *ACS Appl. Mater. Interfaces* **2020**, *12*, 7777, in which we thank the HPC team of Enea (<http://www.enea.it>) for using the ENEA-GRID and the HPC facilities CRESCO (<http://www.cresco.enea.it>) in Portici.

## References

- [1] C. M. Bates, M. J. Maher, D. W. Janes, C. J. Ellison, and C. G. Willson. Block copolymer lithography. *Macromolecules*, 47:2, 2014.
- [2] S. B. Darling. Directing the self-assembly of block copolymers. *Prog. Polym. Sci.*, 32:1152, 2007.
- [3] G. Milano and T. Kawakatsu. Hybrid particle-field molecular dynamics simulations for dense polymer systems. *J. Chem. Phys.*, 130:214106, 2009.
- [4] G. Munaò, A. Pizzirusso, A. Kalogirou, A. De Nicola, T. Kawakatsu, F. Müller-Plathe, and G. Milano. Molecular structure and multi-body interactions in silica-polystyrene nanocomposites. *Nanoscale*, 10:21656, 2018.
- [5] T. Kawakatsu. *Statistical Physics of Polymers*. Springer, Berlin, 2004.
- [6] H. Morita, T. Kawakatsu, and M. Doi. Dynamic density functional study on the structure of thin polymer blend films with a free surface. *Macromolecules*, 34:8777, 2001.

# CALCULATION OF LATTICE THERMAL EXPANSION FOR A CLASS OF ORGANIC CRYSTALS

Andrea Ienco<sup>1\*</sup>, Patrizia Rossi<sup>2</sup> and Paola Paoli<sup>2</sup>

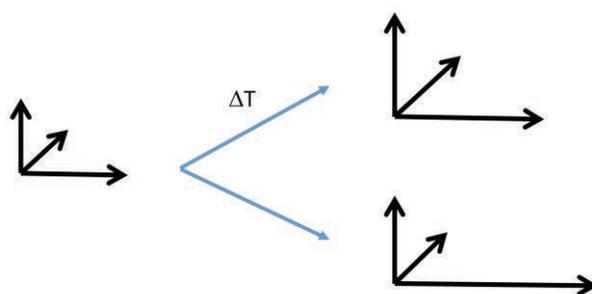
<sup>1</sup>*Consiglio Nazionale delle Ricerche—Istituto di Chimica dei Composti OrganoMetallici (CNR-ICCOM) Via Madonna del Piano 10, Sesto Fiorentino, I-50019 Firenze, Italy<sup>1</sup>*

<sup>2</sup>*Department of Industrial Engineering, University of Florence Via Santa Marta 3, I-50139 Firenze, Italy*

**ABSTRACT.** Thermal anisotropical lattice expansions have been calculate for a series of close related  $\beta$ -blocker organic molecules showing similar packing in solid state. The fast HF-3c method was able to reproduce the experimental trend and on the basis of the calculated results, the role of hydrogen bond column has been highlighted.

## 1 Introduction: Isotropic vs Anisotropic Expansion

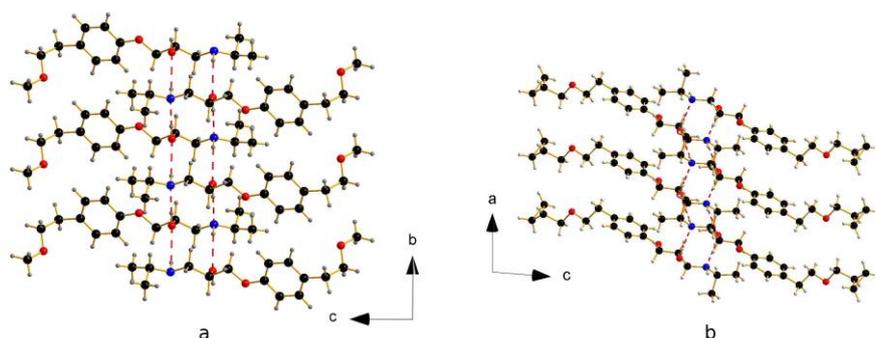
In a crystal, a temperature variation is followed by a change of a cell volume. As shown in Figure 1, when the same proportional change is observed for all cell axes, an isotropical lattice expansion is observed. On the other case, the lattice undergoes to an anisotropical expansion. This is well studied for inorganic materials, while for organic crystals is usually not consider. On the other side for organic crystals and for APIs (Active Pharmaceutical Ingredients), an unrecognized anisotropic thermal expansion could cause misunderstanding and wrong assignments when comparing the room temperature experimental powder pattern with calculated one based on single crystal data obtained at low temperature. [1]



**Fig.1:** Anisotropical vs isotropical lattice expansion.

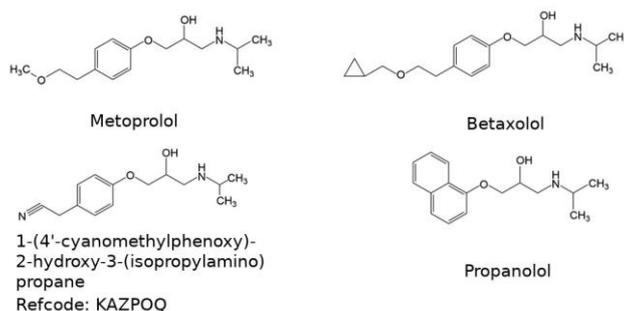
Few years ago, we analysed a series of closed related metoprolol salts. Metoprolol is drug used to treat heart failure and cardiovascular diseases with a  $\beta_1$  selective  $\beta$ -adrenoreceptor blocking activity. We noticed that the tartrate salt expands isotropically, while metoprolol succinate salts undergo an anisotropic lattice expansion. [2] Unfortunately the systems were too large to be treated computationally but also the metoprolol free acid undergoes to the anisotropic thermal expansion. [3] Instead the polymorph I of the  $\beta$ -blocker drug betaxolol (see scheme 1) [4] with the same identical H-bond motifs and packing arrangement in the solid state, as shown in Figure 2, expands isotropically.

<sup>1</sup> Corresponding author. E-mail: andrea.ienco@iccom.cnr.it



**Fig 2:** Comparison of the packing arrangement in metoprolol (a) and betaxolol I (b). Reproduced and adapted from ref. [9].

A search in the Cambridge Database [5] revealed other close related structures, namely 1-(4'-cyanomethylphenoxy)-2-hydroxy-3-(isopropylamino) propane (refcode KAZPOQ) [6], racemic propranolol (refcode PROPRA10) [7] and (S)-propranolol (refcode IMITON) [8] (Scheme 1). Common to all the structures is the presence of hydrogen bonded chains along one cell axis. The length of the axis varies from 4.980Å for PROPRA10 to 5.449Å for KAZPOQ.



**Scheme 1.** Adapted from ref. [9].

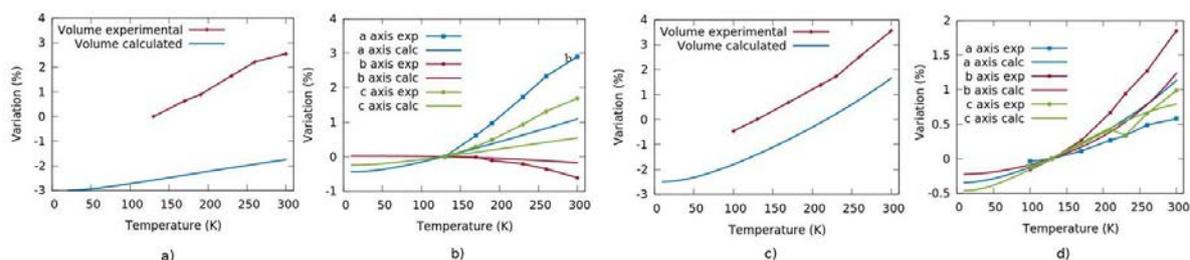
For these systems, a computational approach was feasible and we calculated the thermal expansion for all the crystal structures and for a hypothetical polymorph of metoprolol optimized in betaxolol lattice identified by the acronym POLMET. This allowed to find a correlation between structural arrangement and thermal behaviour for this class of organic molecules. [9]

## 2 Computational methods

The dispersion corrected DFT (DFT-D) methods are nowadays a powerful tool for the calculation of noncovalent interacting systems such as organic crystals. [10] It has been shown that electronic lattice energies have low mean absolute errors (1 – 3 kcal/mol). [11] On the other side, recently the HF-3c method was tested for an expansion study of an organic molecule comparable in size and complexity to our systems. [12] HF-3c is based on Hartree-Fock calculation optimized for speed using different corrections to include long-range London dispersion interactions, the basis set superposition error (BSSE), and the short-range basis set incompleteness. [13] The small MINIX basis set [14] is also used assuring a fast calculation time. Beside the large semiempirical character, it provides on average correct bond lengths and it is not depending on the size of the grid for its analytical nature.

The computational work has been carried out using the CRYSTAL17 software package. [14] The atom coordinates and the lattice constants have been optimized and the final points have been characterized by frequency calculations. In order to calculate the thermal expansion, Quasi Harmonic

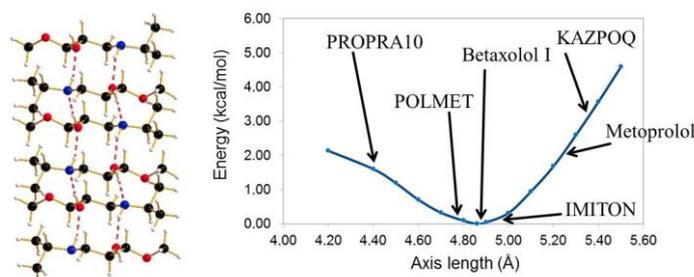
Approximation (QHA) is used to introduce the missing volume dependence of phonon frequencies by retaining the harmonic expression for the Helmholtz free energy. A routine of CRYSTAL17 code allows calculation of the thermal expansion after four constrained volume optimizations and frequency phonon computations. [15] The four points are used: the optimized volume  $V_0$ , a volume with 2% compression and two volumes with 2% and 4% expansion. For metoprolol and betaxolol I, the experimental variations of the cell volume and axes are well reproduced and the trends are as expected. The error on the cell volume is around 3% and 6% in line with the difference usually observed in literature as shown in Figure 3. [12] The computational work has performed on CRESCO6 ENEA HPC Cluster [16]



**Fig.3:** Comparison of the calculated and experimental (dot plot) volume and axes for metoprolol (a, c) and betaxolol (b, d). The data are normalized with respect to the experimental value at 130K. Reproduced and adapted from ref. [9].

### 3 Calculated expansion trends

The calculated thermal expansion for metoprolol and KAZPOQ showed a shrinkage of the axis of the hydrogen bond column, while an increase is instead observed for the other structures. We also noticed that the value of the axis for the metoprolol and KAZPOQ structures was greater than  $5\text{\AA}$ , while for the other four ones it was smaller than  $5\text{\AA}$ . In order to verify this trend, we build a simplified model of our molecules organized as one dimensional chain with hydrogen bond column. The length of the optimized axis is  $4.86\text{\AA}$ . We also re-optimized the model system fixing the mono dimensional axis between  $4.20\text{\AA}$  and  $5.50\text{\AA}$ . The energy variation as function of the axis length is presented in Figure 4.



**Fig. 4:** A picture of the simplified model (left) and the graph of Energy versus axis length (right). The axis length for the given structure are highlighted by arrows. Reproduced and adapted from ref. [9].

The graph tells that a shrinkage of the axis is favorite when in a given system the axis of the hydrogen bond column (e.g. metoprolol or KAZPOQ) is longer than the equilibrium distance of the mono dimensional model. When the axis is shorter, an elongation of the axis instead is expected. In other words, for these systems, hydrogen bonds are stronger than the other forces acting in the crystals. So at the increase of the temperature, the axis parallel to the hydrogen bond will relax toward the preferred length of the hydrogen bond column at a different rate respect the other axes. In

conclusion, anisotropic thermal expansion will be easily predictable, when the value of the crystal axis parallel to the hydrogen bond column is greater and far away the calculated equilibrium distance of the mono-dimensional simplified system.

## References

- [1] G. A. Stephenson. Anisotropic lattice contraction in pharmaceuticals: The influence of cryo-crystallography on calculated powder diffraction patterns. *Journal of Pharmaceutical Sciences* **95**, pp. 821–827, (2006).
- [2] P. Paoli, P. Rossi, E. Macedi, A. Ienco, L. Chelazzi, G. L. Bartolucci and B. Bruni. Similar but different: The case of metoprolol tartrate and succinate salts. *Crystal Growth & Design*, **16**, pp. 789–799, (2016).
- [3] P. Rossi, P. Paoli, L. Chelazzi, L. Conti and A. Bencini, The solid-state structure of the  $\beta$ -blocker metoprolol: A combined experimental and in silico investigation. *Acta Cryst C*, **75**, pp. 87–96, (2019).
- [4] J. Canotilho, R. A. E. Castro, M. T. S. Rosado, M. Ramos Silva, A. Matos Beja, J. A. Paixão and J. Simões Redinha. *Journal of Molecular Structure* **891**, pp. 437–442, (2008).
- [5] C. R. Groom, I. J. Bruno, M. P. Lightfoot and S. C. Ward. The cambridge structural database. *Acta Cryst B*, **72**, 171–179, (2016).
- [6] J. Akisanya, A. W. Parkins and J. W. Steed. A synthesis of atenolol using a nitrile hydration catalyst. *Org. Process Res. Dev.* **2**, pp. 274–276, (1998).
- [7] H. L. Ammon, D.-B. Howe, W. D. Erhardt, A. Balsamo, B. Macchia, F. Macchia and W. E. Keefe. The crystal structures of dichloroisoproterenol, propranolol and propranolol hydrochloride. *Acta Cryst B*, **33**, pp. 21–29, (1977).
- [8] A. A. Bredikhin, D. V. Savel'ev, Z. A. Bredikhina, A. T. Gubaidullin and I. A. Litvinov. Crystallization of chiral compounds. 2. Propranolol: Free base and hydrochloride. *Russian Chemical Bulletin*, **52**, pp. 853–861, (2003).
- [9] P. Paoli, S. Milazzo, P. Rossi and A. Ienco. Rationalization of Lattice Thermal Expansion for Beta-Blocker Organic Crystals *Crystals* **10**, 350, (2020).
- [10] J. G. Brandenburg and S. Grimme. Organic crystal polymorphism: A benchmark for dispersion corrected mean field electronic structure methods. *Acta Crystallogr., Sect. B: Struct. Sci., Cryst. Eng. Mater.* **72**, pp. 502–513, (2016).
- [11] M. Cutini, B. Civalleri, M. Corno, R. Orlando, J.G. Brandenburg, L. Maschio, P. Ugliengo. Assessment of Different Quantum Mechanical Methods for the Prediction of Structure and Cohesive Energy of Molecular Crystals. *J. Chem. Theory Comput.* **12**, pp. 3340–3352, (2016).
- [12] J.G. Brandenburg, J. Potticary, H.A. Sparkes, S.L. Price, S.R.; Hall. Thermal Expansion of Carbamazepine: Systematic Crystallographic Measurements Challenge Quantum Chemical Calculations. *J. Phys. Chem. Lett.* **8**, pp. 4319–4324, (2017).
- [13] R. Sure, S. Grimme. Corrected Small Basis Set Hartree-Fock Method for Large Systems. *J. Comput. Chem.* **34**, pp. 1672–1685, (2013).
- [14] R. Dovesi, A. Erba, R. Orlando, C. M. Zicovich-Wilson, B. Civalleri, L. Maschio, M. Rérat, S. Casassa, J. Baima, S. Salustro and B. Kirtman. Quantum-mechanical condensed matter simulations with CRYSTAL. *WIREs Computational Molecular Science* **8**, e1360, (2018).
- [15] A. Erba. On combining temperature and pressure effects on structural properties of crystals with standard ab initio techniques. *J. Chem. Phys.* **141**, 124115, (2014).
- [16] F. Iannone, F. Ambrosino, G. Bracco, M. De Rosa, A. Funel, G. Guarnieri, S. Migliori, F. Palombi, G. Ponti, G. Santomauro and P. Procacci "CRESCO ENEA HPC Clusters: A Working Example of a Multifabric GPFS Spectrum Scale Layout." 2019 International Conference on High Performance Computing & Simulation (HPCS), Dublin, Ireland, pp. 1051-1052, (2019), doi:10.1109/HPCS48598.2019.9188135.

# KINETIC ANALYSIS OF ACCELERATOR DRIVEN SYSTEM WITH MCNP 6.2 AND JEFF 3.3 NUCLEAR DATASET ON THE CRESCO COMPUTATIONAL FACILITY: THE KUCA REACTOR MODEL

Nunzio Burgio<sup>1\*</sup>, Alfonso Santagata<sup>1</sup>

ENEA *FSN-FISS-RNR, CR Casaccia,*  
*S.P. Anguillarese, S.M. di Galeria Rome – Italy*<sup>1</sup>

**ABSTRACT.** We develop an MCNP model of the KUCA (Kyoto University Critical Assembly) in an ADS (Accelerator Driven System) experimental layout with the primary objective of testing the novel cross-section data set JEFF 3.3 and implementing post-processing methodologies to produce Pulsed Neutron Source (PNS) responses from time-resolved MCNP simulation. Since the computational demand is prohibitive for our laboratory resources, the use of the CRESCO facility was mandatory, and it results in a two order of magnitude decreases of the average simulation time. Furthermore, the overall agreement between KUCA simulated and experimental PNS results was somewhat encouraging.

## 1 Introduction

ADS (Accelerator Driven System) are nuclear system based on the coupling of a sub-critical core (i.e. a fissile mass that cannot sustain a fission reaction chain autonomously) with an external source that furnishes a fraction of neutron necessary to maintain the nuclear chain reaction. The ADS external source belongs from the interaction of protons, accelerated by a cyclotron or in a linear system, that impinging into a metallic target, generates spallation neutrons [1]. KUCA (Kyoto University Critical Assembly) is a multi-core facility used for reactor physics studies [2] that also allows the investigation of ADS due to its capacity of coupling nuclear cores with the Fixed-Field Alternating Gradient (FFAG) proton accelerator [3]. What makes KUCA attractive for neutronic simulations is its relatively simple geometric layout, the presence of a limited number of materials, and high heterogeneity of the fuel. For this reason, following the indication find in [4], we develop an MCNP model of the KUCA-ADS system. One of the crucial aspects of an ADS is the continuous monitoring of the core to avoid uncontrolled reactivity increases that could trigger an incidental sequence. The PNS (Pulsed Neutron Source) is an experimental method, among others, candidate for the online monitoring of core reactivity. The main goals of our activity are to test the new nuclear data package JEFF 3.3 [5], refining some post-processing techniques for MCNP to simulate PNS (Pulsed Neutron Source) response [6] and estimate, via the Area Method [7], the degree of subcriticality of the KUCA-ADS system. Since the scope of the present paper is to report some of the aspects of this activity related to the use of the CRESCO computational facility, we indicate only a limited number of results. However, the reader can find a more comprehensive discussion of the results in [10]. The simulation results are in relatively good agreement with the experimental data, and future work will focus on the improvement of sensitivity analysis on cross-section data to define the margin of uncertainties in the simulation results and to assess their accuracy.

## 2 MCNP 6.2 code on CRESCO 4

As denoted by its acronym MCNP 6.2 (Monte Carlo N Particles) is a general-purpose, three-dimensional, time-resolved, MC (Monte Carlo) code for the transport of nuclear particles. It is used in the fields of Reactor Physics, Nuclear Medicine, Nuclear instrumentation design, Nuclear Imaging,

---

<sup>1</sup> Corresponding author. E-mail: nunzio.burgio@enea.it

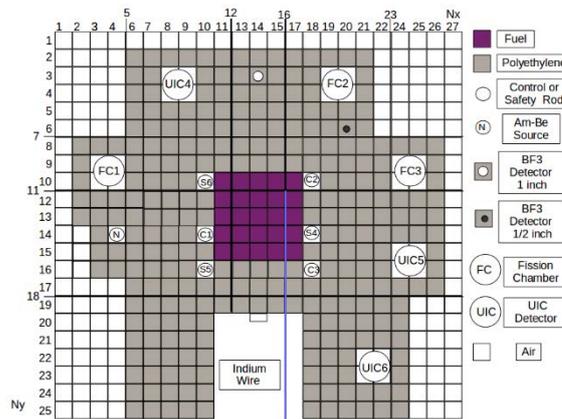
Shielding, Dosimetry, Accelerator and Target design, Aerospace [8]. As a Monte Carlo codes, it infers the macroscopic quantities relevant for nuclear particles transport from the sampling of the cross-section data on a user-defined geometry without any needs to resolve the transport equations. The main drawback of the MC technique is a slow convergency of the results that request, in some cases, the use of parallel programming techniques. We use an MPI executable obtained compiling the MCNP 6.2 source code with the Intel Fortran compiler version 14.0.1 and linking it with OpenMPI library version 1.6.5. A typical PNS simulation on KUCA request 320 tasks (1 thread each) with of sample of  $2e9$  source protons with a simulation time of about 510 minutes. The use of CRESCO decreases of two orders of magnitude the computational time of the simulations.

### 3 Description of the KUCA model

The essential elements of the KUCA geometric layout, described in great details in [4], are:

- Control Rods: height 1524 mm, neutron absorber Boron Oxide.
- Moderator: Polyethylene Blocks of 1524 mm of height.
- Fuel Rods: Thin Layer of metallic Uranium alternative with polyethylene in the central part (height 376 mm). Polyethylene buffers above and below the central part.

A matrix of 25 rows and 27 columns reports the position of each element (Figure 1). For example, the fuel element located aside of the safety rod S6 is the element E(11,10) in the matrix, and the 1-inch BF3 detector is E(3,14).

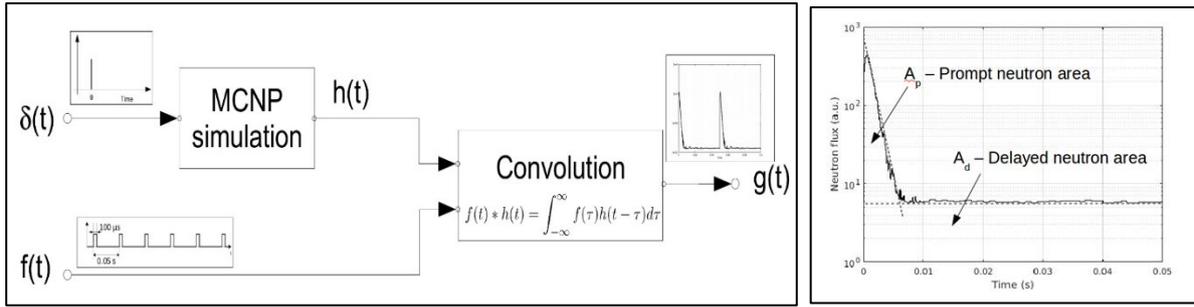


**Figure 1:** Schematic view of the KUCA geometry.

As reported in Figure 1, the KUCA I-3 configuration has a  $5 \times 5$  matrix of fuel rods and the six control rods inserted in the core to realise a negative reactivity of  $-2390$  pcm. We implement such configuration in an MCNP formalism using a script in Octave programming language [9] which accepts in input the geometrical and material specifications for the position of each element of the schematisation matrix, and it produces as output the correspondent MCNP input file. We add the beamline and the spallation Tungsten target manually to finalise the MCNP input. The source is a uniformly-distributed, circular-shaped (radius 2.0 cm) proton beam with an energy of 100 MeV. Finally, we add the material compositions and their densities in the MCNP input according to the specifications reported in [4].

### 4 Area method and implementation of the simulated PNS response

The complete simulation of a PNS experiment is a time-consuming task. Therefore, we have adopted, beside the use of the calculation power CRESCO, a two-step simulation method to reduce the computation time further. In the first step (see Figure 2, left frame), an MCNP simulation has been performed having no time dependence in the proton source while having a suitable time binning in the secondary neutron fluence estimators.



**Figure 2:** Calculation procedure to obtain the PNS response (left frame); Identification of prompt and delayed neutron areas for the calculation of the reactivity level in a PNS response (right frame).

In this modality, all the source particles start at time zero, as for a single instantaneous pulse that formally obeys to a Dirac's Delta time distribution  $\delta(t)$ . Consequently, a Monte Carlo estimator collects, in its time bins, the particle scores according to their time of arrival (counted from the birth of their source proton ancestor) at the estimator site generating the detector Instantaneous Answer Function (IAF)  $h(t)$ . In the second step, a dedicated computer program convolves  $h(t)$  with the user-selected time-dependent source  $f(t)$ , according to the relation

$$g(t) = f(t) * h(t) = \int_{-\infty}^{+\infty} f(\tau) h(t - \tau) d\tau \quad (1)$$

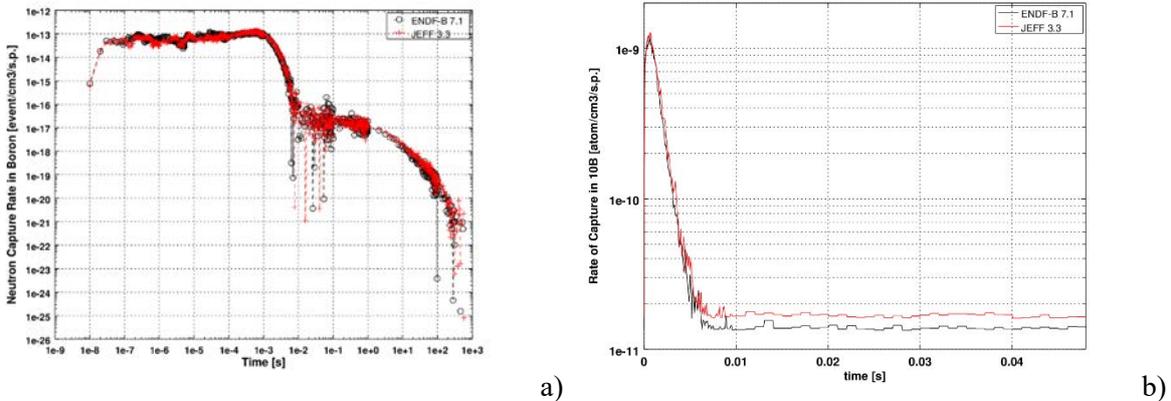
It is worth noticing that  $f(t)$  is a succession of pulses sufficiently long to ensures stable neutron delayed level in the output response. Also, the pulse's period must be longer than the neutron prompt decay time. The final simulated PNS response  $g(t)$  has the shape reported in Figure 2, that is similar to the one obtained experimentally, and allows the estimation of the system reactivity level, via the area method, using the well-known relation

$$-\frac{\rho}{\beta} = \frac{A_p}{A_d} \quad (2)$$

where  $A_p$  and  $A_d$  are respectively the prompt neutron area and the delayed neutron area in the PNS response reported in the right frame of Figure 2.

## 5 Simulation Results and Discussion

Figure 3a reports the time distribution of the  $^{10}\text{B}$  capture reaction rate as estimated by MCNP. This outcome is the IAF  $h(t)$  that convolved according to Equation 1 with the accelerator pulses (period 20 Hz, width 100  $\mu\text{s}$ ,  $f(t)$  function) yield in the PNS response ( $g(t)$  function of equation 1) of the  $^{10}\text{B}$  neutron capture rate (Figure 3b) expected after the accelerator shutdown (PNS response): the delayed neutron constant background is identifiable starting from 0.01 seconds.



**Figure 3:** a) IAF of  $^{10}\text{B}$  capture reaction rate as estimated by MCNP (red curve); b) The PNS response after accelerator shutdown for the  $^{10}\text{B}$  capture detector (red curve).

Table 2 reports the reactivity level calculated from the simulated PNS responses using equation 2 (area method).

**Table 2 :** Comparison of the calculated and measured level of reactivity at the BF3#1 detector in KUCA sub-critical configuration

Simulated Area Method		Experimental Area Method [4]	
P [PCM]	Error [PCM]	$\rho$ [PCM]	Error [PCM]
-2259	66	-2398	35

The simulated PNS responses and the reactivity estimated with the area method are in relatively good agreement with measurements. Further work will be carried out in a systematic sensitivity analysis to assess the overall simulation uncertainties. Finally, the possibility to use the CRESCO computational facility represents for us an advantage not only shortening the calculation time but also allows us to consider a more significant number of simulation alternatives for the interpretation of the experimental outcomes.

## References

- [1] M. Salvatores. Physics features comparison of TRU burners: Fusion/Fission Hybrids, Accelerator-Driven Systems and low conversion ratio critical fast reactors. *Ann. Nucl. Energy* 36 (2009) 1653–1662.
- [2] <https://www.rri.kyoto-u.ac.jp/en/facilities/ca>
- [3] C.H. Pyeon, J.-Y. Lim, Tsuyoshi Misawa. Experiments on injection of spallation neutrons by 100-MeV protons into the Kyoto university critical assembly. January 2010, *Transactions of the American Nuclear Society* 103:21-22
- [4] C.H. Pyeon. ‘Neutronics on Solid Pb-Bi in Accelerator-Driven System with 100 MeV Protons at Kyoto University Critical Assembly’, Rev. 5 an internal report of the Research Reactor Institute, Kyoto University, Japan.
- [5] <https://www.oecd-nea.org/dbdata/jeff33/index.html>
- [6] N. Burgio et al., GUINEVERE Experiment: Kinetic analysis of some reactivity measurements methods by Deterministic and Monte Carlo codes. *PHYSOR 2012, – Advances in Reactor Physics – Linking Research, Industry, and Education-April 15-20, 2012*
- [7] N.G. Sjöstrand. Measurements on a Subcritical Reactor using a Pulsed Neutron Source. *Arkiv för Fysik*, 11, 13, 1956.
- [8] C.J. Werner (editor). *MCNP User’s Manual Code Version 6.2. – October 27, 2017 Rev. 0. LA-UR-17-29981.*
- [9] <https://www.gnu.org/software/octave/>
- [10] N. Burgio et al. ANALYSIS OF CRITICAL AND SUB-CRITICAL CONFIGURATIONS OF THE KUCA REACTOR WITH MCNP6.2 USING JEFF 3.3 NUCLEAR DATASET. *RRFM-IGORR Conference Proceedings , 24-28 March 2019, Amman Jordan. ISBN-978-92-95064-31-7.*

# STRATEGY DIVERSITY PROMOTES TIME DEPENDENCE COEXISTENCE OF COOPERATION AND DEFECTION IN EVOLUTIONARY GAMES

Irene Sendiña-Nadal<sup>1,2\*</sup>, Inmaculada Leyva<sup>1,2</sup>, Juan A. Almendral<sup>1,2</sup> and Stefano Boccaletti<sup>3,4,5</sup>

<sup>1</sup>*Universidad Rey Juan Carlos & GISC, Complex Systems Group, C/Tulipán s/n, 28933 Móstoles, Madrid, Spain*

<sup>2</sup>*Universidad Politécnica de Madrid, Center for Biomedical Technology, 28223 Pozuelo de Alarcón, Madrid, Spain*

<sup>3</sup>*Northwestern Polytechnical University, Unmanned Systems Research Institute, 710072 Xi'an, China*

<sup>4</sup>*CNR - Institute of Complex Systems, Via Madonna del Piano 10, I-50019 Sesto Fiorentino, Italy*

*Moscow Institute of Physics and Technology (National Research University), 9 Institutskiy per., 141701 Dolgoprudny, Moscow Region, Russian Federation*

**ABSTRACT.** Evolutionary games provide the theoretical backbone for many aspects of our social life: from cooperation to crime, from climate inaction to imperfect vaccination and epidemic spreading, from antibiotics overuse to biodiversity preservation. An important, and so far overlooked, aspect of reality is the diverse strategic identities of individuals. In this work, we show that allowing individuals to adopt different strategies with different partners yields a very rich evolutionary dynamics, including time-dependent coexistence of cooperation and defection, system-wide shifts in the dominant strategy, and maturation in individual choices.

## 1 Introduction

Evolutionary games are the paradigm for studying agent interaction dynamics in collective and networked systems, and they are used to investigate competition in economic, social, and technological systems [1]. In particular, much effort has been devoted to the evaluate the impact of the connectivity structure in the evolution of social dilemmas on networks [5, 4].

So far, studies of evolutionary dynamics in networked ensembles have limited to the case were each individual has a same cooperation-defection strategy with all their neighbours [7]. In this work we abandon that limitation in favour of explicitly allow the players to adopt different strategies with each partner. As a result, new and complexer dynamics flourish, as time-dependent coexistence of cooperation and defection. Our results show that switching the focus from node to link dependent strategies are critical for capturing a wide number of behaviors observed in human societies.

---

\*Corresponding author. E-mail: irene.sendina@urjc.es.

## 2 Model

We start by considering a network,  $G$ , made of  $N$  players (or units, or nodes), connected according to a symmetric adjacency matrix,  $A = (a_{ij})$ , where  $a_{ij} = 1$  if units  $i$  and  $j$  interact, and  $a_{ij} = 0$  otherwise. Moreover,  $k_i = \sum_{j=1}^N a_{ij}$  is the degree of player  $i$ . The strategy matrix  $S = (s_{ij})$ , is defined as  $s_{ij} = 1(2)$  if player  $i$  cooperates(defects) with player  $j$ , and  $s_{ij} = 0$  if  $a_{ij} = 0$ . Each player  $i$  is then here associated to a  $k_i$ -dimensional vector of independent strategies. The prisoner's dilemma interaction game [6] is governed by the payoff matrix:

$$P = \begin{array}{c|cc} & C & D \\ \hline C & b-c & -c \\ D & b & 0 \end{array}$$

where parameters  $b > c > 0$  account for the gain or losses each player obtains in all the possible defection (D) or cooperation (C) pairwise interactions. The Nash equilibrium is mutual defection [2].

The link gain or loss is calculated as  $g_{ij} = P(s_{ij}, s_{ji})$ . The average payoff of player  $i$  is therefore  $g_i = \frac{1}{k_i} \sum_{j \in \Lambda_i} g_{ij}$  where  $\Lambda_i$  is the set of neighbors of  $i$ . The network state can be tracked by the ensemble averaged cooperation level  $\rho(t) = \frac{1}{N} \sum_i \frac{k_i^C(t)}{k_i}$  where  $k_i^C(t)$  is the number of cooperative links of node  $i$  at time  $t$ . Moreover,  $\rho = \langle \rho(t) \rangle_T$  and  $\sigma_\rho$  are, respectively, the time-averaged cooperation density over an observation time  $T$  and the standard deviation of  $\rho(t)$ .

Each player  $i$  is assumed to imitate the strategy (C or D) used against it by its best neighbor along each one of its  $k_i$  links with probability

$$p = \frac{1}{1 + e^{-\frac{\Delta g_i}{\alpha}}}, \quad (1)$$

where  $\alpha$  is a free parameter playing the role of an effective temperature in the above Fermi-Dirac function, and  $\Delta g_i$  is the payoff balance (in units of  $b$ ) between  $i$  and its best performer neighbour.

## 3 Numerical results

### 3.1 Numerical methods and use of computational resources

We used the CRESCO resources for the computational study of the prisoner's social dilemma in networks allowing individuals to use different strategies in all its interactions, considering different topologies, network size and different average degrees. Homemade MATLAB scripts were written and compiled with MATLAB Compiler and executed at CRESCO with MATLAB Runtime. Extensive distributed computations have been performed for large parameters ranges, with statistical validation of the results. Calculations were performed in CRESCO4 and CRESCO5, using the h144 queues for full evolution simulations. Homemade MATLAB scripts were used for visualizing the results.

### 3.2 Game evolution in scale-free networks

In Figure 1 we show (a) the cooperation density  $\rho$ , (b) the associated standard deviation  $\sigma_\rho$ , and (c) the link strategy index defined by  $\ell_i = 1 - \frac{|k_i^D - k_i^C|}{k_i}$  in the  $(b/c, \alpha)$  parameter space, obtained on scale-free,  $N=1000 \langle k \rangle=4$  networks.

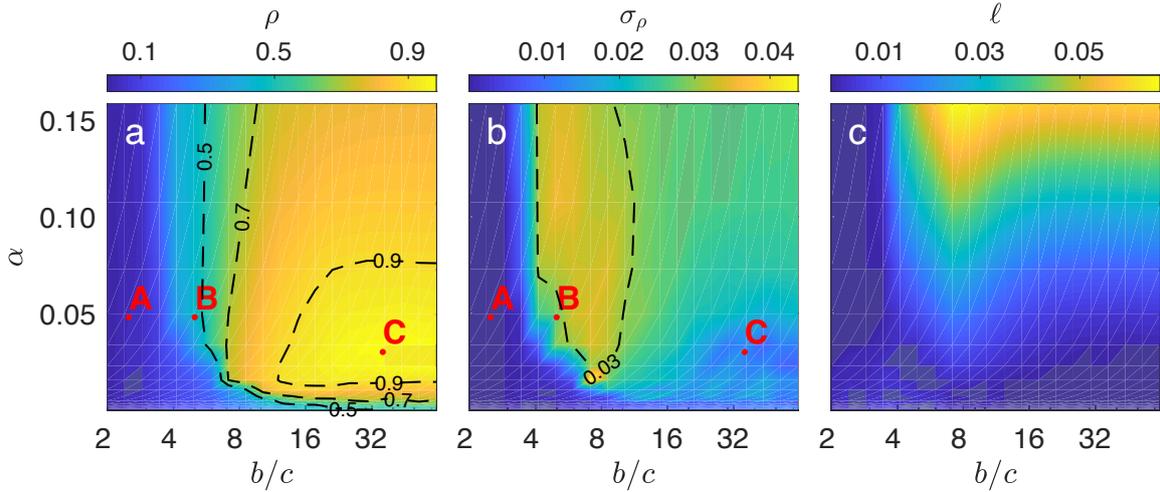


Figure 1: **Diversity in strategic identities induces time-dependent states.** Heatmaps (with colorbar on top) of  $\rho$  (a),  $\sigma_\rho$  (b), and  $\ell$  (c) in the parameter space  $(\alpha, b/c)$ . See text for the definitions of all quantities. Simulations were performed over  $T_f = 10,000$  time units, while the time-averaging of all quantities was made over the last  $T = 2,000$  steps. Each point is obtained by a further ensemble average over 10 SF network realizations with size  $N = 1,000$  and average degree  $\langle k \rangle = 4$ . Isolines at the values of  $\rho = 0.5, 0.7$ , and  $0.9$  are drawn in panel a. The isoline at the value of  $\sigma_\rho = 0.03$  is reported in panel b. Three different regimes can be identified: i) a stationary state in which cooperation dominates the network, at sufficiently large  $b/c$  ratios and sufficiently small  $\alpha$  values (the red point labeled with C and located at  $(36, 0.03)$  in panels a,b); ii) the classical regime for  $b/c < \langle k \rangle$  where defection becomes dominant and cooperators can no longer survive (the red point labeled with A and located at  $(2.5, 0.048)$  in the same panels); and iii) a novel *dynamical state* with coexisting cooperators and defectors (the red point labeled with B and located at  $(5, 0.048)$ ). The emergence of the new state is a direct consequence of diversity in the players' strategic identities.

Figure 1(a) reveals that cooperation dominates for large  $b/c$  ratios and small  $\alpha$  values. This regime is marked with the red point C. When decreasing  $b/c$ , a transition occurs at  $b/c \sim \langle k \rangle$  toward the classical regime marked with the red point A [3], where defection dominates. In between, a novel *dynamical state* emerges, where time-dependent coexistence of cooperators and defectors persists in time. In panels (b) and (c) one can observe that this coexistence is characterized by a large  $\sigma_\rho$  and  $\ell$ , as indicated by the red point B in panels (a) and (b), whereas regimes A and C are characterized by small values of those quantities. The new state, therefore, is a result of the diversity in the players' strategic identities.

Figure 2 reveals the details behind the dynamics emerging at the three points marked by letters A, B, and C in Fig. 1a. In case A (panel A1) defection dominates over cooperation in a stationary way. In case C (panel C1) the cooperators dominate after a longer transitory dynamics.

Case B (Fig. 2, panel B1) shows that the coexistence of cooperators and defectors is characterized by system-wide shifts in the dominant strategy. At odds with case C, maturity never takes place, and players are forever trapped in a cycle of individual changes that allows time-dependent cooperation. In this case, diverse strategic identities are thus a successful vehicle for maintaining the survival of cooperation even at strongly unfavorable  $b/c$  ratios, but not for assuring the elimination of defectors.

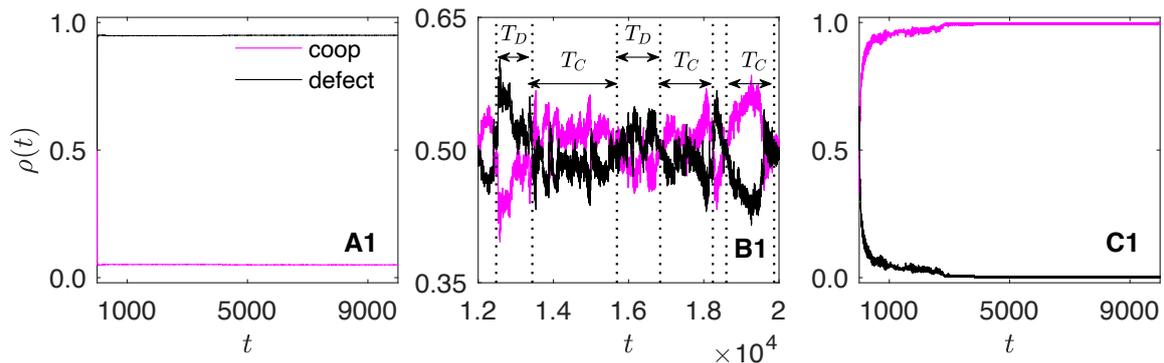


Figure 2: **The coexistence of cooperators and defectors is characterized by system-wide shifts in the network dominant strategy.** Time evolution of  $\rho(t)$  of a single run of the game on a single realization of an SF network with  $N = 1,000$  and  $\langle k \rangle = 4$ . Each panel corresponds to the game parameters marked by red points in Fig. 1a). In panel A1, defection dominates over cooperation in a steady state. Conversely, in panel C1 cooperators ultimately dominate, though temporal fluctuations persist over a relatively large time scale. In the novel state (panel B1), the coexistence of cooperators and defectors is characterized by system-wide shifts in the dominant strategy. The horizontal lines in panel B1 mark the transitions from periods during which cooperation is dominant (intervals with label  $T_C$ ) to time lapses during which defection dominates instead (intervals with label  $T_D$ ).

## 4 Conclusions

Unlike simple forms of life which use the same strategy for all their interactions, humans and many animals do not. In this study of evolutionary games on structured networks, we allow individuals to use different strategies in its interactions. We show that, regardless of the particularities of the implementation, diverse strategic identities give rise to evolutionary dynamics that is unseen in the monotonic case. Our results include sudden systemic shifts and oscillations of dominant strategies in the network. We have observed time-dependent cooperation and maturation in individual strategic choices, where players initially change their strategies frequently and use different strategies with different neighbors, only to eventually converge to an essentially monotonic strategy in time.

Given the high and robust degree to which diverse strategic identities affect the evolutionary dynamics on networks, we argue that truly applicable models of the future should relax the *one strategy for all* limitation and embrace the freedom and complexity of more realistic diverse strategic identities. As our research shows, this allows individuals to carefully gauge their strategic vectors, which can lead to socially optimal outcomes that would otherwise remain hidden or completely unattainable.

## References

- [1] Marco Alberto Javarone. *Statistical Physics and Computational Methods for Evolutionary Game Theory*. Springer, Cham, 2018.
- [2] John Nash. Equilibrium points in n-person games. *Proc. Natl. Acad. Sci. USA*, 36:48–49, 1950.

- [3] Hisashi Ohtsuki, Christoph Hauert, Erez Lieberman, and Martin A. Nowak. A simple rule for the evolution of cooperation on graphs and social networks. *Nature*, 441:502–505, 2006.
- [4] M. Perc, J. Gómez-Gardeñes, A. Szolnoki, and L. M. Floría and Y. Moreno. Evolutionary dynamics of group interactions on structured populations: a review. *J. R. Soc. Interface*, 10:20120997, 2013.
- [5] M. Perc and A. Szolnoki. Coevolutionary games – a mini review. *BioSystems*, 99:109–125, 2010.
- [6] Anatol Rapoport and Albert M Chammah. *Prisoner’s dilemma: A study in conflict and cooperation*. University of Michigan Press, Michigan, 1965.
- [7] György Szabó and Gábor Fáth. Evolutionary games on graphs. *Phys. Rep.*, 446:97–216, 2007.

# COMPUTATIONAL DETERMINATION OF CRITICAL CURRENTS IN GRANULAR SUPERCONDUCTORS

Giorgio Mancini\*

*Università Di Camerino, Sezione di Fisica della Scuola di Scienze e Tecnologie, Via Madonna delle Carceri 9B,  
62002 Camerino (MC), Italia*

**ABSTRACT.** Disordered granular superconductors present one or more small resistivity peaks curves at lower temperatures than the critical one  $T_c$  [1–9]. A model based on Josephson junctions array and Coulomb blockage capable to explain re-entrant resistive peaks in diamond-like carbon-silicon films containing tungsten was proposed by Chudinov et al. [10]. A computational method [11] developed according to such a model for the determination of superconducting percolation paths clearly showed their link to the peak-type re-entrant superconductivity [10]. Based on the same model and approach, a further method was developed to compute the critical current densities  $j_c$  through simulated C—Si—W granular superconductors in the temperature region  $T \leq T_c$ . The strict correspondence of experimental and computed results [12] constituted a deeper confirmation of the validity of the proposed underlying physical model and its capability to describe richer, more structured aspects of the re-entrant phenomenon.

## 1 Introduction

Re-entrant or quasi-re-entrant superconductivity phenomena in disordered metals have interested experimental and theoretical researchers since many years [13]. Re-entrant superconductivity has been observed for the first time in magnetic superconductors of the type  $HoMo_6S_8$  and  $ErRh_4B_4$  [14] consisting in a superconductor - dielectric transition at a certain temperature  $T_r$  under the critical temperature of the superconductor  $T_c$ , the normal state then maintained down to  $T = 0K$ . Amorphous carbon-silicon films containing tungsten (C—Si—W films: they are superconducting granular metals consisting in superconducting metallic grains, approximately the same dimension, embedded in a dielectric matrix) revealed another kind of re-entrant superconductivity phenomenon [4, 5], indicated as *peak-type* re-entrant superconductivity: some samples show a very small but clear increase in resistance when going down in temperature, followed by a return of the resistance to the former value, so that one or more resistive peaks appear in a range of temperatures that should correspond to the superconducting state. Chudinov et al. [12] showed that multiple re-entrant peaks are associated to grains whose sizes present distinct, well separated mean values. For example, the results we report in this paper for two re-entrant peaks come from simulation boxes containing 65% and 45% grains with average radii  $\bar{R}_1 = 220\text{\AA}$  and  $\bar{R}_2 = 250\text{\AA}$ , respectively (Figure 3). Based on *path messages* through communication networks [15–17] and on *deferred path messages* for an efficient reduction of information handling, a computational method to determine percolation paths trough granular superconductors showing peak-type re-entrant superconductivity was recently presented and illustrated in details [11]. Considering the interplay of Josephson energy, thermal energy and Coulomb blockage, such a method proved capable to explain the appearance of re-entrant resistivity peaks experimentally observed in diamond-like carbon-silicon films containing tungsten. It showed that the re-entrant resistivity was the consequence

---

\*Corresponding author. E-mail: giorgio.mancini@unicam.it.

of the absence of percolation paths on activated superconducting grains crossing the entire samples from side to side. The computed and experimental data for the temperature range and location of the peaks resulted in strict agreement. From a physical point of view, the method allowed to describe the paths the Cooper pairs walk from an electrode to the other one. Additional, richer information can be obtained from the study of the current traversing the superconductor. As will be exposed in details later, the intensity of the super-current that can flow through each couple of connected superconducting grains presents a maximum value, depending on radii and distance. This maximum super-current is called the *critical current* of the couple, due to the fact that if a higher current than the critical one flows through two coupled grains, it provokes their transition into the dielectric state. For a superconductor as a whole, the critical current consists in the maximum current flowing through it such that no local inter-granular critical current is exceeded. In other words, the critical current for a superconductor is the maximum current it can be traversed by while maintaining its superconducting phase. Using the model and approach for percolation as a starting point, a further method was developed to compute the critical current densities  $j_c$  through simulated C—Si—W granular superconductors in the temperature region  $T \leq T_c$ . Experimental measurements and simulations for critical currents in C—Si—W granular superconductors with one or two re-entrant peaks were carried out and compared: they showed strict, clear correspondences in all cases.

## 2 The physical bases

Metallic granular systems are superconductor when entirely crossed by Cooper pairs. In order to obtain a superconducting shunt in a percolative system, there must be a path consisting of superconducting connected grains. The absence of such a path at temperatures lower than  $T_c$  characterizes the appearance of re-entrant resistivity peaks [10]. According to theoretical analyses [18–20], superconducting metallic granular systems are governed -at low temperatures- by the competition between two energies  $E_c$  and  $E_J$ , being  $E_c$  the energy of the Coulombian blockage acting on grains  $i$  and  $j$  (Figure 1):

$$E_c^{i,j} = \frac{1}{2} \frac{Q_i Q_j}{C_{ij}}, \quad (1)$$

(with  $C_{ij}$  the inter-granular capacitance,  $Q_i$  and  $Q_j$  the charges excess due to Cooper pairs (charge  $2e$ ) and  $E_J$  the characteristic energy of Josephson junction coupling for nearest neighbouring grains ( $i = j \pm 1$ ):

$$E_J^{i,j}(T) = \frac{\pi \hbar}{4e^2} \frac{1}{R_N^{i,j}} \Delta(T) \tanh\left(\frac{\Delta(T)}{2k_B T}\right) \quad (2)$$

where  $\Delta(T) = \Delta(0)(1 - \frac{T}{T_c})^{\frac{1}{2}}$  is the superconducting gap at temperatures  $T < T_c$ ,  $\Delta_0$  the superconducting gap at  $T = 0K$  [21],  $R_N^{i,j}$  the resistance for single electron tunnelling between neighbouring grains and  $k_B$  the Boltzmann constant. Finally, taking into account all energies involved at each temperature  $T < T_c$ , (namely  $E_c$ ,  $E_J$  and  $k_B T$ ), two granules are connected -or activated- when one of the conditions

$$(a) \quad k_B T \geq E_c^{i,j} \quad (b) \quad k_B T < E_c^{i,j} \text{ and } z E_J^{i,j} > E_c^{i,j} + k_B T \quad (3)$$

holds, with  $z$  -the nearest neighbours average number- accounting for shielding effect. Finally, for each couple of connected grains ( $i, j$ ) a critical current  $I_c^{i,j}$  and its associate density  $J_c^{i,j}$  are defined which depend on the grains distance and radii  $R_i, R_j$  according to the relations:

$$I_c^{i,j} = \frac{\pi}{2e R_N^{i,j}} \Delta(T) \tanh\left(\frac{\Delta(T)}{2k_B T}\right) \quad \text{and} \quad J_c^{i,j} = \frac{I_c^{i,j}}{\pi (R_{min}^{i,j})^2} \quad (4)$$

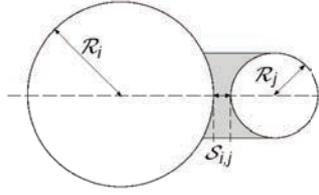


Figure 1: Two neighbouring grains immersed the dielectric matrix, separated by the distance  $S_{i,j}$ .

where  $R_{min}^{i,j}$  is the smallest radius of the two grains.

### 3 The model

As for the originating model, the numerical representations of our disordered granular systems were generated positioning spherical grains in a simulation box, their concentration, radii and mutual distances distributed gaussianly around known experimental values. Two main, different sets have been used to simulate samples with one or two distinct grain species. For the samples with one grain specie we generated two random gaussian distributions for radii and inter-granular distances characterised by mean values  $\bar{R}$  and  $\bar{\sigma}$ , whereas for the samples with two grain species we generated two random gaussian distributions for radii and one for the inter-granular distances characterised by mean values  $\bar{R}_1, \bar{R}_2$  and  $\bar{\sigma}$ . The closest grains to two opposite sides of the boxes constituted the injection (source) and ejection (drain) electrodes. In correspondence of a given temperature  $T < T_c$  a graph (or network - the two terms used interchangeably in this paper) was determined by the association of a node to each grain and a weighted edge to each couple of activated grains (3). Dead-end paths were then removed by subsequent searches for simple nodes (i.e. the ones insisted on by just one edge) [11]. This procedure has no effect on - and returns no information about - isolated rings and isolated clusters of multiple-edged grains. If such information are needed, equivalent, reduced graphs made only of active grains, presenting no rings and no isolated cluster can be obtained by a simple variant of the method illustrated in [11] for percolation paths, modifying the definition of exhausted edges and a single point of the sending process (*sect.* 6.1), such that:

- An edge traversed in **either** direction will be indicated as an exhausted edge;

1. (a) Edges can be traversed no more than once in **either** direction.

As soon as there are no more active *Tnodes*, the nodes and edges in the sets of the complete and deferred paths constitute the desired reduced graphs. It is to notice that this variant runs much faster than the original, due to the reduction in the number of propagating path-messages (edges can be now crossed just once) and to the lack of the reconstruction process from deferred paths.

### 4 Critical current computation

Critical currents computations can be performed both on reduced and larger graphs, being the computational times the only difference. From the starting graph, a weighted network can be derived so that

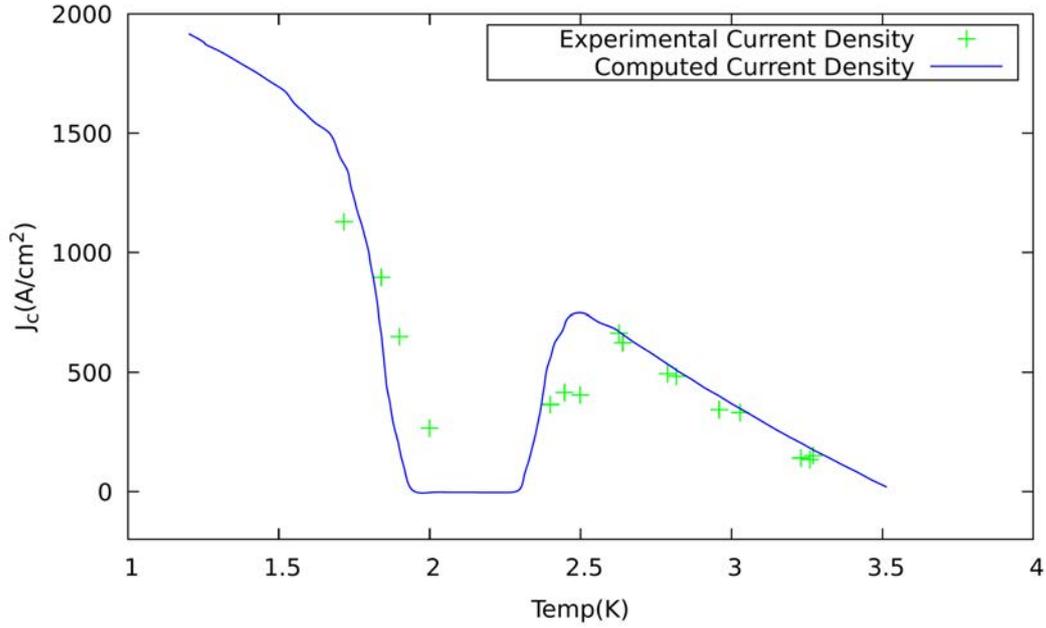


Figure 2: Experimental (+) and numerical results for the critical current density in a single-peak C—Si—W granular superconductor.

the determination of the critical current through the system can be obtained solving a maximum net flow problem:

given a network  $\mathcal{G} = (\mathcal{N}, \mathcal{A})$  with non-negative capacities  $u_{i,j}$  associated to each arc  $(i, j) \in \mathcal{A}$ , be  $s, t \in \mathcal{N}$  the *source* and *terminal (drain)* node, respectively, then a flow  $v$  is to be determined such that:

$$\begin{cases} \mathbf{max} & v \\ \sum_{\{j:(i,j) \in \mathcal{A}\}} x_{i,j} - \sum_{\{j:(i,j) \in \mathcal{A}\}} x_{j,i} = & \begin{cases} 0, & \text{if } i \in \mathcal{N} - \{s, t\}; \\ -v, & \text{if } i = t \end{cases} \\ 0 \leq x_{i,j} \leq u_{i,j} & \end{cases} \quad (5)$$

The correspondence between our granular system and the network was built associating each grain to a node in  $\mathcal{N}$ , each pair of connected grains to an arch  $(i, j) \in \mathcal{A}$ , setting  $u_{i,j} = I_c^{i,j}$  (4) and introducing two fictitious nodes  $s$  and  $t$  linked to the injection and drain electrodes, respectively, by arches of infinite capacity. The critical current flow has then been computed by means of the well known Ford-Fulkerson method [22].

## 5 Results

The results of the simulations for the critical current  $I_c$  are shown in Figures 2 and 3. They clearly show the good agreement, both qualitative and quantitative, between the experimental data, the results from the numerical simulation for the critical super-current, the positions of re-entrant peaks [10, 12].

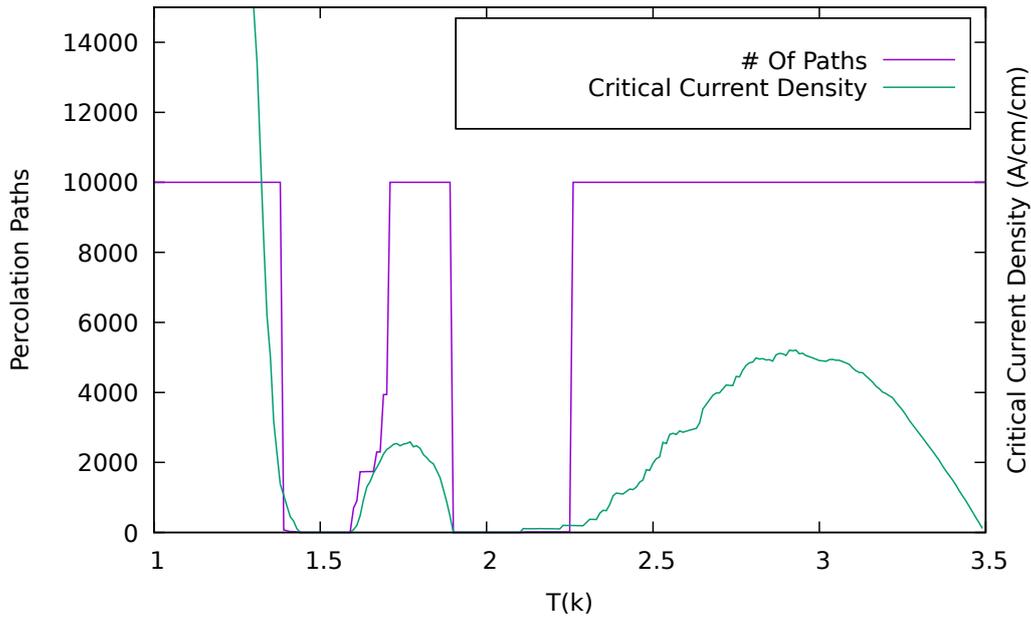


Figure 3: Simulations results for percolation paths and critical current in a two-peak C—Si—W granular superconductor. The number of processed percolation paths is reported up to a cut-off value of 10000 (flat red line).

## 6 Conclusions

Based on the physical model proposed by Chudinov et al. [10], a numerical method has been presented for the computational determination of the critical currents in metallic granular superconductors. The method proved its effectiveness and confirmed the validity of the implemented model showing strict correspondences with experimental results.

## 7 Acknowledgements

The experimental measurements were carried out in the Low Temperature Laboratory at the Physics Department of the Università di Camerino. The simulation software -written in Fortran90- was developed there, too. The calculations were performed using the facilities, services and compilers available at the ENEAGRID infrastructure (Italy), on CRESCO4 cluster, 800GB of disk storage has been granted on its GPFS file system.

## References

- [1] Yu. G. Morozov, I. G. Naumenko, and V. I. Petinov. *Sov. J. Low Temp. Phys.*, 2:484, 1976.
- [2] B. I. Belevtsev, Yu. F. Komnik, and A. V. Fomin. Superconductivity effects near the metal-insulator transition in granular indium films. *J. Low Temp. Phys.*, 69:401–417, 1987.
- [3] A. F. Hebard and M. A. Paalanen. Pair-breaking model for disorder in two-dimensional superconductors. *Phys. Rev. B*, 30:4063–4066, Oct 1984.

- [4] M. Kunchur, P. Lindenfeld, W. L. McLean, and J. S. Brooks. Absence of superconductivity in metallic granular aluminum. *Phys. Rev. Lett.*, 59:1232–1235, Sep 1987.
- [5] S. Kobayashi, Y. Tada, and W. Sasaki. *Physica B*, 107:129, 1981.
- [6] A. M. Glukhov, I. M. Dmitrienko, and A. A. Shablo. Conductivity and superconducting properties of granulated sn-ge films. *Sov. J. Low Temp. Phys.*, 9:13, 1983.
- [7] R. C. Dynes, J. P. Garno, and J. M. Rowell. Two-dimensional electrical conductivity in quench-condensed metal films. *Phys. Rev. Lett.*, 40:479–482, Feb 1978.
- [8] Alice E. White, R. C. Dynes, and J. P. Garno. Destruction of superconductivity in quench-condensed two-dimensional films. *Phys. Rev. B*, 33:3549–3552, Mar 1986.
- [9] H. M. Jaeger, D. B. Haviland, A. M. Goldman, and B. G. Orr. Threshold for superconductivity in ultrathin amorphous gallium films. *Phys. Rev. B*, 34:4920–4923, Oct 1986.
- [10] S M Chudinov, R Ferretti, S Fusari, G Mancini, and S Stizza. Reentrant superconductivity mechanisms in amorphous carbon-silicon films containing tungsten. *Physical Review B*, 62:12516–12521, 10 2000.
- [11] G. Mancini. Percolation paths through simulated granular superconductors. *High Performance Computing on CRESCO infrastructure: research activities and results 2018*, pages 97–102, December 2019. ISBN: 978-88-8286-390-6.
- [12] S M Chudinov, G Mancini, M Minestrini, R Natali, S Stizza, and A Bozhko. Critical current in granular superconductor c-si-w with peak-type re-entrant superconductivity. *Journal of Physics: Condensed Matter*, 14(2):193–209, dec 2001.
- [13] B. I. Belevtsev. Superconductivity and localization of electrons in disordered two-dimensional metal systems. *Sov. Phys. Usp.*, 33:36–54, 1990.
- [14] L. N. Bulaevskii, A. I. Buzdin, Kulic M. L., and S. V. Panjukov. *Advances in Physics*, 8:115, 1982.
- [15] R. Balducci and R.S. Pearlman. Efficient exact solution of the ring perception problem. *Journal of Chemical Information and Computer Sciences*, 34(4):822–831, 1994.
- [16] Mancini G. A redundancy eliminating approach to linearly independent rings selection in the ring perception problem. *Computer Physics Communications*, 143(2):187–197, 2002.
- [17] Rybicki J., Bergmański G., and Mancini G. A new program package for investigation of medium-range order in computer-simulated solids. *Journal of Non-Crystalline Solids*, 293-295:758–763, 2001. 8th Int. Conf. on Non-Crystalline Materials.
- [18] B. Abeles. Effect of charging energy on superconductivity in granular metal films. *Phys. Rev. B*, 15:2828–2829, Mar 1977.
- [19] E. E. Šimánek. Effect of charging energy on transition temperature of granular superconductors. *Solid State Communications*, 31(6):419–421, 1979.
- [20] K.B. Efetov. Phase transition in granulated superconductors. *Zh. Éksp. Teor. Fiz. [Journal of Experimental and Theoretical Physics]*, 78:2017–2032, May 1980.
- [21] V. Ambegaokar and A. Baratoff. Tunneling between superconductors. *Phys. Rev. Lett.*, 10:486–489, Jun 1963.

- [22] L.R. Ford Jr. and D.R. Fulkerson. Maximal flow through a network. *Canad. J. Math.*, 8:399–404, 1956.

# SYNCHRONIZATION IN HETEROGENEOUS ENSEMBLES OF OSCILLATORS: EXPLOSIVE TRANSITIONS AND DYNAMICAL CONTROL

Ricardo Gutiérrez\*

*Complex Systems Interdisciplinary Group (GISC), Department of Mathematics  
Universidad Carlos III de Madrid, 28911 Leganés, Madrid, Spain*

**ABSTRACT.** This document summarizes recent research on the synchronization of networks of heterogeneous oscillators. The work is to a large extent numerical, as it involves the integration of systems of ordinary differential equations and the computation of Lyapunov exponents, as well as synchronization measures. Part of these numerical methods were implemented in standard C codes that I wrote, compiled using gcc, and ran at CRESCO/ENEAGRID High Performance Computing infrastructure. Two papers are discussed. The first one deals with the nature of the transition between incoherent motion and synchronous dynamics in networks of cooperative and competitive oscillators. Its aim is to clarify the mechanisms underlying the phenomenon of explosive synchronization, whereby a system becomes abruptly synchronized under a tiny change of a control parameter. The second presents a pinning method whose aim is to impose a given dynamics on a freely evolving networked system. We illustrate this idea, which is of wide applicability, with a trophic web network of 12 interacting mammalian species inspired by European and North-American ecosystems.

## 1 Explosive synchronization under competition and cooperation

In the standard models of synchronization a population of oscillators undergo a continuous phase transition from incoherence to partial synchronization as the coupling strength reaches a critical value. Under certain conditions, however, the transition becomes explosive: an abrupt onset of synchronization follows an infinitesimal change in the coupling strength and hysteresis loops are observed, as in a thermodynamic first-order phase transition. The mechanism behind the onset of explosive phase transitions has been the object of much recent interest. In most of the scenarios considered, a restrictive condition correlating the natural frequencies to the topology of connections prevents the clustering of similarly evolving nodes that underlies the usual continuous transition. However, a simple adaptive mechanism is already sufficient for the emergence of this intriguing phenomenon: in Ref. [1] we show that the coexistence of cooperative and competitive oscillators changes the nature of the phase transition.

Specifically, we consider an adaptive Kuramoto-like model comprising  $N$  phase oscillators. The instantaneous phase of oscillator  $i$  is denoted by  $\theta_i$  ( $i = 1, 2, \dots, N$ ), which evolves in time as follows

$$\dot{\theta}_i = \omega_i + \lambda \alpha_i \sum_{j=1}^N A_{ij} \sin(\theta_j - \theta_i). \quad (1)$$

---

\*Corresponding author. E-mail: rigutier@math.uc3m.es.

Here,  $\omega_i$  is the natural frequency of the oscillator,  $\lambda$  is the coupling strength and  $A_{ij}$  is the adjacency matrix ( $A_{ij} = 1$  if oscillators  $i$  and  $j$  are coupled, otherwise it is zero). The degree of a given oscillator  $i$  is the number of neighbors it has,  $k_i = \sum_{j=1}^N A_{ij}$ . The novelty of the model lies in its coupling strength being controlled by the Kuramoto global order parameter  $R$  through the local variable  $\alpha_i$  in Eq. (1). The order parameter satisfies  $Re^{i\Phi} = \frac{\sum_{j=1}^N \sum_{k=1}^N A_{jk} e^{i\theta_k}}{\sum_{j=1}^N k_j}$ , where  $\Phi$  is an average phase, and  $R$  quantifies the global synchronization ( $0 \leq R \leq 1$ ). The network is partitioned into two sets of nodes: there are  $N\rho$  competitors, with  $\alpha_i = 1 - R$ , which decouple from their neighbors as global synchrony increases, and  $N(1 - \rho)$  cooperators, for which  $\alpha_i = R$ , which tend to couple more strongly to their neighbors as the network becomes globally more synchronized. The population balance is controlled by the competition fraction  $\rho$ . The case  $\rho = 0$  has been previously shown to lead to explosive synchronization, while for  $\rho = 1$  the ability to form synchronization clusters is enhanced and the transition is continuous.

We explore numerically the synchronization transition and its dependence on the competition fraction  $\rho$  in networks of oscillators with natural frequencies  $\{\omega_i\}$  drawn from a uniform probability distribution in  $[-1, 1]$ . These are Erdős-Rényi (ER) networks and Barabási-Albert (BA) scale-free networks, which are paradigmatic examples of homogeneous and heterogeneous networks, respectively, with an average degree  $\langle k \rangle = 12$ . In Fig. 1 we show the  $R$  vs  $\lambda$  curves for the forward and backward transition to coherence ( $\lambda$  is increased from 0 to 0.5, and then decreased back again) for two values of the competition fraction  $\rho$  in an ER network. The presence of a vast majority of competitors,  $\rho = 0.8$ , leads to a continuous transition. By contrast, the existence of a large fraction of cooperators,  $\rho = 0.2$ , gives rise to an explosive transition with hysteresis effects.

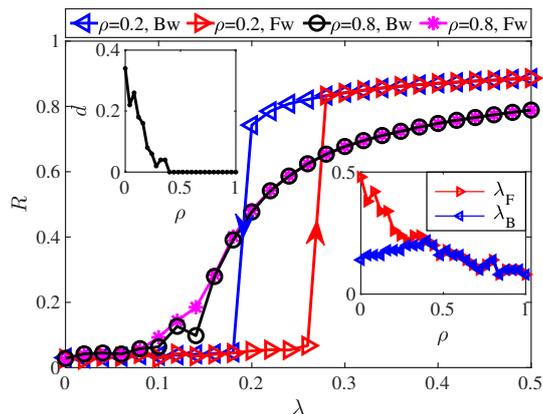


Figure 1: Forward and backward synchronization transitions on an ER network of  $N = 1000$  and mean degree  $\langle k \rangle = 12$ . The insets show the transition points  $\lambda_B$  and  $\lambda_F$  and the distance between them  $d$ .

A continuum mean-field approximation, whereby  $r_i \rightarrow R$  and  $\Phi_i \rightarrow \Phi$ , yields an analytical solution. For  $\rho \in [0, 1]$ , let  $F(k, \omega, \lambda, \alpha, R) = \int_{|\omega| < \lambda \alpha R k} h(k, \omega) k \sqrt{1 - \left(\frac{\omega}{\lambda \alpha R k}\right)^2} d\omega dk$ , where  $h(k, \omega)$  is the joint distribution of frequencies  $\omega$  and degree  $k$ , one gets

$$R = \frac{1}{\langle k \rangle} (\rho F(k, \omega, \lambda, R, R) + (1 - \rho) F(k, \omega, \lambda, 1 - R, R)). \quad (2)$$

A solution for an ER network with  $\langle k \rangle = 12$  and  $\rho = 0.3$  is plotted in Fig. 2, where the existence of an unstable solution is responsible for explosive synchronization. We conclude that the transition to synchronization changes from explosive to continuous by simply modifying the population balance.

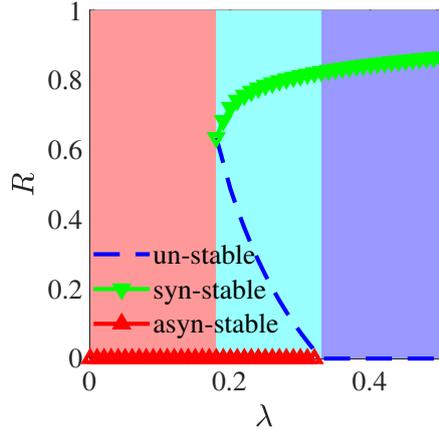


Figure 2: Mean-field solution for ER network when  $\rho = 0.3$ . Green inverted-triangle line: synchronization; red triangle line: incoherence; blue dashed line: unstable solution.

## 2 A pinning method for steering networks toward desired dynamics

In Ref. [2], we propose a procedure for imposing specific dynamic evolutions compatible with the equations of motion on a networked system. Studying how to steer the dynamics of such systems is of great interest in many areas of science. In contrast to previous methods, which rely of stringent simplifying assumptions, this one does not impose any restrictions on the dynamics or the interactions.

We consider a two-layer network: one layer is the *slave* layer, which is the original network over which one wants to impose the desired dynamics, and the other is the *master* layer, identical to the slave, but realizing the desired dynamics. Our control method consists in establishing more and more directed inter-layer links from nodes in the master layer to their counterparts in the slave layer. At each step the selected node is the one whose pinning causes the most rapid approach towards inter-layer synchronization. Before the pinning starts, both master and slave layers evolve spontaneously as

$$\dot{\mathbf{r}}_i = \mathbf{f}(\mathbf{r}_i) + \sigma_1 \sum_{j=1}^N D_{ji}(z_j^2 - z_i^2) = \mathbf{f}(\mathbf{r}_i) + \sigma_1 \sum_{j=1}^N \mathcal{L}_{ji} z_j^2. \quad (3)$$

where  $D_{ji} = 1$  if there is a directed link from node  $j$  to node  $i$ , and is zero otherwise. The graph can be also represented by the Laplacian matrix  $\mathcal{L}_{ji} = D_{ji} - k_{\text{in},i} \delta_{ji}$ , where the in-degree of node  $i$  is  $k_{\text{in},i} = \sum_j D_{ji}$ . The local dynamics of node  $i$  is given by  $\mathbf{f}(\mathbf{r}_i)$ , and  $\sigma_1$  is the intra-layer coupling strength. When the control procedure starts, each node  $i$  in the master network keeps evolving according to the dynamics in Eq. (3),  $\dot{\mathbf{r}}_i^M = \mathbf{f}(\mathbf{r}_i^M) + \sigma_1 \sum_j \mathcal{L}_{ji} (z_j^M)^2$ . In the slave layer dynamics, instead, one has to consider an additional term which accounts for the inter-layer coupling from the master layer (we take a linear coupling through the  $y$  variable). One has  $\dot{\mathbf{r}}_i^S = \mathbf{f}(\mathbf{r}_i^S) + \sigma_1 \sum_j \mathcal{L}_{ji} (z_j^S)^2 + \sigma_2 \chi_i (y_i^M - y_i^S)$ . Here  $\chi_i$  is a binary variable that is one if there is a link coupling node  $i$  in the master layer to node  $i$  in the slave layer and is zero otherwise, and  $\sigma_2$  is the inter-layer coupling strength.

We first explore our method on random graphs of  $N = 50$  nodes with the in-degree  $k_{\text{in}}$  (i.e. the number of links pointing to a given node) and the out-degree  $k_{\text{out}}$  (i.e. the number of links emanating from a given node) uniformly distributed in  $\{5, 6, \dots, 45\}$ . Each node is a chaotic Rössler oscillator coupled quadratically to its neighbors. The results are shown in Fig. 3. Two observables are employed to characterize the synchronization between master and slave as additional inter-layer links are established in

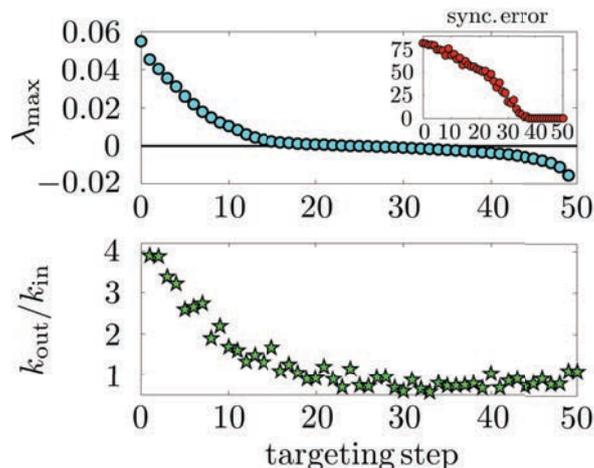


Figure 3: Steering the dynamics of networks of  $N = 50$  Rössler oscillators. (Top). Lyapunov exponent  $\lambda_{\max}$  and synchronization error at each targeting step. (Bottom) Influence  $k_{\text{out}}/k_{\text{in}}$  of newly-pinned nodes. A 4th-order Runge-Kutta method is used for the integration of the  $2 \times 3 \times N = 300$  ODEs.

successive targeting steps. One is the maximum Lyapunov exponent,  $\lambda_{\max}$ . The other is the synchronization error, which is the time average of the Euclidean distance between the states of the master layer and the slave layer. A remarkable correspondence is found between the targeting sequence and the node ranking in terms of the *influence index*  $k_{\text{out}}/k_{\text{in}}$ , as shown in the lower panel of Fig. 3. This index is large when a node has a privileged position for influencing others, while receiving very little influence from the rest of the network. A strong correlation is also observed in BA and ER networks.

We then consider a 12-species model of a mammalian food web inspired by holarctic ecosystems. We illustrate how to control such a trophic web through the pinning of a limited number of species. By controlling the populations of a relatively small number of species, one can make the system abandon its spontaneous evolution from its (typically uncontrolled) initial state towards a target dynamics, or periodically control it so as to make the populations evolve within stipulated bounds. The key information on the trophic web is contained in the sequence of pinned nodes, as this reveals which are the species whose population one must preserve or modify in order to maintain a desired dynamics or disrupt an undesired one. The pinning actions rapidly bring the slave system into the desired dynamical regime.

In conclusion, we propose a versatile method for steering networks toward desired dynamics, which has shown to be valuable for unveiling correlations between node controllability and topological properties. Most importantly, the method is of practical value for the control of systems that do not satisfy the highly idealized requirements of network control methods in the literature.

## References

- [1] X. Dai, X. Li, R. Gutiérrez, H. Guo, D. Jia, M. Perc, P. Manshour, Z. Wang, and S. Boccaletti. Explosive synchronization in populations of cooperative and competitive oscillators. *Chaos, Solitons & Fractals*, 132:109589, 2020.
- [2] R. Gutiérrez, M. Materassi, S. Focardi, and S. Boccaletti. Steering complex networks toward desired dynamics. *Scientific Reports*, 10:20744, 2020.

# THE ENEA-REG SYSTEM, A MULTI-COMPONENT REGIONAL EARTH SYSTEM MODEL

Alessandro Anav<sup>1\*</sup>, Adriana Carillo<sup>1</sup>, Massimiliano Palma<sup>1</sup>, Maria Vittoria Struglia<sup>1</sup>, Ufuk Utku Turuncoglu<sup>2</sup>, Gianmaria Sannino<sup>1</sup>,

<sup>1</sup> *ENEA SSPT-MET-CLIM, Italian National Agency for New Technologies, Energy and Sustainable Economic Development, CR Casaccia, Rome, Italy*

<sup>2</sup> *National Center for Atmospheric Research (NCAR), Boulder, CO, USA*

**ABSTRACT.** We present a newly developed regional Earth system model. The ENEA-REG system is made up of two interchangeable regional climate models as atmospheric components (RegCM and WRF), a river model (HD), and an ocean model (MITgcm); processes taking place at the land surface are represented by land surface models embedded in the atmospheric components.

We performed a multidecadal hindcast simulation for the period 1980-2013 over the Med-CORDEX region. Overall, our results show that the atmospheric components are able to correctly reproduce both large-scale and local features of the Euro-Mediterranean climate. Nevertheless, when WRF is used as the atmospheric component of the Earth system, the performances of the ocean model are remarkably better compared with the RegCM version.

## 1 Introduction

The Mediterranean basin is a complex region, characterized by the presence of pronounced topography and a complex land-sea distribution including a considerable number of islands and several straits. These features generate strong local atmosphere–sea interactions leading to the formation of intense local winds, like Mistral, Etesian and Bora which, in turn, dramatically affect the Mediterranean ocean circulation. Given the relatively fine spatial scales at which these processes take place, the Mediterranean basin provides a good opportunity to study regional climate, with a special focus on the air-sea coupling. For these reasons, regional coupled models have been developed and used to study both present and future Mediterranean climate system; these models, depending on their complexity, include several physical components of the climate system, like atmosphere, ocean, land surface, rivers and biogeochemistry (both for land and ocean). Since the last two decades, an increasing number of studies have been performed over the Mediterranean basin and nowadays there is a coordinated effort for producing hindcast and future simulations over this region using regional coupled climate models sharing some common protocols [1]. In particular, the Coordinated Regional Climate Downscaling Experiment (CORDEX) was designed to produce, worldwide, high-resolution regional climate simulations through a coordinated experiment protocol ensuring that model simulations are

carried out under similar conditions facilitating thus the analysis, intercomparison, and synthesis of different simulations. In the framework of the CORDEX program, regional climate model simulations dedicated to the Mediterranean area belong to the Med-CORDEX initiative [1].

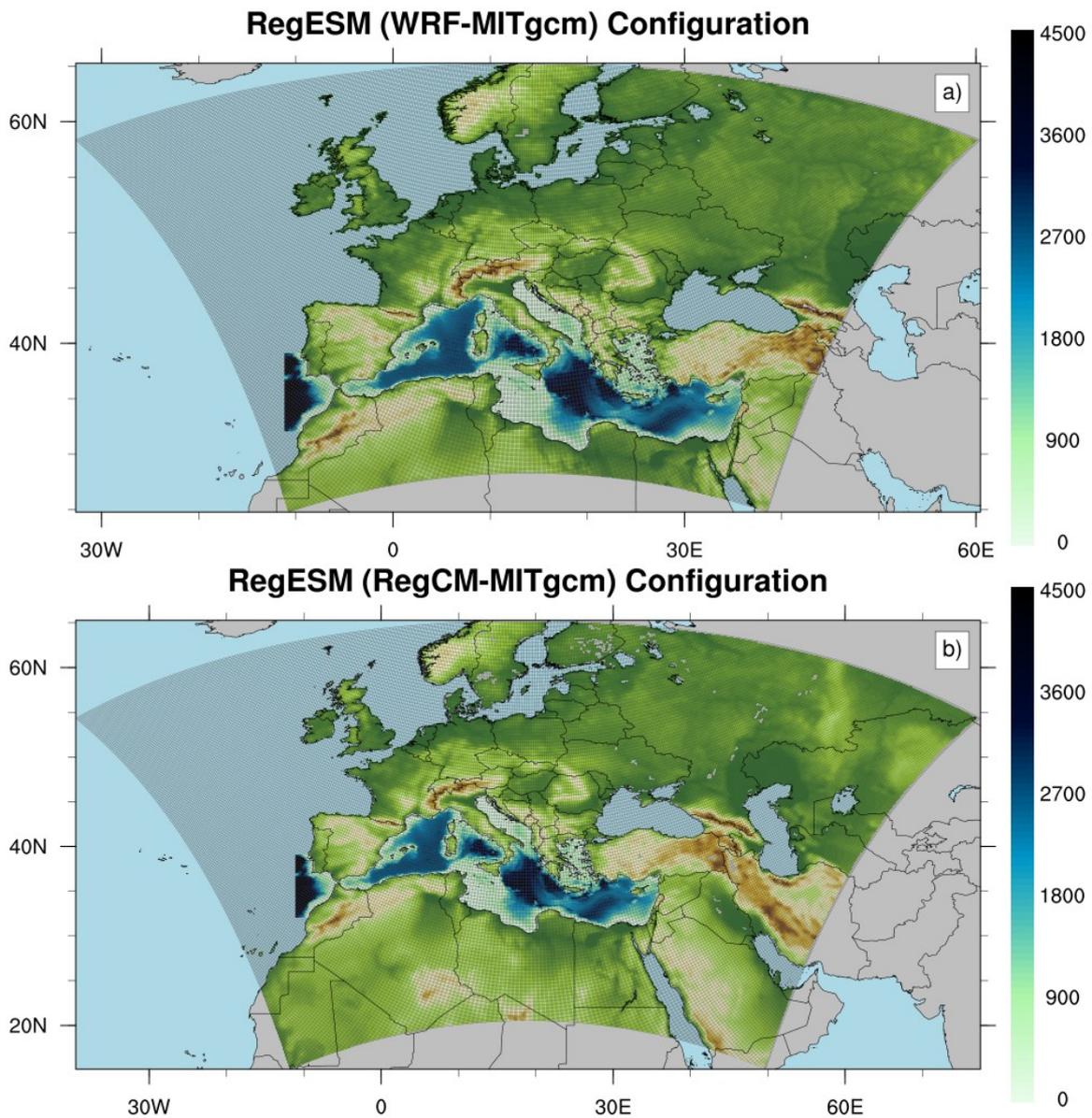
## 2 Setup

The ENEA-REG regional Earth system model has the capability to include several model components (atmosphere, river routing, ocean, wave) to allow different modeling applications. For each simulation, the components of the modeling system can be easily enabled or disabled via the driver's configuration file. In addition, the modeling framework also supports plugging new earth system sub-components (e.g. atmospheric chemistry, sea ice, ocean biogeochemistry) with minimal code changes through its simplified interface, which is called “cap”. The National United Operational Prediction Capability (NUOPC) cap is a Fortran module that serves as interface to a model when it is used in a NUOPC-based coupled system; it is a small software layer that sits on top of a model code, making calls into it and exposing model data structures in a standard way [2].

In the simulations performed on CRESCO6, the modeling system is configured to include three components: a regional atmospheric climate model, a regional ocean model and an hydrological model. The driver used to glue, regrid and exchange data among the three components of ENEA-REG modeling system is RegESM [2]. The driver employs the Earth System Modeling Framework (ESMF) library (version 7.1) and the NUOPC layer to connect and synchronize each model component and perform interpolation among different horizontal grids [2]. While the ESMF library deals with interpolation and regridding of exchanged fields, the NUOPC layer simplifies common tasks of model coupling like component synchronization and run sequence by providing additional wrapper layer between coupled model and ESMF framework. It also allows defining different coupling time intervals among the components to reproduce fast and slow interactions among the model components.

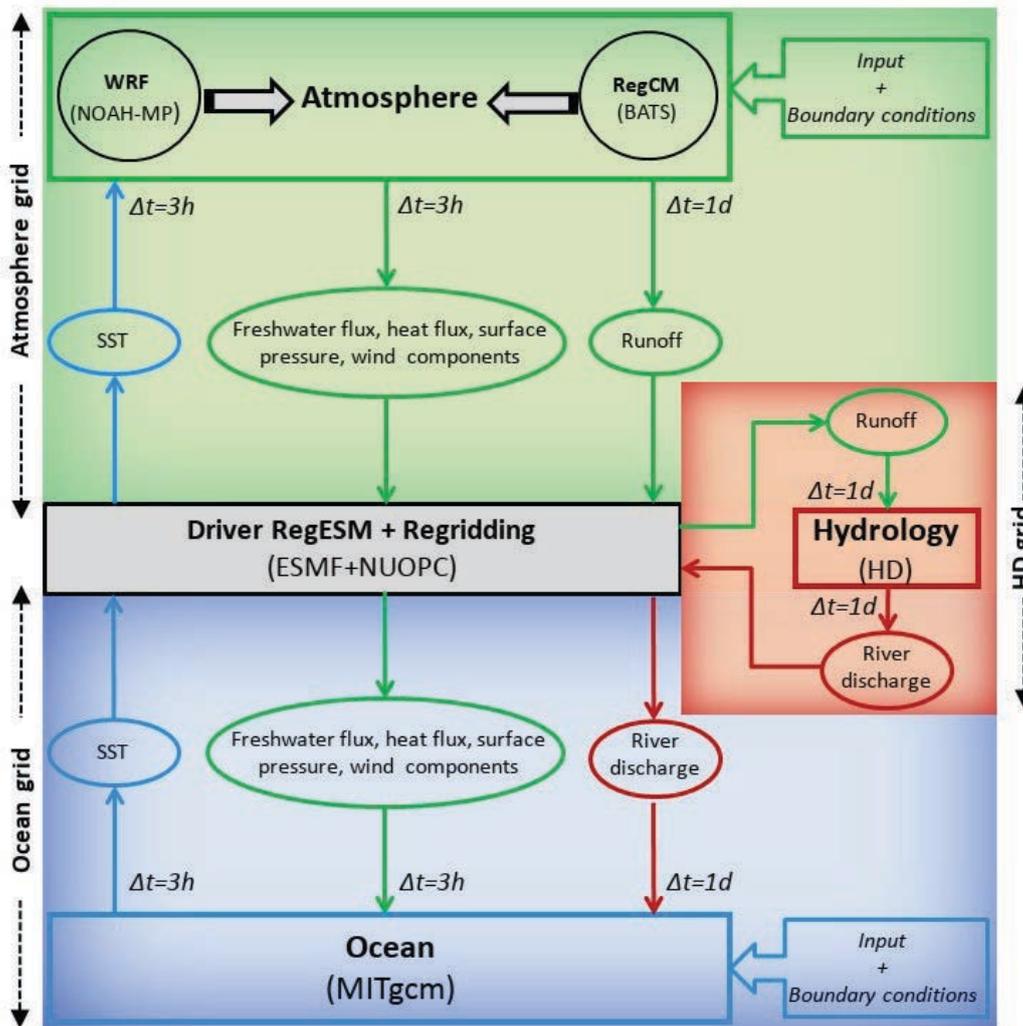
The ENEA-REG regional Earth system model is made up of two interchangeable atmospheric components: the Weather Research and Forecasting (WRF) model and the REGional Climate Model (RegCM). WRF is a limited-area, non-hydrostatic, terrain-following eta-coordinate mesoscale model developed by the NCAR/MMM (National Center for Atmospheric Research, Mesoscale and Microscale Meteorology division). The model domain is projected on a Lambert conformal grid with a horizontal resolution of 15 km and with 35 vertical levels extending from land surface up to 50 hPa (figure 1a). The other atmospheric component is RegCM (version 4.5) a hydrostatic, compressible, sigma-p vertical coordinate model; the model domain is projected on a Lambert conformal grid with a horizontal resolution of 20 km and with 23 vertical levels extending from land surface up to 50 hPa (figure 1b).

The ocean component of the ENEA-REG system is the Massachusetts Institute of Technology General Circulation Model (MITgcm version c65; Marshall et al., 1997). The MITgcm solves both the hydrostatic and nonhydrostatic Navier-Stokes equations under the Boussinesq approximation for an incompressible fluid with a spatial finite-volume discretization on a curvilinear computational grid using the  $z^*$  rescaled height vertical coordinate. The model domain has a horizontal resolution of  $1/12^\circ$ , corresponding to  $570 \times 264$  grid points, and covers the entire Mediterranean Sea with the boundary conditions in the Atlantic Ocean (Figure 1). In the vertical the model is discretized using 75 unevenly spaced Z-levels going from 1 m at the surface to about 300 m in the deepest part of the basin.



**Fig.1:** Different domains of the ENEA-REG system, with green shading representing the topography of the atmospheric models (i.e. WRF and RegCM, solid grey lines indicate the computational domain) and blue shading the bathymetry of the ocean component.

In the experiments presented here, the atmospheric model retrieves sea surface temperature (SST) from the ocean model (where grids are overlapped), while the ocean model collects surface pressure, wind components, freshwater (evaporation-precipitation, i.e. E-P) and heat fluxes from the atmospheric component. Similarly, the hydrological model uses surface and sub-surface runoff simulated by the atmospheric component to compute the river drainage and exchanges this field with the ocean component to close the water cycle (figure 2). The model coupling time step between ocean and atmosphere is set to 3-hours, while the coupling with the hydrological model is defined as 1-day. All the ENEA-REG components has been compiled with the intel compiler (mpiifort) and ran over CRESCO6 using 360 CPUs; the domain decomposition has been achieved with MPI.

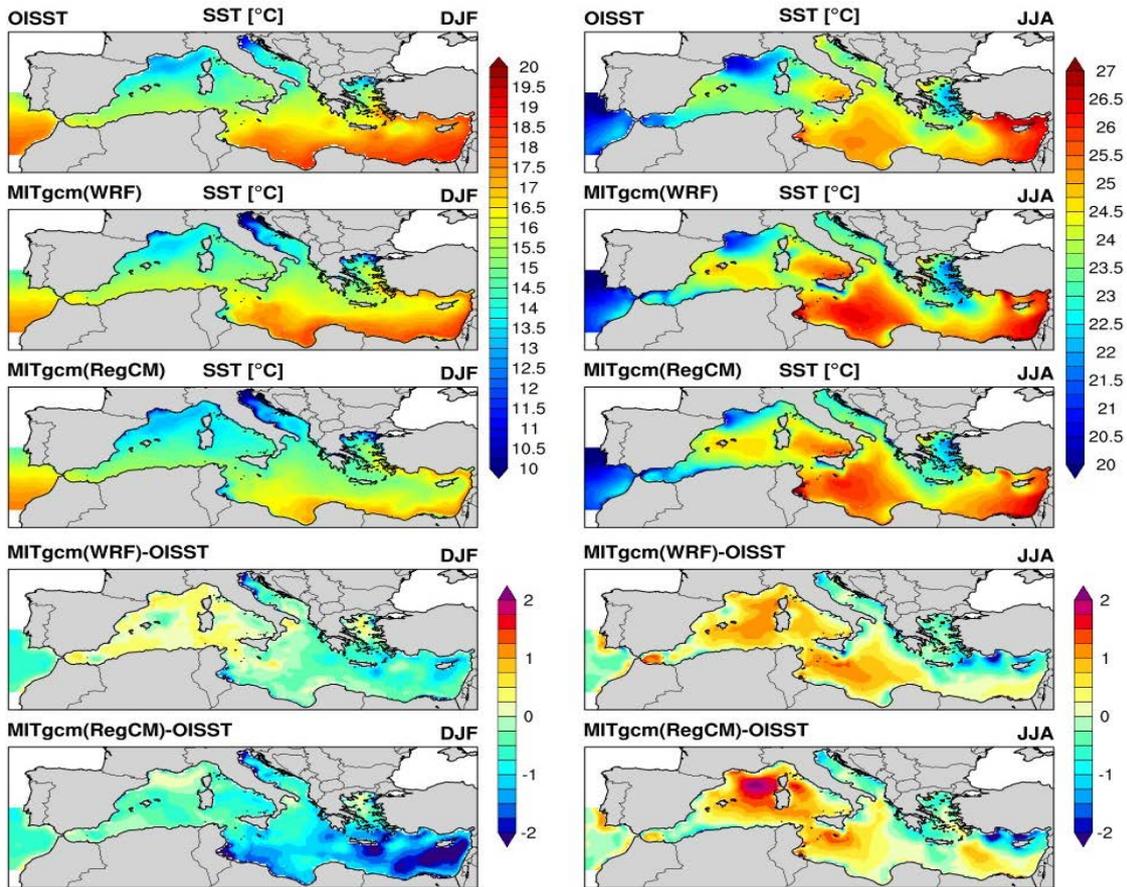


**Fig.2:** Schematic description of the ENEA-REG regional coupled model. The green block represents the atmosphere with the two components that can be selected and used (i.e. WRF and RegCM), the blue block is the ocean component (i.e. MITgcm), the red block represents the river routing component while the grey block is the ESMF/NUOPC coupler which collects, regrids and exchanges variables between the different components of the system.

### 3 Main results

The correct representation of physical processes taking place at the air-sea interface is crucial for the success of a coupled climate simulation. A first evaluation of the goodness with which these processes are simulated is given by the analysis of the ocean surface variables like Sea Surface Temperature (SST). Figure 3 shows the comparison of simulated SST with OISST reference data averaged during winter (DJF) and summer (JJA) seasons for the temporal period 1982-2013. We recall that SST, in a coupled simulation, is actually the same variable for ocean and atmosphere components (where grids overlap), and guides the thermal exchange providing an active feedback among the two components: the higher is the difference among SST and atmosphere temperature, the larger will be the heat exchange at the interface that tends to lower such difference. Looking at Figure 3, the coupled model well reproduces the OISST spatial pattern with an agreement larger than 0.99

for both the configurations and seasons. WRF-MITgcm shows moderate biases during winter ( $-0.24^{\circ}\text{C}$ ) and summer ( $0.23^{\circ}\text{C}$ ) while RegCM-MITgcm has a widespread negative bias in winter ( $-0.9^{\circ}\text{C}$ ) and a positive bias in summer ( $0.25^{\circ}\text{C}$ ), with marked spatial patterns in the eastern part of the Levantine Sea during winter and in the Sardinian Sea during summer; it should be noted that the spatial average over the entire basin reduces the bias within one degree, although the differences can be locally much more relevant, especially in the RegCM-MITgcm configuration.



**Fig.3:** Seasonal winter (DJF) and summer (JJA) spatial pattern (upper three panels) and bias (lower two panels) of sea surface temperature (SST [ $^{\circ}\text{C}$ ]) as simulated by the coupled model using the two atmospheric components as forcing (i.e. WRF and RegCM) and OISST dataset between 1982 and 2013.

[1] Ruti, P. M., Somot, S., Giorgi, F., Dubois, C., Flaouas, E., Obermann, A., Dell'Aquila, A., Pisacane, G., Harzallah, A., and Lombardi, E.: MED-CORDEX initiative for Mediterranean climate studies, *Bulletin of the American Meteorological Society*, 97, 1187-1208, 2016.

[2] Turuncoglu, U. U.: Toward modular in situ visualization in Earth system models: the regional modeling system RegESM 1.1, *Geoscientific Model Development*, 12, 2019.

# A SIMPLE PROCEDURE TO SIMULATE A SMALL ANGLE X-RAY SCATTERING EXPERIMENT

E. Burrese<sup>1</sup>, L. Tapfer<sup>2</sup>

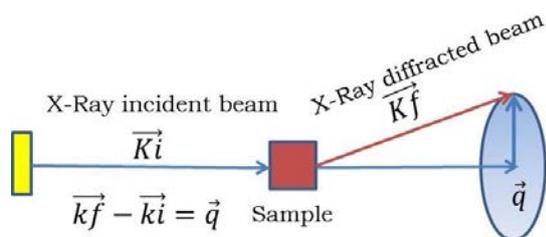
<sup>1</sup>ENEA, SSPT-PROMAS-MATAS, Brindisi Research Center, S.S.7 APPIA km 706 Brindisi <sup>1</sup>

<sup>2</sup>ENEA, SSPT-PROMAS, Brindisi Research Center, S.S.7 APPIA km 706 Brindisi

**ABSTRACT.** We report on a primary version of an algorithm for calculation of the low angle scattering curve for nanoparticles, clusters of nanoparticles, biological molecules or in general for amorphous structures. Currently, the code written in Fortran is suitable for an arbitrary atomic arrangement. The procedure has been applied for small angle scattering intensity profiles of clusters of silicon atoms and cadmium sulphide atoms of different size.

## 1 Introduction

Small Angle X-Ray Scattering (SAXS) is a materials characterization technique based on recording, by means of an opportune detector, the intensity of the X-ray photon beam scattered after interaction with a sample and deviated at low angle with respect to the incident beam direction. The deflection angle is small enough to obtain information about shape, size and other geometrical parameters of the sample but not about its crystalline structure. Indeed, in the case of larger deviation angles crystalline structure information can be obtained and the experimental setup is known as WAXS (Wide Angle X-Ray Scattering). For SAXS measurements a minimal experimental setup is shown in Figure 1. Generally, SAXS experiments are performed by means of a synchrotron radiation but also conventional X-Ray sources which are present in many X-ray laboratories are used for this scope.



**Fig. 1:** A schematic representation of the SAXS experiment. X-Ray incident beam  $\vec{k}_i$  interacts with a sample and it is scattered in order to have  $\vec{k}_f - \vec{k}_i = \vec{q}$ , where  $\vec{q}$  is the scattering vector and is measured for small values.

The modulus of the scattering vector  $|\vec{k}_f - \vec{k}_i| = 4\pi \frac{\sin \theta}{\lambda}$ , where  $\theta$  is the Bragg angle and  $\lambda$  is the X-Ray wavelength of the, is of the order of  $\frac{2\pi}{d}$  which defines the resolution limit: for large values of  $d$  (for example  $d$  size of a nanoparticle) low values of scattering vector were obtained (the “small angle”).

<sup>1</sup> Corresponding author. E-mail: emiliano.burrese@enea.it

The theory for the scattering at low angle is equivalent to that for the scattering by non-crystalline forms of matter with a random orientation of the atoms, such as gases, liquids, amorphous solids, nanocrystals but also crystalline powders and biological molecules[1].

The general equation of the scattering of amorphous materials made up of N atoms is given by

$$I = \sum_i^N \sum_j^N f_i f_j e^{i\frac{2\pi}{\lambda}(\vec{k}_f - \vec{k}_i) \cdot \vec{r}_{ij}} \quad (1)$$

where  $\vec{r}_{ij}$  is the distance between atom pairs,  $\lambda$  is the X-ray wavelength and  $f_j$  is the atomic scattering factor of atom j. The equation (1) can be rearranged as

$$I = \int d\vec{r}_1 \int d\vec{r}_2 \rho(\vec{r}_1) \rho(\vec{r}_2) e^{i\frac{2\pi}{\lambda}(\vec{k}_f - \vec{k}_i) \cdot (\vec{r}_1 - \vec{r}_2)} \quad (2)$$

by inserting an atomic density function  $N = \int \rho(\vec{r}) d\vec{r}$  for the coordinates of the electrons  $\vec{r}_1$  and  $\vec{r}_2$  [2]. Now by introducing a function  $p(r)$  which correlates the radial distances among all the atom pairs of the sample (similar to a Patterson function), the equation (2) should be resolved integrating for all the orientation in space of  $\vec{r}_{12}$ . The intensity of the scattering which comes from this last arrangement by using the modulus  $h = |\vec{k}_f - \vec{k}_i| = 4\pi \frac{\sin \theta}{\lambda}$  and is given by

$$I(h) = \int_0^\infty 4\pi r^2 p(r) \frac{\sin hr}{hr} dr \quad (3)$$

Once the function  $p(r)$  is known, it is possible to obtain the intensity as a function of the scattering in term of h by integrating, for example, equation (3) between  $r=0$  and  $r=D$ , where D is the diameter of the nanoparticles.

Some possible functions for  $p(r)$  are the pair density function (PDF) reported in Eq.(4) or the reduced PDF.

$$p(r) = \frac{1}{4\pi N r^2 \langle f \rangle^2} (\sum_i \sum_{j \neq i} f_i f_j \delta(r - r_{ij})) \quad (4)$$

where  $\langle f \rangle$  corresponds to the average scattering factor and N is the number of the atoms. If we define the radial distribution function (RDF) as  $R(r) = 4\pi r^2 p(r)$ , the intensity of scattering at low angle results to be

$$I(h) = \int_0^\infty R(r) \frac{\sin hr}{hr} dr \quad (5)$$

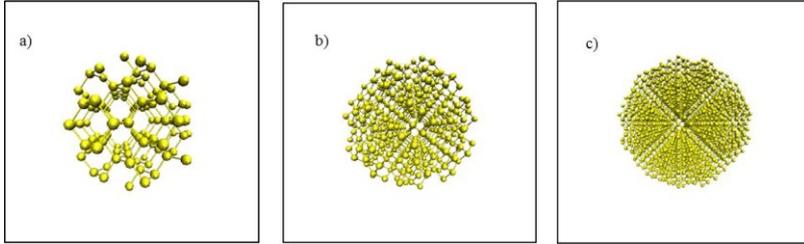
which can be solved in different ways. The algorithm has been tested both for spherical atomic cluster with one kind of atom and with more kinds of atoms. Some of these results are reported in section 2. In particular the calculations regarded silicon clusters with different size (with radius of 8 Å, 12 Å and 20 Å, respectively) and CdS clusters with radius of 12 Å and 18 Å. In addition some geometrical parameters, which are typical of the small angle treatment framework, were also calculated. Finally, in order to test the goodness of the algorithm, all these calculations were compared with respect to the scattering intensity of a sphere of the same radius and with intensity

$$I(h) = I(0) \left[ 3 \frac{\sin hR - hR \cos hR}{(hR)^3} \right]^2.$$

## 2 Results and discussion

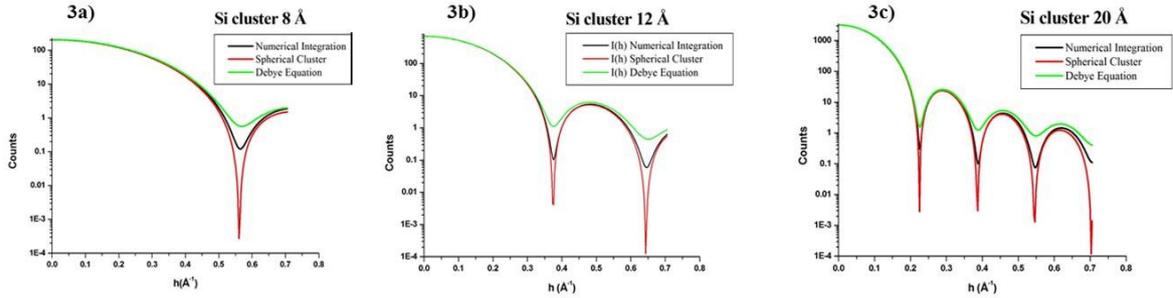
## 2.1 Silicon clusters

Figure 2 shows three clusters of spherical shape and different radius of 8 Å , 12 Å and 20 Å (Figure 2a, 2b and 2c, respectively). The atoms were arranged in an unit cell with cubic symmetry (diamond structure), however, considering that the small scattering does not collect information from crystallographic planes. The SAXS intensity in Eq.(5) has been solved by numerical integration, comparing it with the Debye equation as well as with the scattering intensity of a sphere of the same radius.



**Fig. 2:** Silicon clusters of radius 8 Å (a), 12 Å (b), 20 Å (c) containing 105, 358 and 1652 atoms, respectively.

The respective SAXS intensity curves in logarithmic scale are reported in Figure 3.



**Fig. 3:** SAXS curves of the clusters in Figure 2 calculated by using Eq.(5) (black lines), the Debye equation (green lines) and by using the equation for a sphere with the same radius (red lines).

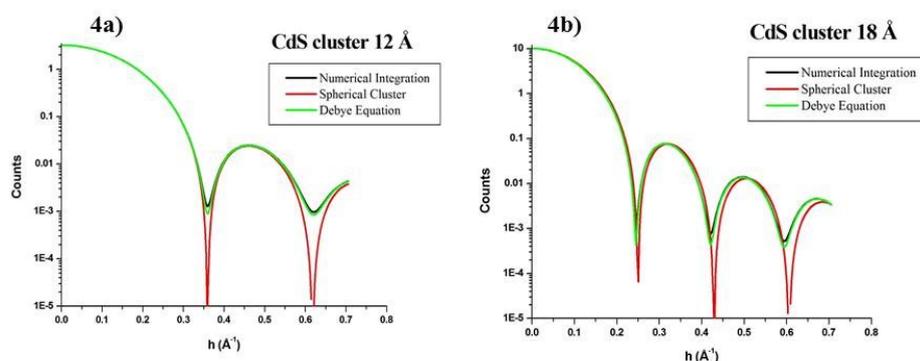
	R=8 Å	R=12 Å	R=20 Å
$R_g$ Sphere	6.198	9.295	15.491
$R_g$ Guinier Analysis	6.177	9.518	17.036

**Table 1:** Radius of gyration  $R_g$  for silicon clusters are compared with spherical  $R_g$ .

The agreement among these curves seem to be pretty good; black curve seems to describe better the oscillating behaviour of the curve as well as around the minimum. In Table 1 the Guinier analysis for the calculation of the radius of gyration is reported for three clusters with respect to the extreme case of the sphere with  $R_g = (3/5)^{1/2} R_0$ , where  $R_0$  is the radius of the sphere. In this case we can observe a greater deviation for clusters in Figure 2c.

## 2.2 Cadmium Sulfide Clusters

We analysed also SAXS curves for clusters with more than one kind of atom. Figure 4 shows SAXS curves for cluster of 12 Å and 18 Å.



**Fig. 4:** SAXS curves of the CdS clusters calculated by using equation Eq.(5) (black lines), Debye equation (green lines) and by using the equation for a sphere with the same radius (red curves).

For these samples the comparison with a sphere was still not completely satisfactory since we can observe a slight phase displacement of the oscillations calculated by Eq.(5). The same discrepancy has been obtained for Guinier analysis with gyration radius  $R_g=3.253$  Å and  $R_g=9.580$  Å for cluster of 12 Å and 18 Å, respectively. These calculations underestimate the radius of gyration calculated for spheres with the same radius.

## 3 Conclusions

In this brief report a procedure to calculate the SAXS intensity for small and medium size spherical clusters both for Silicon and CdS atoms has been presented. The results of the algorithm were compared with the scattering intensity of a sphere with analogue sizes and these results seem to be in good agreement. The procedure requires the calculation of the PDF or RDF functions, which can be obtained also by experimental techniques, and it is very useful to reconstruct the atomic structure of the sample. Moreover, the algorithm is very general and versatile, also in order to obtain the SAXS intensity curves for any atomic arrangement like amorphous body or biological molecules as well as for clusters as reported here.

The computing resources and the related technical support used for this work have been provided by CRESCO/ENEAGRID High Performance Computing infrastructure and its staff [3].

## 4 References

- [1] B.E. Warren, X-ray Diffraction Dover Publications, INC., New York (1990)
- [2] O. Gatter, O. Kratky, Small Angle X-Ray Scattering, Academic Press (1982)
- [3] F. Iannone et al., "CRESCO ENEA HPC clusters: a working example of a multifabric GPFS Spectrum Scale layout," 2019 International Conference on High Performance Computing & Simulation (HPCS), Dublin, Ireland, 2019, pp. 1051-1052, doi: 10.1109/HPCS48598.2019.9188135

# MOLECULAR DYNAMICS CALCULATIONS FOR THE DETERMINATION OF $\text{PuO}_2$ HEAT CAPACITY

Rolando Calabrese<sup>1\*</sup>

<sup>1</sup>*ENEA, Nuclear Safety and Sustainability Division, I-40129, Bologna, Italy*<sup>1</sup>

**ABSTRACT.** A new generation of fast breeder reactors (FBR) is under development with the objective of making nuclear energy more sustainable. Most promising reactor designs are loaded, at least in their early phase of deployment, with  $\text{UO}_2$ - $\text{PuO}_2$  mixed oxide fuel. Heat capacity is important for the evaluation of thermal conductivity and performance under transient conditions. If, on the one hand, the heat capacity of  $\text{UO}_2$  has been deeply studied, on the other hand, measurements on  $\text{PuO}_2$  are scarce or even lacking at high temperatures. Numerical methods such as MD calculations have been proposed to overcome the difficulties encountered in experimental measurements. However, several issues remain open such as the deviation of MD calculations from measurements in the low temperature domain, the existence of a Bredig transition, the formation of defects at high temperature and correlated increase of heat capacity. This study presents results on the heat capacity of  $\text{PuO}_2$  obtained using MD calculations based on an interatomic potential published recently.

## 1 Introduction

Next generation FBRs aim at improving nuclear energy sustainability by overcoming, among others, the limitation in uranium resources. New reactors' designs will require the use of MOX fuel with a concentration of plutonium dioxide up to 30 mol% [1]. Kato and Matsumoto agree on the fact that the thermophysical properties of MOX fuel are relevant for safety assessment under the demanding conditions typical of fast reactors [2]. Toxicity, high radiation level, and behaviour at high temperature are all factors that make the measurement of  $\text{PuO}_2$  thermophysical properties complex [3]. Numerical methods such as MD calculations have been proposed to overcome these constraints. This study presents an analysis of  $\text{PuO}_2$  heat capacity by using an interatomic potential based on the model published by Uchida et al. [4]. Lattice parameter, enthalpy and heat capacity have been calculated in the NPT ensemble. The melting temperature of  $\text{PuO}_2$  has been evaluated as well. Calculations have been performed by means of the LAMMPS code (ver. 2019) [5]. Code runs have been carried out on the CRESCO6 cluster [6]. This machine supports the MPI message-passing library giving the opportunity to take advantage of the built-in parallel structure of the code. LAMMPS uses spatial-decomposition techniques to partition the simulation domain into small sub-domains of equal computational cost, one of which is assigned to each processor.

## 2 Interatomic potential and V&V of the model

The Born-Mayer-Huggins pair potential (BMH) is widely used for studying MOX and actinide oxides [3,7]. The BMH potential partially ionic is presented in Eq. 1. The first term accounts for the

---

<sup>1</sup> Corresponding author. E-mail: rolando.calabrese@enea.it.

long-range Coulomb potential. The second and the third term represent the interactions due to the Pauli's repulsion principle and van der Waals forces, respectively. These phenomena are effective at short-range. The Coulomb potential is the dominant term. Effective electronic charges  $z_i$  and  $z_j$  are defined according to the value of ionicity assumed in the model. Term  $r_{ij}$  is the distance between ion  $i$  and ion  $j$ .

$$U_{ij}(r_{ij}) = \frac{z_i z_j e^2}{4\pi\epsilon_0 r_{ij}} + A_{ij} \exp\left(-\frac{r_{ij}}{\rho_{ij}}\right) - \frac{C_{ij}}{r_{ij}^6} \quad (1)$$

A Morse potential accounts for the covalent bond between anion and cation; see Eq. 2. In this equation,  $r_{ij}^*$  stands for the covalent bond length,  $D_{ij}$  and  $\beta_{ij}$  determine the depth and the shape of potential.

$$U_{ij}(r_{ij}) = D_{ij} \left\{ \exp[-2\beta_{ij}(r_{ij} - r_{ij}^*)] - 2\exp[-\beta_{ij}(r_{ij} - r_{ij}^*)] \right\} \quad (2)$$

The coefficients used in calculations are based on the model published by Uchida et al. [4]. Initial values have been tailored to comply with two conditions: a value of lattice constant at room temperature (300 K) of 5.3946 Å; a linear thermal expansion (LTE) in agreement with the curve recommended in [4]. A value of 0.565 was assumed for ionicity with effective charges +2.26 and -1.13 for plutonium and oxygen, respectively. The PuO<sub>2</sub> supercell is composed of 6144 atoms arranged in a *fcc* lattice. The elementary cell is formed by 12 atoms. It has been replicated 8x8x8 along the  $x$ ,  $y$ ,  $z$  orthogonal axes. Following its initialization, the system has been equilibrated for a period of 30 ps. Equilibration of temperature and pressure was attained by means of the Nose-Hoover thermostat/barostat [5]. A period of 30 ps was employed for measurement. The time step used in calculations was 1 fs.

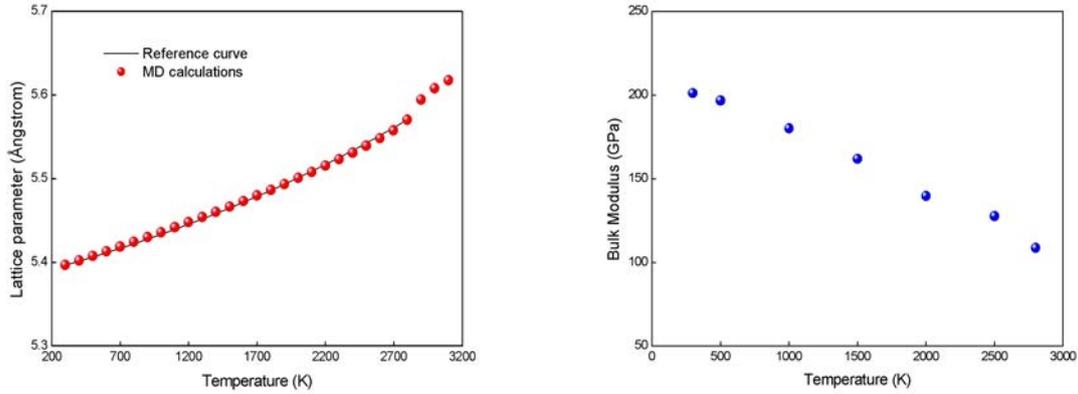
**Table 1:** Coefficients of the interatomic potential.

Pair	$A_{ij}$ (eV)	$\rho_{ij}$ (Å)	$C_{ij}$ (eV·Å <sup>6</sup> )	$D_{ij}$ (eV)	$\beta_{ij}$ (Å <sup>-1</sup> )	$r_{ij}^*$ (Å)
O-O	483029.685	0.178	96.0			
Pu-Pu	0.0	0.200	0.0			
Pu-O	5700.0	0.252	0.0	0.165	2.0	2.37

The model used in MD simulations has been verified and validated using two quantities: lattice constant and bulk modulus at room temperature. X-ray diffraction measurements suggested a lattice parameter at room temperature of 5.3946 Å [8]. This value was assumed as reference. With this assumption a theoretical curve of lattice thermal expansion was constructed. This curve complies with the correlation of LTE recommended in [4]. Comparison of calculations with the curve taken as reference is presented in Fig. 1a. This graph suggests that the dimensional change of the elementary cell with increasing temperature is predicted rather well.

The bulk modulus is defined as  $-1/\beta$  being  $\beta$  the compressibility which is the derivative of molar volume with respect to pressure at constant temperature. For this purpose, 4 values of external pressure ranging from 0.1 MPa to 1.5 GPa have been considered. Calculations have been replicated at 500 K, 1000 K, 1500 K, 2000 K, 2500 K, and 2800 K. Results are presented in Fig. 1b. The calculated bulk modulus at room temperature was 201 GPa. This result slightly underestimates the

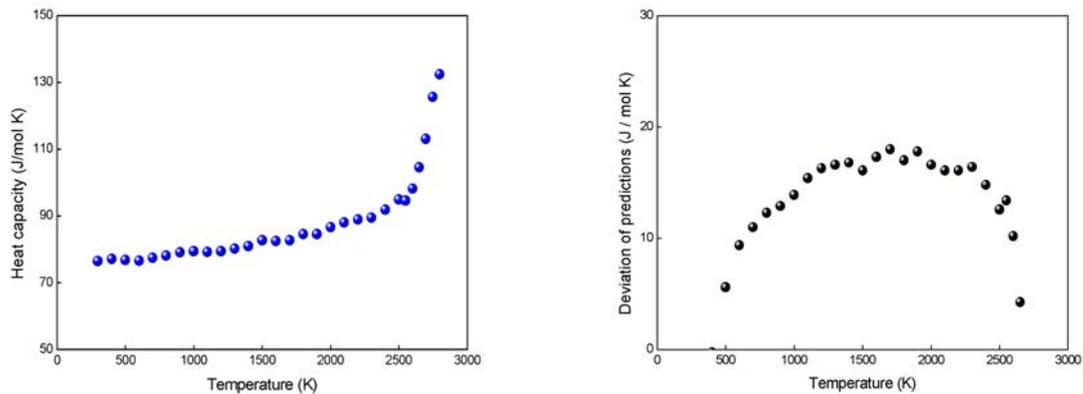
indication reported in [9]. The decrease of the bulk modulus with increasing temperature is consistent with the data published in [10]. This reference reports values that lie in the interval 300-1500 K. A discontinuity of the lattice parameter and internal energy was noted between 2800 K and 2830 K. This estimation of the melting temperature is close to the value measured by Kato et al. (2843 K) [11].



**Fig.1:** Lattice parameter (a) and Bulk Modulus (b).

### 3 Heat capacity results

The enthalpy evaluations obtained in the NPT ensemble have been used to determine the heat capacity as the derivative of enthalpy increment vs. temperature at constant pressure. A qualitative analysis suggested that the increase noted above 2000 K could be represented by a term consistent with the occurrence of electronic disorder. The estimation of activation energy (about 3.9 eV) is close to what reported in [12]. Results of heat capacity are presented in Fig. 2a. They have been compared to the model proposed by Fink that was developed taking into account an extensive experimental dataset [13]. MD calculations underestimate predictions above 500 K. Deviations are rather constant in the interval 1000-2500 K with absolute values that are close to the proposed excitation levels of  $5f$  electrons (Schottky term) [2].



**Fig.2:** Heat capacity (a) and deviations of predictions from the model by Fink [13] (b).

## 4 Conclusions

This study presents MD simulations on the heat capacity of PuO<sub>2</sub>. Lattice constant and bulk modulus have been predicted rather well. However, an underestimation at low temperature and the need for a Schottky term have been confirmed. The phonon and dilation contributions to heat capacity showed some increase beyond 2500 K that is in the temperature domain where a pre-melting transition could occur (85% of the melting temperature). A moderate increase of heat capacity consistent with the onset of electronic disorder has been noted at high temperature. Melting temperature predictions were close to the value measured by Kato et al. [11]. As concluding remark, this study has confirmed the capabilities of MD calculations as a valuable tool to investigate temperature and pressure domains where experimental measurements are scarce or unreliable.

## References

- [1] A.K. Sengupta, K.B. Khan, J. Panakkal, H.S. Kamath, S. Banerjee, Evaluation of high plutonia (44% PuO<sub>2</sub>) MOX as a fuel for fast breeder test reactor, *Journal of Nuclear Materials* **385**, pp. 173-177, (2009).
- [2] M. Kato, T. Matsumoto, Thermal and Mechanical Properties of UO<sub>2</sub> and PuO<sub>2</sub>, NEA/NSC/R(2015)2, pp. 172-177, OECD/Nuclear Energy Agency, Paris, (2015).
- [3] H. Balboa, L. Van Brutzel, A. Chartier, Y. Le Bouar, Assessment of empirical potential for MOX nuclear fuels and thermomechanical properties, *Journal of Nuclear Materials* **495**, pp. 67-77, (2017).
- [4] T. Uchida, T. Sunaoshi, K. Konashi, M. Kato, Thermal expansion of PuO<sub>2</sub>, *Journal of Nuclear Materials* **452**, pp. 281-284, (2014).
- [5] S. Plimpton, Fast Parallel Algorithms for Short-Range Molecular Dynamics, *Journal of Computational Physics* **117**, pp. 1-19, (1995).
- [6] High Performance Computing on CRESCO infrastructure: research activities and results 2018, Italian National Agency for New Technologies Energy and Sustainable Economic Development, Rome, (2019).
- [7] R. Calabrese, Molecular Dynamics Approaches to the Determination of MOX Thermophysical Properties, ADPFISS-LP1-091, ENEA, Rome, (2017).  
(in <https://iris.enea.it/handle/20.500.12079/7993?mode=full.49#.X2TWRz9xezS>)
- [8] O.L. Kruger, H. Savage, Heat Capacity and Thermodynamic Properties of Plutonium Dioxide, *Journal of Chemical Physics* **49**, pp. 4540-4544, (1968).
- [9] S. Hirooka, M. Kato, Sound speeds in and mechanical properties of (U,Pu)O<sub>2-x</sub>, *Journal of Nuclear Science and Technology* **55(3)**, pp. 356-362, (2017).
- [10] S. Minamoto, M. Kato, K. Konashi, Y. Kawazoe, Calculations of thermodynamic properties of PuO<sub>2</sub> by the first-principles and lattice vibration, *Journal of Nuclear Materials* **385**, pp. 18-20, (2009).
- [11] M. Kato, K. Morimoto, H. Sugata, K. Konashi, M. Kashimura, T. Abe, Solidus and liquidus temperatures in the UO<sub>2</sub>-PuO<sub>2</sub> system, *Journal of Nuclear Materials* **373**, pp. 237-245, (2008).
- [12] J.-M. Bonnerot, Propriétés Thermiques des Oxydes Mixtes d'Uranium et de Plutonium, CEA-R-5450, (1988).
- [13] J.K. Fink, Enthalpy and Heat Capacity of the Actinide Oxides, *International Journal of Thermophysics* **3(2)**, pp. 165-200, (1982).

# *HYBRID PARTICLE FIELD MODEL OF LIPID A*

Antonio De Nicola<sup>1,2\*</sup>, Antonio Pizzirusso<sup>2</sup>, Greta Donati<sup>2</sup>, Stefano Caputo<sup>2</sup>

<sup>1</sup>*Department of Organic Material Science, Yamagata University, Yonezawa 992-8510, Japan*

<sup>2</sup>*Dipartimento di Chimica e Biologia "A. Zambelli", Università di Salerno, Fisciano 84084, Italy*

**ABSTRACT.** Lipid A is one of three components of bacterial lipopolysaccharides (constituting the outer membrane of Gram-negative bacteria) and is recognized to have an important biological role in inflammatory response of the immune system. Its biological activity is modulated by the number of acyl-chains and from the electrostatic interactions with the different counter-ions. In this paper we report a coarse-grained model of poly-acyl Lipid A based on the hybrid particle field molecular dynamics approach (hPF-MD).

## **1 Introduction**

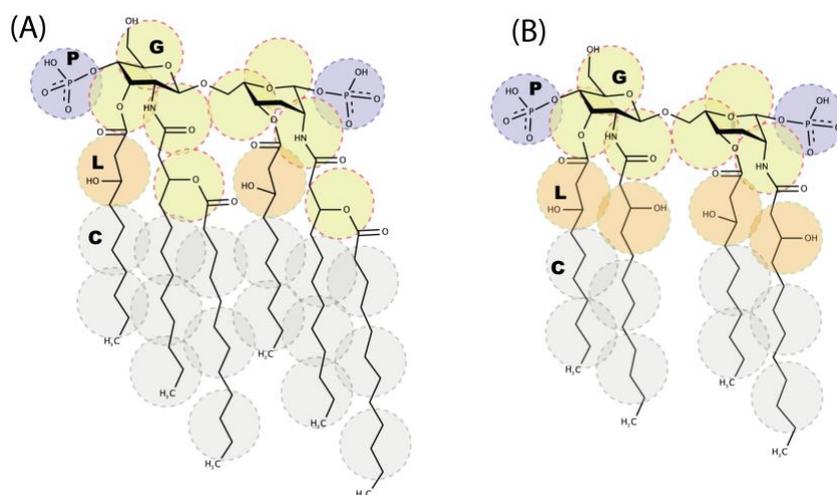
Lipopolysaccharides (LPS) are the major lipid constituents of the external leaflet of the asymmetric outer membrane (OM) in Gram-negative bacteria.[1,2] The OM acts as a structural and functional barrier against the penetration of xenobiotic agents like hydrophobic antibiotics, digestive enzymes, heavy metals, detergents, or bile salts.[3] The LPS molecule can contain three different regions: (i) the polymeric O-antigen (repetitive monosaccharide subunits) responsible for the bacterium immunospecificity,[4] (ii) a branched oligosaccharide brush (composed of 8-12 monosaccharide units) and (iii) Lipid A unit (composed by typically 4-6 acyl chains attached to a phosphorylated glucosamine disaccharide) responsible for the anchoring of the LPS leaflet onto the phospholipid leaflet of the OM.[5,6] Lipid A has a fundamental biological role as a stimulator of the inflammatory response by the immune system.[7-9] The immune response to the presence of Lipid A is modulated by variations of its chemical structure such as the number, length and position of acyl chains as well as by the overall molecular charge.[9-11]

In the past decades, several MD studies have been reported on both LPS and Lipid A. In particular, all-atom[7,13-19] and coarse-grained (CG) models[5,20-23] have been developed and are continuously improved.[14,15,18] However, due to the time and length scale limits, simulations of large systems and/or phenomena occurring on time scales larger than the  $\mu$ s are still challenging even for CG models. Overcoming such bottlenecks may be achieved by the development of models for lipopolysaccharides within a density-functional approach,[24] and in particular by the hybrid particle-field/molecular dynamics (hPF-MD) model.[25-27] The main advantage of hPF-MD is that the non-bonded pair interactions between particles are replaced by the evaluation of an external potential based on the local particle density. Thanks to an efficient parallelization scheme,[28] it is possible to overcome the current time and scale limits of standard MD simulations while retaining a sufficiently high molecular resolution in the model to retain chemically resolved moieties. Past studies demonstrated the ability of this

approach to investigate diverse soft matter systems, including lamellar and non-lamellar phospholipid moieties,[29–31] biocompatible block-copolymers,[32,33] non-charged surfactants,[34,35] and peptides.[36] In this report, we show the development of specific hybrid CG model for the biologically relevant Lipid A (hexa- and tetra-acyl chains).

## 2 Lipid A Model

In this study, the single-molecule Hamiltonian  $H_0(\{r\}_m)$  is based on a CG representation of both Lipid A, the ions, and the solvent (Figure 1). All intramolecular interactions have the same functional form as in standard force fields, [25] using the parameters developed in our previous studies,[29,37] apart from the length  $l_B$  between two L bead types. Specifically, we employ lengths  $l_B = 0.47$  nm for hexa- and  $l_B = 0.34$  nm for the tetra-acylated Lipid A in order to improve the agreement with the all atom structures. The complete list of Lipid A bonded parameters is reported in reference[38].



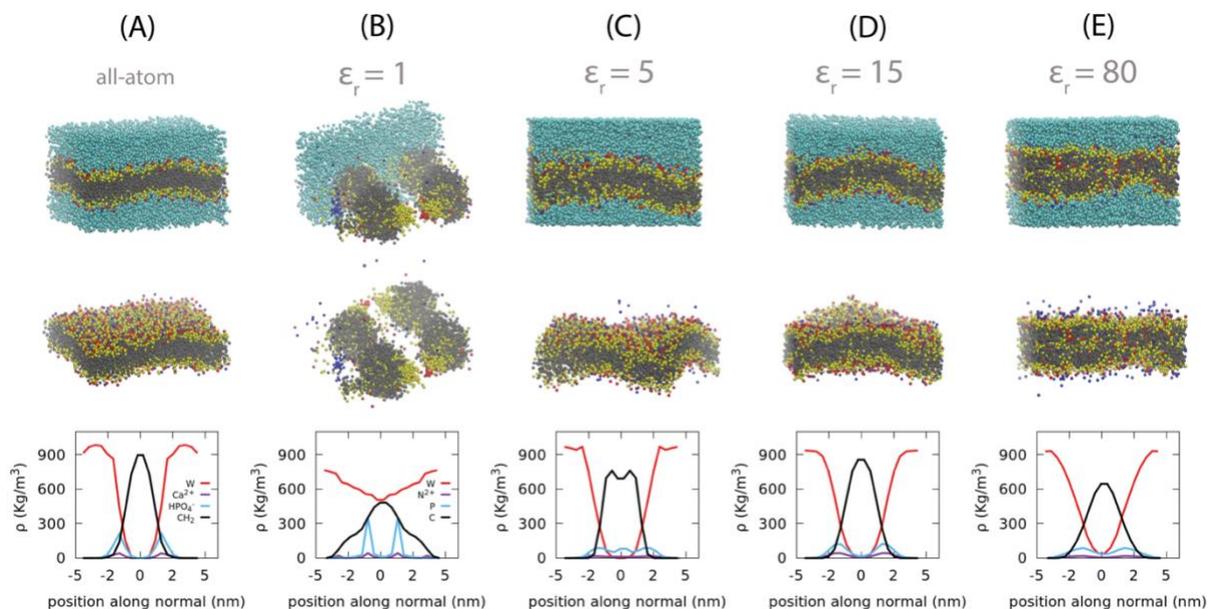
**Figure 1.** Chemical structures of (A) hexa- and (B) tetra-acylated forms of Lipid A. The drawings compare the atomic molecular structures and the corresponding coarse-grained mapping used in this work. The different bead types are indicated by the P, G, L, C letters, the same color in the circles indicates beads of the same type. The counter-ion bead (N), and the water CG bead (W), mapping four H<sub>2</sub>O molecules, are not shown in the figure.

## 3 Results

*Effect of the dielectric constant  $\epsilon_r$  on the stability of Lipid A bilayer.* Appropriate calibration of the effective dielectric constant  $\epsilon_r$  is fundamental for the correct description of charged lipid bilayers, as demonstrated by past work on a POPG membrane. For this reason, we first investigated the dependency of the stability of Lipid A bilayers as a function of different values of the relative dielectric constant.

Starting from an initial CG configuration of hexa-acylated Lipid A mapped on the reference atomistic configuration (HPF1, Table 2), we ran four hPF-MD simulations, each 300 ns long,

at  $\epsilon_r=1,5,15,80$  values, reporting the corresponding average structures and density profiles for the different chemical moieties along the direction normal to the plane of the bilayer in Figure 2.



**Figure 2.** Equilibrated structures of Lipid A bilayers. Top panels: Representative snapshots of the equilibrium configurations for (A) all-atom and (B-E) hPF-MD simulations of hexa-acylated Lipid A using different relative dielectric constant ( $\epsilon_r$ ) values. The CG bead types are colored as: P (red), G and L (yellow), C (gray), N2+ (blue), and W (cyan). The same configuration is also represented without the water beads for the sake of clarity. Bottom panels: Average mass density profiles for the corresponding systems shown above.

As expected, the organization of the Lipid A layers depends on the value of the dielectric constant (Figure 2). A very low value  $\epsilon_r=1$  produces an unstable bilayer with complete disruption of the bilayer structure within the first 30 ns of simulation, while larger values of  $\epsilon_r$  maintain the global lamellar structure of the membrane. hPF-MD simulations using a value of  $\epsilon_r=15$  (Figure 2D) yield a stable bilayer structure within the simulation time scale and reproduces the most accurately the density profiles of the atomistic reference. Conversely, the density profiles indicate that the organization of Lipid A inside the membrane is not correct for  $\epsilon_r=5$  and 80. In both cases we observe the appearance of strong surface undulation and a migration of the hydrophilic sites to the middle plane of the bilayer (see snapshots with  $\epsilon_r=80$ ), with the negatively charged P beads being distributed along the whole cross-section of the membrane (Figures 2C,E), indicating that the lipids are tilted by an angle greater than 45 degrees (Figure 2E).

## References

- [1] H. Nikaido, Molecular Basis of Bacterial Outer Membrane Permeability Revisited, *Microbiol. Mol. Biol. Rev.* 67 (2003) 593–656. doi:10.1128/MMBR.67.4.593–656.2003.
- [2] S.G. Wilkinson, Bacterial lipopolysaccharides—Themes and variations, *Prog. Lipid Res.* 35 (1996) 283–343. doi:https://doi.org/10.1016/S0163-7827(96)00004-5.
- [3] N. Kučerka, E. Papp-Szabo, M.P. Nieh, T.A. Harroun, S.R. Schooling, J. Pencer, E.A. Nicholson, T.J. Beveridge, J. Katsaras, Effect of cations on the structure of bilayers formed by lipopolysaccharides isolated from *Pseudomonas aeruginosa* PAO1, *J. Phys. Chem. B.* 112 (2008) 8057–8062. doi:10.1021/jp8027963.
- [4] M. Caroff, D. Karibian, Structure of bacterial lipopolysaccharides, *Carbohydr. Res.* 338 (2003) 2431–2447. doi:https://doi.org/10.1016/j.carres.2003.07.010.
- [5] H. Ma, F.J. Irudayanathan, W. Jiang, S. Nangia, Simulating Gram-Negative Bacterial Outer Membrane: A Coarse Grain Model, *J. Phys. Chem. B.* 119 (2015) 14668–14682. doi:10.1021/acs.jpcc.5b07122.
- [6] A. Molinaro, O. Holst, F. Di Lorenzo, M. Callaghan, A. Nurisso, G. D’Errico, A. Zamyatina, F. Peri, R. Berisio, R. Jerala, J. Jiménez-Barbero, A. Silipo, S. Martín-Santamarí, Chemistry of lipid a: At the heart of innate immunity, *Chem. - A Eur. J.* 21 (2015) 500–519. doi:10.1002/chem.201403923.
- [7] F.J.S. Pontes, V.H. Rusu, T.A. Soares, R.D. Lins, The effect of temperature, cations, and number of Acyl chains on the lamellar to non-lamellar transition in Lipid-A membranes: A microscopic view, *J. Chem. Theory Comput.* 8 (2012) 3830–3838. doi:10.1021/ct300084v.
- [8] A.B. Schromm, K. Brandenburg, H. Loppnow, A.P. Moran, M.H.J. Koch, E.T.H. Rietschel, U. Seydel, Biological activities of lipopolysaccharides are determined by the shape of their lipid A portion, *Eur. J. Biochem.* 267 (2000) 2008–2013. doi:10.1046/j.1432-1327.2000.01204.x.
- [9] K. Brandenburg, U. Seydel, A.B. Schromm, H. Loppnow, M.H.J. Koch, E.T. Rietschel, Conformation of lipid A, the endotoxic center of bacterial lipopolysaccharide, *Innate Immun.* 3 (1996) 173–178. doi:10.1177/096805199600300302.
- [10] S.I. Miller, R.K. Ernst, M.W. Bader, LPS, TLR4 and infectious disease diversity, *Nat. Rev. Microbiol.* 3 (2005) 36–46. doi:10.1038/nrmicro1068.
- [11] K. Brandenburg, M.H.J.H.J. Koch, U. Seydel, Phase diagram of deep rough mutant lipopolysaccharide from *Salmonella minnesota* R595, *J. Struct. Biol.* 108 (1992) 93–106. doi:10.1016/1047-8477(92)90010-8.
- [12] H. Reichelt, C.A. Faunce, H.H. Paradies, The phase diagram of charged colloidal lipid A-diphosphate dispersions, *J. Phys. Chem. B.* 112 (2008) 3290–3293. doi:10.1021/jp711720j.
- [13] S. Khalid, T.J. Piggot, F. Samsudin, Atomistic and Coarse Grain Simulations of the Cell Envelope of Gram-Negative Bacteria: What Have We Learned?, *Acc. Chem. Res.* 52 (2019) 180–188. doi:10.1021/acs.accounts.8b00377.
- [14] R.D. Lins, T.P. Straatsma, Computer simulation of the rough lipopolysaccharide membrane of *Pseudomonas aeruginosa*, *Biophys. J.* 81 (2001) 1037–1046. doi:10.1016/S0006-3495(01)75761-X.
- [15] K.N. Kirschner, R.D. Lins, A. Maass, T.A. Soares, A glycam-based force field for simulations of lipopolysaccharide membranes: Parametrization and validation, *J. Chem. Theory Comput.* 8 (2012) 4719–4731. doi:10.1021/ct300534j.
- [16] T.A. Soares, T.P. Straatsma, Assessment of the convergence of molecular dynamics simulations of lipopolysaccharide membranes, *Mol. Simul.* 34 (2008) 295–307. doi:10.1080/08927020701829880.
- [17] A. Li, J.W. Schertzer, X. Yong, Molecular dynamics modeling of *Pseudomonas aeruginosa* outer membranes, *Phys. Chem. Chem. Phys.* 20 (2018) 23635–23648. doi:10.1039/c8cp04278k.
- [18] E.L. Wu, O. Engström, S. Jo, D. Stuhlsatz, M.S. Yeom, J.B. Klauda, G. Widmalm, W. Im, Molecular dynamics and NMR spectroscopy studies of *E. coli* lipopolysaccharide structure and dynamics, *Biophys. J.* 105 (2013) 1444–1455. doi:10.1016/j.bpj.2013.08.002.
- [19] D.E.S.S. Santos, L. Pol-Fachin, R.D. Lins, T.A. Soares, Polymyxin Binding to the Bacterial Outer Membrane Reveals Cation Displacement and Increasing Membrane Curvature in Susceptible but Not in Resistant Lipopolysaccharide Chemotypes, *J. Chem. Inf. Model.* 57 (2017) 2181–2193. doi:10.1021/acs.jcim.7b00271.
- [20] V.H. Rusu, R. Baron, R.D. Lins, PITOMBA: Parameter interface for oligosaccharide molecules based on atoms, *J. Chem. Theory Comput.* 10 (2014) 5068–5080. doi:10.1021/ct500455u.
- [21] B. Van Oosten, T.A. Harroun, A MARTINI extension for *Pseudomonas aeruginosa* PAO1 lipopolysaccharide, *J. Mol. Graph. Model.* 63 (2016) 125–133. doi:10.1016/j.jmgm.2015.12.002.
- [22] P.C. Hsu, B.M.H. Bruininks, D. Jefferies, P. Cesar Telles de Souza, J. Lee, D.S. Patel, S.J. Marrink, Y. Qi, S. Khalid, W. Im, CHARMM-GUI Martini Maker for modeling and simulation of complex bacterial membranes with lipopolysaccharides, *J. Comput. Chem.* 38 (2017). doi:10.1002/jcc.24895.
- [23] F.A. Baltoumas, S.J. Hamodrakas, V.A. Iconomidou, The gram-negative outer membrane modeler: Automated building of lipopolysaccharide-rich bacterial outer membranes in four force fields, *J. Comput. Chem.* 40 (2019) 1727–1734. doi:10.1002/jcc.25823.
- [24] T Kawakatsu., *Statistical physics of polymers: an introduction.* Springer-Verlag, Berlin, Heidelberg, 2004. ISBN 3-540-43440-2. pp 216, *Polym. Int.* (2005). doi:10.1002/pi.1772.
- [25] G. Milano, T. Kawakatsu, Hybrid particle-field molecular dynamics simulations for dense polymer systems., *J. Chem. Phys.* 130 (2009) 214106. doi:10.1063/1.3142103.
- [26] T.A. Soares, S. Vanni, G. Milano, M. Cascella, Toward Chemically Resolved Computer Simulations of Dynamics and Remodeling of Biological Membranes, *J. Phys. Chem. Lett.* 8 (2017) 3586–3594. doi:10.1021/acs.jpcclett.7b00493.

- [27] S. J. Marrink, V. Corradi, P. C.T. Souza, H. I. Ingólfsson, D. Peter Tieleman, M. S.P. Sansom, S.J. Marrink, P.C.T. Souza, V. Corradi, D.P. Tieleman, H.I. Ingólfsson, M.S.P. Sansom, Computational Modeling of Realistic Cell Membranes, *Chem. Rev.* 119 (2019) 6184–6226. doi:10.1021/acs.chemrev.8b00460.
- [28] Y. Zhao, A. De Nicola, T. Kawakatsu, G. Milano, Hybrid particle-field molecular dynamics simulations: Parallelization and benchmarks, *J. Comput. Chem.* 33 (2012) 868–880. doi:10.1002/jcc.22883.
- [29] A. De Nicola, Y. Zhao, T. Kawakatsu, D. Roccatano, G. Milano, Hybrid Particle-Field Coarse-Grained Models for Biological Phospholipids, *J. Chem. Theory Comput.* 7 (2011) 2947–2962. doi:10.1021/ct200132n.
- [30] A. De Nicola, Y. Zhao, T. Kawakatsu, D. Roccatano, G. Milano, Validation of a hybrid MD-SCF coarse-grained model for DPPC in non-lamellar phases, *Theor. Chem. Acc.* 131 (2012) 1167. doi:10.1007/s00214-012-1167-1.
- [31] E. Sarukhanyan, A. De Nicola, D. Roccatano, T. Kawakatsu, G. Milano, A. De Nicola, D. Roccatano, T. Kawakatsu, G. Milano, Spontaneous insertion of carbon nanotube bundles inside biomembranes: A hybrid particle-field coarse-grained molecular dynamics study, *Chem. Phys. Lett.* 595–596 (2014) 156–166. doi:10.1016/j.cplett.2014.01.057.
- [32] A. De Nicola, T. Kawakatsu, G. Milano, A Hybrid Particle-Field Coarse-Grained Molecular Model for Pluronic Water Mixtures, *Macromol. Chem. Phys.* 214 (2013) 1940–1950. doi:10.1002/macp.201300214.
- [33] A. De Nicola, S. Hezaveh, Y. Zhao, T. Kawakatsu, D. Roccatano, G. Milano, Micellar drug nanocarriers and biomembranes: how do they interact?, *Phys. Chem. Chem. Phys.* 16 (2014) 5093. doi:10.1039/c3cp54242d.
- [34] A. De Nicola, T. Kawakatsu, C. Rosano, M. Celino, M. Rocco, G. Milano, Self-Assembly of Triton X-100 in Water Solutions: A Multiscale Simulation Study Linking Mesoscale to Atomistic Models, *J. Chem. Theory Comput.* 11 (2015) 4959–4971. doi:10.1021/acs.jctc.5b00485.
- [35] A. Pizzirusso, A. De Nicola, G.J.A. Sevink, A. Correa, M. Cascella, T. Kawakatsu, M. Rocco, Y. Zhao, M. Celino, G. Milano, Biomembrane solubilization mechanism by Triton X-100: a computational study of the three stage model, *Phys. Chem. Chem. Phys.* 19 (2017) 29780–29794. doi:10.1039/C7CP03871B.
- [36] S.L. Bore, G. Milano, M. Cascella, Hybrid Particle-Field Model for Conformational Dynamics of Peptide Chains, *J. Chem. Theory Comput.* 14 (2018) 1120–1130. doi:10.1021/acs.jctc.7b01160.
- [37] Y.-L. Zhu, Z.-Y. Lu, G. Milano, A.-C. Shi, Z.-Y. Sun, Hybrid particle–field molecular dynamics simulation for polyelectrolyte systems, *Phys. Chem. Chem. Phys.* 18 (2016) 9799–9808. doi:10.1039/C5CP06856H.
- [38] A. De Nicola, T.A. Soares, D.E.S. Santos, S.L. Bore, G.J.A. Sevink, M. Cascella, G. Milano, Aggregation of Lipid A variants: A hybrid particle-field model, *Biochim. Biophys. Acta - Gen. Subj.* (2020) 129570. doi:10.1016/j.bbagen.2020.129570.

# DOSE CALCULATION UN DEMO PHTS

Chiovaro P.<sup>\*1</sup>, Del Nevo A.<sup>2</sup>

<sup>1</sup>*University of Palermo, Italy*

<sup>2</sup>*ENEA C.R. Brasimone, ENEA, Camugnano (BO), Italy*

**ABSTRACT.** Within the framework of the activities foreseen by EUROfusion action on the “Optimization of WCLL PHTSs for N16 and N17 dose rate minimization”, the University of Palermo is involved in the assessment of the absorbed dose in the isolation valves envisaged for cooling circuits of the DEMO reactor equipped with a Water Cooled Lithium Lead Breeding Blanket (WCLL BB). In particular, the aim of this research activity is to assess the spatial distribution of the absorbed dose, due to the decay of nitrogen isotopes produced by coolant activation, around some key components of WCLL BB cooling circuit, focussing the attention on the Primary Heat Transfer System (PHTS), in the Upper Pipe Chase (UPC) of the reactor. A 3-D map of the absorbed dose in the volume under investigation is of considerable interest in order to be able to evaluate the conditions influencing devices whose availability is undermined by intense ionizing radiation fields. To this regard the assessment of the dose absorbed in the aforementioned valve is of uttermost importance. Photonic and neutronic calculations, to evaluate the spatial distribution of the absorbed dose, have been performed following a computational approach based on the Monte Carlo method and adopting the Monte Carlo N-Particle (MCNP5-1.60) code along with the JEFF-3.2 transport cross section libraries.

## 1 Introduction

Within the framework of the activities foreseen by EUROfusion action on the “Optimization of WCLL PHTSs for N16 and N17 dose rate minimization”, the University of Palermo is involved in the assessment of the absorbed dose in the isolation valves envisaged for cooling circuits of the DEMO reactor equipped with a Water Cooled Lithium Lead Breeding Blanket (WCLL BB) [1]. In such a nuclear system the endothermic charged-particle reactions  $^{16}\text{O}(n,p)^{16}\text{N}$  (with a threshold energy of  $\sim 10.2$  MeV) and  $^{17}\text{O}(n,p)^{17}\text{N}$  (with a threshold energy of  $\sim 8.4$  MeV) are the principal sources of water radioactivity during operation [2,3].  $^{16}\text{N}$  decays by emission of  $\beta$  particle and emits  $\gamma$  rays with an half-life of 7.13 s while  $^{17}\text{N}$  decays by  $\beta$  particles and emits neutrons with an half-life of 4.173 s. The aim of this research activity is to assess the spatial distribution of the absorbed dose, due to the decay of nitrogen isotopes produced by coolant activation, around some key components of WCLL BB cooling circuit, focussing the attention on the Primary Heat Transfer System (PHTS), in the Upper Pipe Chase (UPC) of the reactor. The neutron and photon dose maps have been evaluated starting from the spatial distribution of the activity volumetric densities of nitrogen isotopes, obtained in a previous work, as they allow the set-up of photon ( $^{16}\text{N}$ ) and neutron ( $^{17}\text{N}$ ) sources for the related nuclear analyses. A 3-D map of the absorbed dose in the volume under investigation is of considerable interest in order to be able to evaluate the conditions influencing devices whose availability is undermined by intense ionizing radiation fields. To this regard the assessment of the dose absorbed in the aforementioned valve electronic actuators and hydraulic seals, if made of organic materials, is of uttermost importance. As far as the PHTS and the UPC are concerned, attention has been focused on 2018 design [3-4]. Photonic and neutronic calculations, to evaluate the spatial distribution of the absorbed dose, have been performed following a computational approach based on the Monte Carlo method and adopting the Monte Carlo N-Particle (MCNP5-1.60) code [5]

along with the JEFF-3.2 transport cross section libraries [6].

## 2 Primary Heat Transfer Systems

DEMO presents four independent PHTSs. The largest PHTS is devoted to remove the thermal power from the BB, two PHTSs are necessary to extract heat from the Divertor while the last one is intended to cool the Vacuum Vessel (VV). The main function of the PHTS is to provide cooling water to the first wall and blanket systems and to transfer the thermal power for its conversion into electricity. The WCLL PHTS is constituted by two independent PHTSs: the Breeder Zone (BZ) PHTS and the First Wall (FW) PHTS. The cooling water is distributed and collected by means of four ring headers, hosted in a proper that runs all around the tokamak [7]. So, its PHTS foresees four heat exchangers, two for the Outboard Blanket (OB) and two for the Inboard Blanket (IB) and it is located in a reactor building with a peculiar lay-out (Fig. 1).

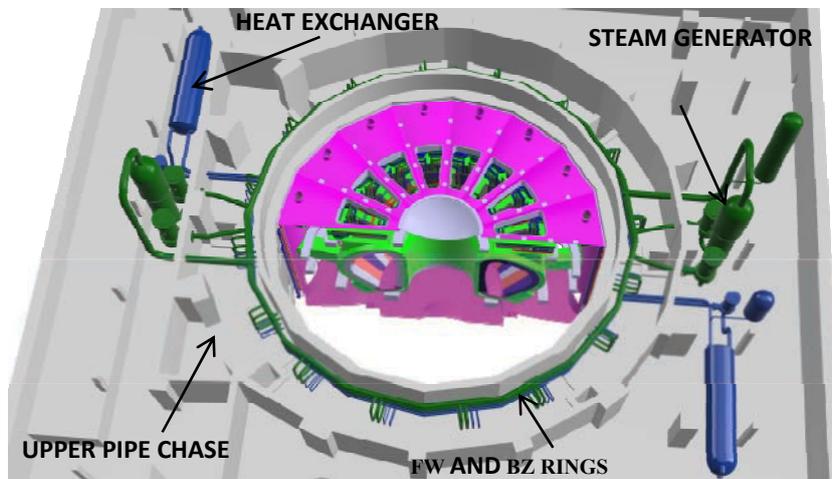


Figure 1. View of DEMO PHTS in the UPC.

On the basis of the ITER experience, the DEMO BB PHTS could foresee a number of valves greater than 128 which have to work in a harsh environment characterized by high gamma and neutron fluences [9].

## 3 PHTS isolation valves

In the framework of the R&D activities for the DEMO Balance of Plant, one of the most interesting solution for the mitigation of the risk of an In-Vessel LOCA envisages the use of gate valves for on/off operation. The proposed solution involves the use of a gate-type valve operated by a pneumatic linear actuator equipped with solenoid valves for on/off operation. The valve body chosen for the isolation valve corresponds to a flexible wedge gate valve from the VELAN catalog [8]. The component is mainly made of stainless steel and does not have materials that are particularly susceptible to gamma radiation. The seals are and can be arranged in series with a lip weld to reduce leakage probability. Therefore, attention has been focused on the PHTS piping arranged in the UPC where isolation valves could be located. In particular, a slice of 22.5° (corresponding to one blanket sector) has been considered as it is shown in figure 3 where it is possible to notice the borated heavy concrete walls of the UPC, the pipes (AISI 316 LN) in which the radioactive water flows and the considered model of valve, that is a simplified version of a gate valve (DN 150, rating 1500) [9].

## 4 MCNP model

A fully heterogeneous 3-D MCNP model of such a slice has been set up for converting the CAD geometry representation into the MCNP constructive solid geometry representation. Reflecting surfaces have been used in the toroidal direction to take into account the geometric continuity of the calculation domain in that direction (Fig. 2) [9].

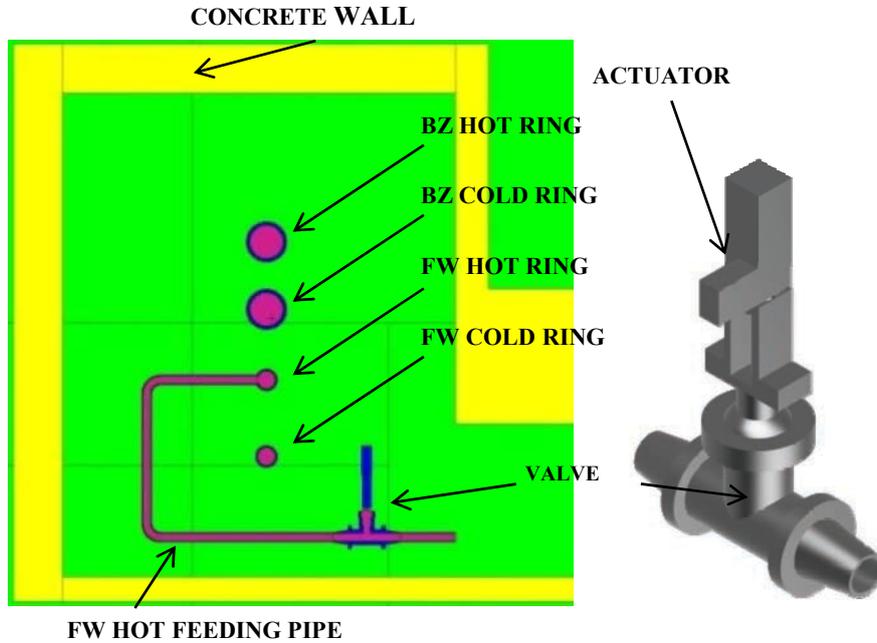


Figure 2. Poloidal-radial section of the MCNP model of the UPC.

It can be observed (Fig. 2) that the neutron flux profiles have the same trend in both the blankets and how the flux in the WCLL is always higher than in the WCCB. With regard to the volumetric density of the nuclear power radial profiles (Fig. 3), they are slightly different and the WCCB curve is above the WCLL one. As far as the photon and neutron sources are concerned, they were modelled as diffuse sources, since water of each tube, in the considered portion of UPC, was conceived as a uniform volume source. Therefore, the emission probability of each source has been calculated from its relative activity (Tabs. 1 and 2) with reference to the total activity of the water present in the domain.

Table 1. Nitrogen activity in the FW circuit [GBq].

COMPONENT	$^{16}\text{N}$	$^{17}\text{N}$
Hot Feeding Pipe	4.893E+02	5.132E-02
Hot Ring	2.089E+03	1.903E-01
Hot Leg	9.368E+02	8.316E-02
Cold Leg	1.683E+02	4.425E-03
Cold Ring	3.488E+02	8.937E-03
Cold Feeding Pipe	5.826E+01	1.466E-03
Valve	6.973E+01	7.313E-03

Table 2. Nitrogen activity in the BZ circuit [GBq].

COMPONENT	$^{16}\text{N}$	$^{17}\text{N}$
Hot Feeding Pipe	2.718E+02	3.093E-02
Hot Ring	6.941E+02	6.781E-02
Hot Leg	1.484E+02	1.163E-02
Cold Leg	9.741E+01	5.664E-03
Cold Ring	3.261E+02	1.845E-02
Cold Feeding Pipe	8.526E+01	4.717E-03

Regarding the energy distributions of both  $^{16}\text{N}$  photons and  $^{17}\text{N}$  neutrons they have been biased taking into account their discrete spectra with the related branching from each energy level [9]. Furthermore, a compromise has been reached between the need to obtain a detailed spatial distribution of the dose and the extent of the computational time by implementing a super-imposed mesh ( $500 \times 500 \times 500$ ) with a uniform voxel size of  $2.52 \times 5.56 \times 1.86 \text{ cm}^3$  [9].

## 5 Dose Calculation results

Detailed 3D photonic and neutronic analyses have been carried out for the assessment of the dose nearby FW and BZ cooling circuit key-points, due to  $\gamma$  radiation from  $^{16}\text{N}$  and neutron emission from  $^{17}\text{N}$ . A period of 7 FPY has been considered as it is assumed to be the expected life of the breeding blanket of DEMO taking into account an availability factor of 33% [10].

A steady state scenario has been taken into account considering the plasma flat-top phase of the reactor. Analyses have been carried out by simulating  $10^{11}$  histories for photon transport and  $6 \cdot 10^{10}$  for neutron transport, so that the results obtained are affected by relative errors lower than 1% in the most of the calculation domain.

Figure 5 and 6 show the spatial distribution of the  $\gamma$  absorbed dose in the slice of the UPC taken into account. The full-scale of 2 MGy has been selected, in the most of the images, as it represents the reference figure for the qualified valves in NPP, in order to verify their suitability for DEMO. It can be observed that a large part of the volume is characterized by a dose above the threshold value identified while, clearly, maximum values are localized in the water, that is, the source volumes. It is also possible to notice shaded areas to the radiation due to the mutual shielding effect of the pipes.

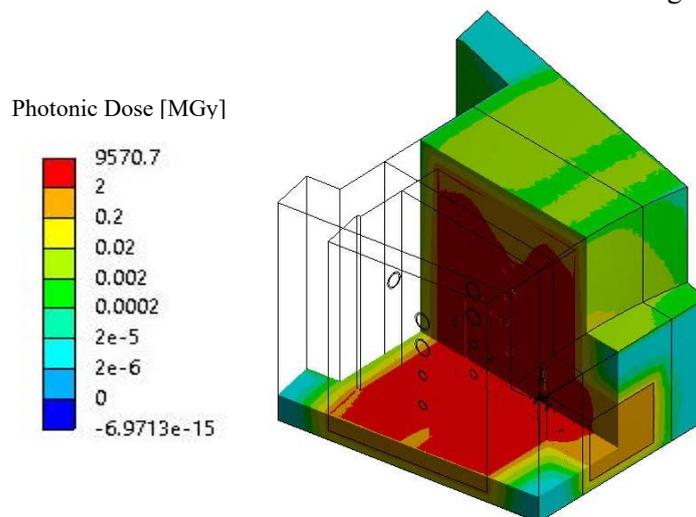


Figure 3. Absorbed  $\gamma$  dose spatial distribution in the UPC slice.

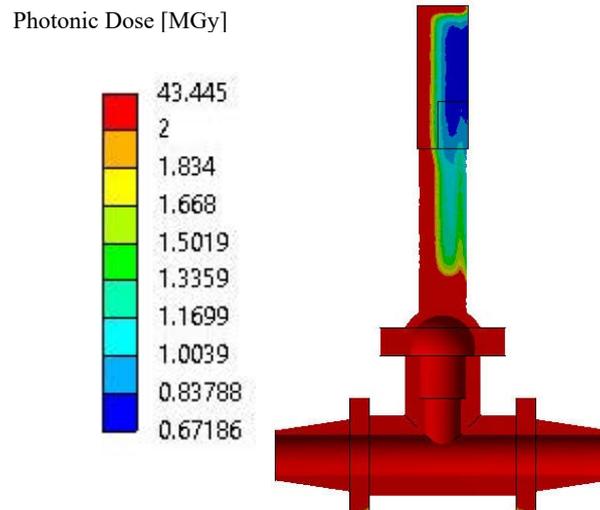


Figure 4. . Absorbed  $\gamma$  dose spatial distribution in the valve.

The outcomes related to the valve, of course, imply that some design modifications of the BB and/or the PHTS must be foreseen. Figure 5 suggests some solutions as the location of the most rad-sensitive equipment away from the highest dose (e.g. the valve actuators), as well some R&D to replace such components or materials with other more rad-resistant capability in order to have less constraint for the location of valves in such complex and crowded layout.

The spatial distribution of neutron absorbed dose in the UPC slice is shown in figures 5 and 6. The dose due to neutron radiation is several order of magnitude lower than photon dose, nevertheless,  $^{17}\text{N}$  activity during the DEMO pulse (tabs. 2 and 3) deserves, for completion, further neutron transport studies to evaluate the inventory of activation products and to evaluate their effects on the radiation field in the UPC.

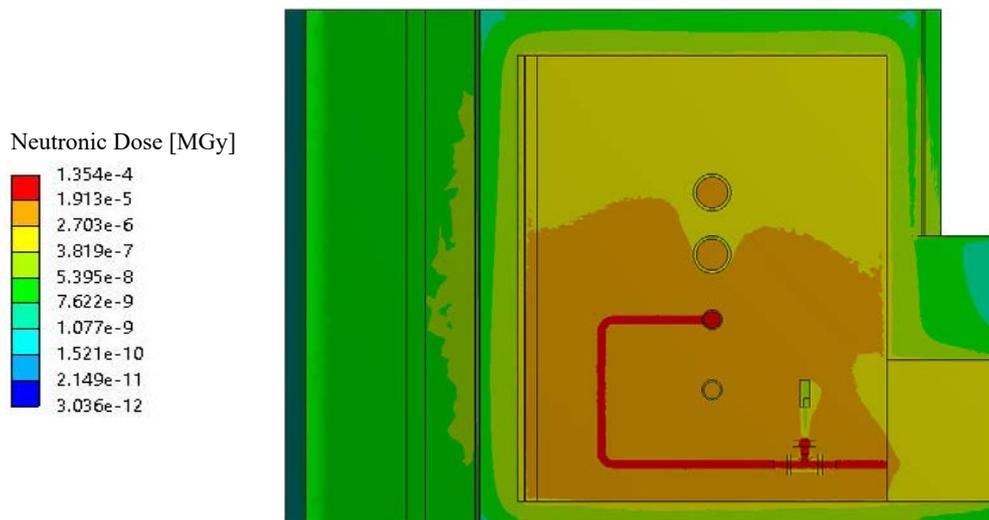


Figure 6. Absorbed neutron dose spatial distribution in the UPC slice.

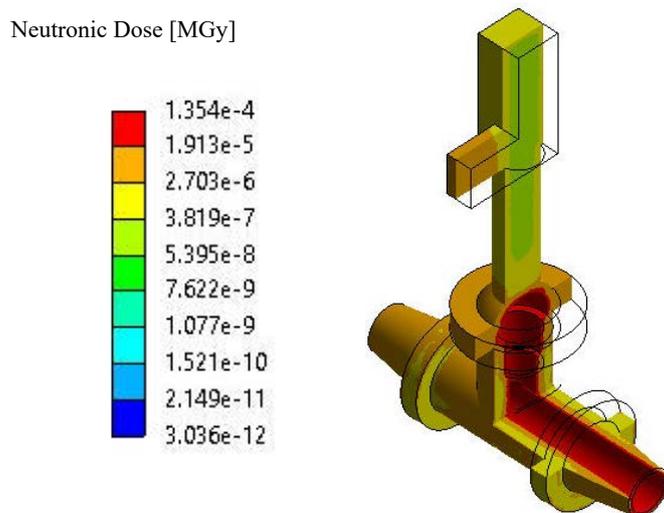


Figure 7. Absorbed neutron dose spatial distribution in the valve.

The results obtained show, as expected, that the main contribution to the absorbed dose by matter in the UPC comes from the photons emitted by the  $^{16}\text{N}$  decay and that, on the other hand, the estimated  $^{17}\text{N}$  activity leads to consider specific analysis for the evaluation the extent of neutron activation in the pipes. Furthermore, these outcome lead to a further development of the present research activity aimed at developing PHTS design changes in order to lower the dose absorbed by the valve. In particular, the peculiarities of the spatial distribution of dose found suggest some simple modifications such as the use of bulkheads to shield the valve and/or change the lay-out of the pipes and/or develop more rad-resistant material for specific cases.

## Conclusion

Within the framework of EUROfusion action, at the University of Palermo a research campaign has been performed in order to assess the response of an hypothetical isolation valve of the DEMO BB PHTS to the photon and neutron irradiation, due to water activation, in terms of absorbed dose. In particular, the work has aimed at investigating some technical strategies to reduce valve aging due to the aforementioned irradiation conditions.

Neutron and photon transport analyses have been performed in a proper model of the UPC in order to assess the a spatial distribution of the absorbed dose in the target locations. Results obtained show, as expected, that the main contribution to the absorbed dose by matter in the UPC comes from the photons emitted by the  $^{16}\text{N}$  decay and that, on the other hand, the estimated  $^{17}\text{N}$  activity leads to consider specific analysis for the evaluation the extent of neutron activation in the pipes.

## References

- [1] F. Cismondi et al., Progress of the conceptual design of the European DEMO Breeding Blanket, Tritium Extraction and Coolant Purification Systems. Presented at the 14th International Symposium On Fusion Nuclear Technology, Budapest, Hungary, 2019.
- [2] Pierluigi Chiovaro, I. Moscato, et al., Investigation of the DEMO WCLL Breeding Blanket Cooling Water Activation, Presented at the 14<sup>th</sup> International Symposium on Fusion Nuclear Technology (ISFNT), 22 – 27 Sept 2019, Budapest (Hungary). (Pinboard ID 22297)
- [3] G. Gliss, I. Moscato, et al., Integration of DEMO hazard piping into the tokamak building, Presented at the 14th International Symposium on Fusion Nuclear Technology (ISFNT), 22 – 27

- Sept 2019, Budapest (Hungary). (ID 22319)
- [4] L. Barucca, et al., Status of EU DEMO Heat Transport and Power Conversion Systems, Fusion Eng. Des., 136 (2018) 1557-1566.
  - [5] X-5 Monte Carlo Team: MCNP – A General Monte Carlo N-Particle Transport Code, Version 5, LANL, Los Alamos, New Mexico, USA, 2003 (April 2003).
  - [6] JEFF3.2 Nuclear Data Library, (2018) [http://www.oecd-neo.org/dbforms/data/eva/evatapes/jeff\\_32/](http://www.oecd-neo.org/dbforms/data/eva/evatapes/jeff_32/).
  - [7] Martelli E., et Ali, Study of EU DEMO WCLL breeding blanket and primary heat transfer system integration, Fusion Engineering and Design, Volume 136, Part B, November 2018, Pages 828-833.
  - [8] CAT-PS-10-11, VELAN Pressure Seal & Bolted Bonnet Gate, Globe and Check Valves.
  - [9] Pierluigi Chiovaro, I. Moscato, et al., Assessment of the Dose Rates due to Water Activation on an Isolation Valve of the DEMO WCLL Breeding Blanket Primary Heat Transfer System, Presented at the 14th International Symposium on Fusion Nuclear Technology (ISFNT), 22 – 27 Sept 2019, Budapest (Hungary). (ID 22303)
  - [10] G. Federici et al., Overview of the DEMO staged design approach in Europe, Nuclear Fusion, vol. 59, no. 6, p. 066013, 2019.

# PRELIMINARY COMPUTATIONAL TESTS FOR EXTRACTING ATMOSPHERIC TRANSFER MATRICES IN GAINS-ITALY

Gino Briganti<sup>1</sup>, Andrea Cappelletti<sup>1</sup>, Luisella Ciancarella<sup>2</sup>, Ilaria D'Elia<sup>3</sup>, Massimo D'Isidoro<sup>2</sup>, Antonio Piersanti<sup>2</sup>, Mario Adani<sup>2</sup>, and Guido Guarnieri<sup>4</sup>

<sup>1</sup>ENEA, Office of Pisa, c/o CNR Research Area

<sup>2</sup>ENEA, Bologna Research Centre

<sup>3</sup>ENEA headquarters, Rome

<sup>4</sup>ENEA Portici Research Centre

**ABSTRACT.** This paper concerns a synthesis of tests performed in order to find the optimal set-up for updating and calculating the new atmospheric transfer matrices (ATMs) in GAINS-Italy over the Italy. ATMs are source-receptor relationships, allowing to quickly estimate the impact of variations of regional emissions on atmospheric pollution and ground deposition in the whole Italian domain. It is presented an update of the previous ATMs, calculated at 20 km resolution. The performed tests demonstrated the possibility to improve the horizontal resolution at 4 km.

## 1 Introduction

Several air pollution policies aim to improve air quality so to mitigate the negative impacts of anthropogenic emissions on human health and the environment. Chemical air quality models represent valuable tools to evaluate the air pollutant concentrations/depositions deriving from anthropogenic emission changes. Although providing detailed information, these models are getting more and more sophisticated, requiring both specialist skills and a considerable amount of computational time, therefore they are not suitable tools to give quick responses about the impacts of different policy scenarios. There is an urgent need for computationally efficient tools, providing an integrated assessment of the different types of impact (both on environment and on health) of the different policies to improve air quality.

The development of simplified relationships between emissions and concentrations/depositions/related impacts are therefore extremely useful tools for a rapid assessment of different air quality policies.

A possible scheme for simplifying the chemical transport and diffusion dynamics is the ATMs (Atmospheric Transfer Matrices) approach, implemented in GAINS (Greenhouse and Air Pollution Interaction and Synergies, [1]). Within the MINNI project (National Integrated Model to support International Negotiation on air pollution issues ([2], [3], [4], [5], [6]), ENEA has already developed an integrated assessment tool, the GAINS-Italy model ([7], [8], [9]), based on ATMs [10]. The present work is an update of [10]: thanks to the massive improvement of CRESCO architecture, we are now able to elaborate the new ATMs in a faster way improving the horizontal resolution from 20 km to 4 km.

## 2 Methodology

Within the ATMs scheme, the source/receptor relationships, built for an air quality model, link the emissions of a certain pollutant in a given geographical area (source) to the relative concentrations/depositions and impacts calculated at a given point (receptor), simplifying the atmospheric processes involved (meteorological dynamics, physical and chemical processes on

pollutants). In the case of the GAINS-Italy model, the source terms are the aggregate emissions on each of the 20 Italian administrative regions, while the receptors coincide with the points of the model calculation grid.

The ATMs describe the response of the concentrations/depositions at a given receptor to the variation of the emissions of each region/source, obtained by applying a perturbation to the emissions of each of the primary compounds of interest, also called precursors of the GAINS indicators: sulphur oxides (SO<sub>x</sub>), nitrogen oxides (NO<sub>x</sub>), ammonia (NH<sub>3</sub>), non-methanic volatile organic compounds (NMVOC) and particulate matter (PM10). The response is calculated by conducting a complete simulation of the AMS-MINNI model for each emitted compound.

ATMs are then a sort of Taylor expansion: emissions of the five GAINS-Italy precursors are grouped at regional level and concentrations/depositions are considered as function of regional emissions only. Concentration/deposition variations over the domain are therefore expressed summing up all the terms of the Taylor expansion. First order approximation requires then one base scenario and 20\*5 variation runs, where 20 is the number of Italian regions and 5 is the number of precursors. 2<sup>nd</sup> order expansion requires (100\*101)/2 terms and so on. Tests were conducted to verify linearity and correlation among the different precursors and impact indicators so to choose the approximation order of the ATMs with a horizontal resolution of 4 km. Most of 2<sup>nd</sup> order coefficients are negligible since SO<sub>x</sub>, PM10, NMVOC give linear dependencies; moreover, most of non linear relations are local, then it is enough to calculate 40 additional coefficients only (20\*2, where 2 is the number of non linear precursors). Table 1 resumes the results of the linearity and correlation tests, conducted over three regions: Lombardia, Lazio and Campania. The expansion may be written as:

$$C_j(x, y) = C_{j,base}(x, y) + \sum_{k=1}^{regions} \sum_{i=1}^{precursors} \left[ A1_{j,k}^i(x, y) \cdot \Delta E_i^k + 0.5 \cdot A2_{j,k}^i(x, y) \cdot (\Delta E_i^k)^2 \right]$$

where (x,y) are the receptor coordinate, k is an index spanning over the 20 regions, i is the precursor index and j the label for GAINS indicators. A1 and A2 are respectively 1<sup>st</sup> and 2<sup>nd</sup> order coefficients and ΔE the emission variation.

**Table 1: Results of linearity and correlation tests per each precursors/indicator with a horizontal resolution of 4 km (in bold the leading terms).**

		Precursors				
		SO <sub>x</sub>	NO <sub>x</sub>	PM10	NH <sub>3</sub>	NMVOC
<b>GAINS Indicators</b>	Sulphur deposition	<b>linear</b>	absent	absent	absent	absent
	Total nitrogen deposition	absent	<b>linear</b>	absent	<b>non linear</b>	absent
	Red. nitrogen deposition	linear	non linear	absent	<b>linear</b>	linear
	O <sub>3</sub> as SOMO35 and AOT40	absent	<b>non linear</b>	absent	absent	<b>linear</b>
	PM2.5	linear	non linear	<b>linear</b>	non linear	linear
	PM10	linear	non linear	<b>linear</b>	non linear	linear
	NO <sub>2</sub>	absent	<b>non linear</b>	absent	absent	linear

The calculation of the ATMs requires significant resources and execution times. In order to optimize the execution times and minimize the overall time for the calculation, some numerical tests were conducted to evaluate the calculation accuracy, aimed to optimize the temporal step of integration (DTS), and the possible improvement of the current ATMs horizontal spatial resolution (20 km). The tests concerned both accuracy relative to a reference emission scenario (BS) and relative to differences between a perturbed (AS) and base emissions scenarios and exploring the new horizontal spatial resolution of 8 km and 4 km. The model setup used to perform all the tests is summarized in Table 2.

The value of 30 s for DTS may be considered the best choice to guarantee the maximum accuracy, up to 4 km horizontal resolution. Table 3 shows the maximum errors obtained on annual averaged fields from the base scenario (BS), varying DTS from 60 s to 200 s; the absolute differences are evaluated by subtracting mean values produced with DTS=30 s, that represents the optimal simulation:

$$\text{diff}=\max\{ | \langle \text{field}(x,y) \rangle_{\text{DTS}} - \langle \text{field}(x,y) \rangle_{\text{DTS}=30\text{s}} | \mid (x,y) \in D \}$$

where the maximum of the absolute values is calculated over the whole grid domain,  $D$ . Negative values in the table mean underestimation induced by non-optimal DTS.

Table 4 shows the maximum errors on the difference between altered case (AS) and BS, in function of DTS, that is:

$$\text{diff}(\text{DTS})=\max\{ | \langle \text{field}(x,y) \rangle_{\text{AS,DTS}} - \langle \text{field}(x,y) \rangle_{\text{BS,DTS}} | \mid (x,y) \in D \}$$

As previously underlined, the maximum of absolute values is computed over the domain and the negative sign means underestimation.

As regards the calculation speed, Fig. 1 illustrates the speeds for three paradigm options: pure MPI, hybrid MPI/OMP with 4 OMP dedicated cores, and hybrid MPI/OMP with 8 OMP dedicated cores. A multiple of 48 core was chosen to occupy the nodes entirely. The aim is to guarantee to complete each simulation in 24 hours: being each annual run splitted in 12 parallel runs, it is enough to get an elapsed time of 2750 s/day. For this purpose, the most efficient configuration, guaranteeing to end each simulation in 24 hours, is to use 48 core with 4 OMP dedicated ones. Pure MPI does not seem much efficient, not being FARM a fully vector code.

**Table 2:** Setup of the model.

Parameters	Description	
	4 km horiz. res.	8 km horiz. res.
Reference Emission Scenario (RES)	ISPRA 2015 national inventory, Emission Manager (EMGR) v. 6.7 intel	
Altered Emission Scenario (AES)	- 25% emissions over Lombardia region	
Meteorology	MINNI scenario 2015	Interpolated from 4km
Version of chemical model (FARM)	4.14	
Parallelization paradigm	Hybrid 48/4 OMP	
Cresco section	Cresco6	
DTS	30, 60, 100, 200, 300 s	
Considered species	O <sub>3</sub> , NO <sub>2</sub> , PM2.5, PM10, TS, TN, TNH at ground	
Postprocessing	Annual averages	

### 3 Conclusions

The illustrated sensitivity tests were carried out to check the possible horizontal spatial resolution of the new ATMs and the most appropriate DTS to use in FARM simulations so to improve the calculation speed now possible thanks to the massive improvement of CRESCO architecture.

The current ATMs have a resolution of 20 km and these tests demonstrated that is now possible to elaborate ATMs with a spatial resolution of 4 km. About the choice of the DTS, more in general, we can affirm that the errors induced by higher DTS are pronounced in BS, but less important in evaluating difference between AS and BS, as reported in Tables 3 and 4, and this seems quite independent from the spatial resolution chosen. In other terms, a short DTS leads to a great accuracy but requires more CPU time; on the contrary, longer DTS is most efficient concerning CPU time but decreases the accuracy. It has been noted that variations of concentrations following emission abatements are less affected by uncertainty with respect to absolute calculations: so, the idea is to

keep the temporal step of 60 s for the complete simulation of the BS, by guaranteeing enough accuracy, and to set DT=150 s for the 140 emission abatement runs (ASs).

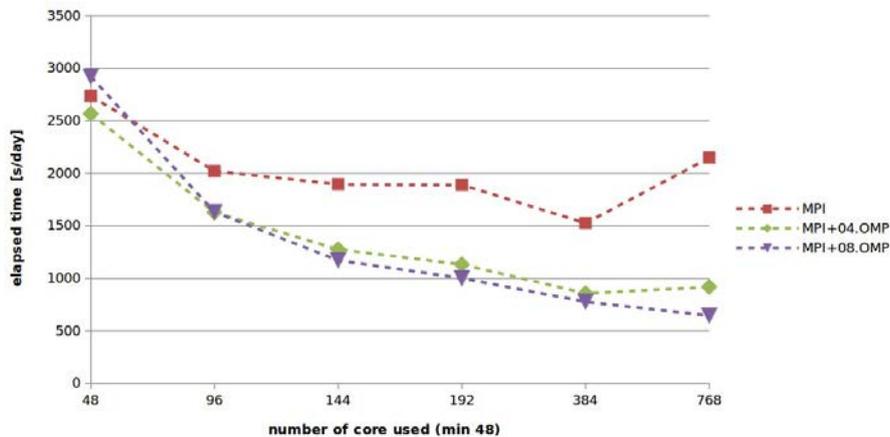
It is worth noting that the possibility to increase DTS allows to go down to 4 km of spatial resolution. In particular, the value DTS=150 s is the maximum value that keep elapsed time under 24 hours on cresco6\_48h24 queue, by using one Cresco6 node only in hybrid parallel run with 4 OMP cores. The choice of using 4 OMP cores is the one that allows to speed up simulations as much as possible.

**Table 3:** BS, annual averages, minimum and maximum values and maximum errors for each species on the domains at 8 and 4 km of resolution (differences DTSXXX-DTS030).

SPEC	UNIT	8KM						4KM				
		MIN	MAX	DTS060	DTS100	DTS200	DTS300	MIN	MAX	DTS060	DTS100	DTS200
O <sub>3</sub>	µg/m <sup>3</sup>	38.2940	94.5680	-1.0290	-2.3560	-5.5500	-8.3510	25.5980	97.8500	1.4930	3.4230	7.2560
NO <sub>2</sub>	µg/m <sup>3</sup>	0.3030	49.0820	0.2840	0.5920	1.1470	1.5600	0.4510	67.7080	0.8890	1.9310	4.1240
PM <sub>25</sub>	µg/m <sup>3</sup>	2.4880	29.0780	-0.0970	-0.2240	-0.5330	-0.8180	2.6360	41.9940	0.2390	0.5610	1.3880
PM <sub>10</sub>	µg/m <sup>3</sup>	3.2490	30.3920	-0.1050	-0.2420	-0.5750	-0.8800	3.4870	44.6320	0.2360	0.5530	1.3790
TS	mg/m <sup>2</sup> /h	0.0094	4.3085	0.0915	0.2160	0.5458	0.9171	0.0114	14.1290	0.5213	1.2523	3.4059
TN	mg/m <sup>2</sup> /h	0.0059	0.0845	-0.0003	-0.0008	-0.0021	-0.0033	0.0077	0.1018	0.0008	0.0018	0.0040
TNH	mg/m <sup>2</sup> /h	0.0004	0.5867	0.0005	0.0011	0.0027	0.0041	0.0003	0.6109	0.0015	0.0034	0.0076

**Table 4:** Differences AS-BS, annual averages, minimum, maximum values and maximum errors for each species on the domains at 8 and 4 km of resolution, in function of DTS.

SPEC	UNIT	8KM						4KM				
		MIN	MAX	DTS060	DTS100	DTS200	DTS300	MIN	MAX	DTS060	DTS100	DTS200
O <sub>3</sub>	µg/m <sup>3</sup>	-0.16840	4.60660	-0.02315	-0.05360	-0.12835	-0.18596	-0.17860	5.06650	0.03806	0.08656	0.16426
NO <sub>2</sub>	µg/m <sup>3</sup>	-7.62630	0.01590	0.02864	0.06611	0.15899	0.23071	-8.77370	0.02916	0.04562	0.10589	0.20668
PM <sub>25</sub>	µg/m <sup>3</sup>	-0.47800	0.00067	0.00511	0.01175	0.02701	0.04004	-0.48553	0.01020	0.00749	0.01702	0.03620
PM <sub>10</sub>	µg/m <sup>3</sup>	-0.47800	0.00073	0.00508	0.01173	0.02699	0.04001	-0.48554	0.01025	0.00750	0.01700	0.03623
TS	mg/m <sup>2</sup> /h	-0.00003	0.00002	0.00003	-0.00019	-0.00003	-0.00036	-0.00007	0.00002	0.00011	0.00007	0.00048
TN	mg/m <sup>2</sup> /h	-0.00926	0.00000	0.00006	0.00013	0.00034	0.00053	-0.00987	0.00000	0.00007	0.00018	0.00042
TNH	mg/m <sup>2</sup> /h	-0.00025	0.00138	-0.00003	-0.00003	-0.00007	-0.00009	-0.00038	0.00150	0.00002	0.00004	0.00007



**Fig.1:** Calculation speed (elapsed time) in seconds per day for the simulation with a horizontal spatial resolution of 4x4 km.

## References

[1] Amann, M., Bertok, J., Borcken-Kleefeld, J., Cofala, J., Heyes, C., Höglund-Isaksson, L., Klimont, Z., Nguyen, B., Posch, M., Rafaj, P., Sandler, R., Schöpp, W., Wagner, F., Winiwarer, W., 2011. Cost-effective control of air quality and greenhouse gases in Europe: modeling and policy applications. *Environmental Modelling Software*, 26, pp. 1489–1501.

[2] Mircea, M., Zanini, G., Briganti, G., Cappelletti, A., Pederzoli, A., Vitali, L., Pace, G., Marri, P., Silibello, C., Finardi, S., Calori G., 2011. Modelling Air Quality over Italy with MINNI Atmospheric

Modelling System: from regional to local scale. In: Steyn DG and Trini Castelli S, editors. Air Pollution Modeling and its Application XXI, 491-498, ISSN 1874-6519, [http://dx.doi.org/10.1007/978-94-007-1359-8\\_82](http://dx.doi.org/10.1007/978-94-007-1359-8_82).

[3] Mircea, M., Ciancarella, L., Briganti, G., Calori, G., Cappelletti, A., Cionni, I., Costa, M., Cremona, G., D'Isidoro, M., Finardi, S., Pace, G., Piersanti, A., Righini, G., Silibello, C., Vitali, L., Zanini, G., 2014. Assessment of the AMS-MINNI system capabilities to predict air quality over Italy for the calendar year 2005. *Atmospheric Environment*, 84, 178–188, ISSN 1352-2310, <http://dx.doi.org/10.1016/j.atmosenv.2013.11.006>.

[4] Mircea, M., Grigoras, G., D'Isidoro, M., Righini, G., Adani, M., Briganti, G., Ciancarella, L., Cappelletti, A., Calori, G., Cionni, I., Finardi, S., Larsen, B.R., Pace, G., Perrino, C., Piersanti, A., Silibello, C., Zanini, G., 2016. Impact of grid resolution on aerosol predictions: a case study over Italy. *Aerosol and Air Quality Research*, 16,1253–1267, <https://doi.org/10.4209/aaqr.2015.02.0058>.

[5] Ciancarella, L., Adani, M., Ciucci, A., Cremona, G., D'Isidoro, M., Mircea, M., Piersanti, A., Russo, F., Vitali, L., Righini, G., Zanini, G., Briganti, G., Cappelletti, A., D'Elia, I., 2016. La simulazione nazionale di AMS-MINNI relativa all'anno 2010. Simulazione annuale del Sistema Modellistico Atmosferico di MINNI e validazione dei risultati tramite confronto con i dati osservati. ENEA Technical Report, RT/2016/12/ENEA, <http://openarchive.enea.it/bitstream/handle/10840/7628/RT-2016-12-ENEA.pdf>

[6] Vitali, L., Adani, M., Briganti, G., Cappelletti, A., Ciancarella, L., Cremona, G., D'Elia, D'Isidoro, M., Guarnieri, G., Mircea, M., Piersanti, A., Righini, G., Russo, F., Villani, M.G., Zanini, G., 2019. AMS-MINNI National Air Quality Simulation on Italy for the calendar year 2015. Annual air quality simulation of MINNI Atmospheric Modelling System: results for the calendar year 2015 and comparison with observed data. ENEA Technical Report, RT ENEA/15/2019.

[7] D'Elia, I., Piersanti, A., Briganti, G., Cappelletti, A., Ciancarella, L., Peschi, E., 2018. Evaluation of mitigation measures for air quality in Italy in 2020 and 2030. *Atmospheric Pollution Research*, 9, 977-988, <http://dx.doi.org/10.1016/j.apr.2018.03.002>.

[8] D'Elia, I., Bencardino, M., Ciancarella, L., Contaldi, M., Vialetto, G., 2009. Technical and Non-Technical Measures for air pollution emission reduction: The integrated assessment of the regional Air Quality Management Plans through the Italian national model. *Atmospheric Environment*, 43, 6182-6189, ISSN 1352-2310, <http://dx.doi.org/10.1016/j.atmosenv.2009.09.003>.

[9] Ciucci, A., D'Elia, I., Wagner, F., Sander, R., Ciancarella, L., Zanini, G., Schöpp, W., 2016. Cost-effective reductions of PM<sub>2.5</sub> concentrations and exposure in Italy. *Atmospheric Environment*, 140, 84-93, <https://doi.org/10.1016/j.atmosenv.2016.05.049>.

[10] Briganti, G., Calori, G., Cappelletti, A., Ciancarella, L., D'Isidoro, M., Finardi, S., Vitali, L., 2011. Determination of Multi-year Atmospheric Transfer Matrices for GAINS-Italy Model. In: *High Performance Computing on CRESCO Infrastructure: Research Activities and Results*, vol. 2010-2011, pp. 45-51. Report Cresco.

# FROM PROTOTYPE TO PRODUCTION: MONTE CARLO STUDIES FOR UPDATING THE ENEA NEUTRON ACTIVE INTERROGATION SYSTEM

Giada Gandolfo<sup>1\*</sup>, Luigi Lepore<sup>2</sup>, Nadia Cherubini<sup>2</sup>, Giuseppe A. Marzo<sup>2</sup>

<sup>1</sup>*ENEA, Italian National Agency for New Technologies, Energy and Sustainable Economic Development, C.R. Frascati, Rome, Italy*

<sup>2</sup>*ENEA, Italian National Agency for New Technologies, Energy and Sustainable Economic Development, C.R. Casaccia, Rome, Italy*

**ABSTRACT.** The Italian National Agency for New Technologies, Energy and Sustainable Economic Development (ENEA) developed a new device to improve CBRNe resilience, the Neutron Active Interrogation system (NAI). The original experimental setup was tested on the field during the live demo open to the public at the EDEN Project Demonstration occurred in September 2015 at ENEA Frascati Research Centre in Rome. Since then, the setup has been modified to improve the device detection capabilities. In 2018 the device has been selected for the ENEA Proof of Concept Program with the aim to finance the upgrade of the prototype towards industrial production.

## 1 Introduction

Since September 2013, ENEA - the Italian National Agency for New Technologies, Energy, and Sustainable Economic Development – Nuclear Material Characterization and Radioactive Waste Management Laboratory has been involved in the EDEN (End-user DEMo for cbrNe) project in order to design and realize the Neutron Active Interrogation (NAI) system whose purpose is the non-destructive inspection of samples suspected of containing fissile or fertile material and explosives [1]. Over the years several upgrades have been identified and implemented on the first prototype [2]. In 2018 ENEA financed a Proof of Concept (PoC) Program to verify the technological feasibility and the scaling-up of technologies with a low degree of technological maturity developed in its own laboratories. The ENEA NAI system has been selected to be eligible for funding in this program, which involve the collaboration with industrial partners to encourage the technology transfer. This paper describes the feasibility study on which the improvement of NAI prototype is based.

## 2 Project objectives

The NAI device is currently composed by:

1. a compact d-t neutron generator;
2. a sample cavity composed by 4 polyethylene blocks;
3. ten He-3 proportional counters housed in two of the polyethylene blocks (detection blocks);
4. 1 mm thick cadmium sheets surrounding lateral surfaces of detection blocks to cut the background effects caused by thermalized interrogation neutrons;
5. moderating structure surrounding the sample;
6. electronics for data acquisition.

Fast neutrons emitted by the d-t neutron generator thermalize in moderator and, if fissile material is present in the inquired object, induce fission on fissile nuclei generating fission neutrons detected by the He-3 tubes.

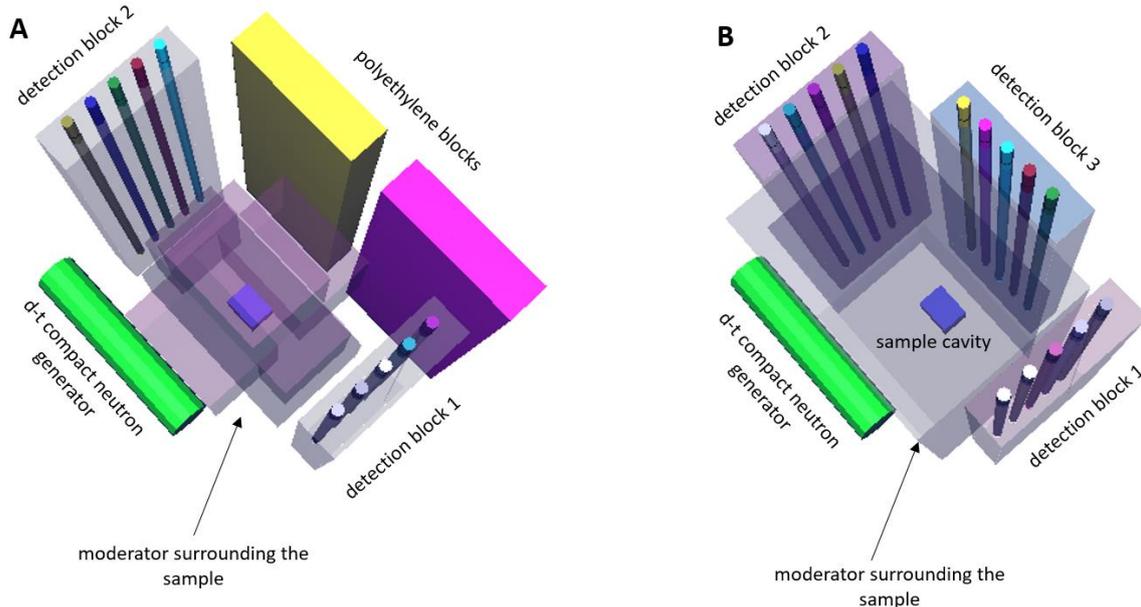
The operating principle of NAI device is the Differential Die-Away time Analysis, an active neutron interrogation technique based on the difference between the die-away time of fast interrogation neutrons and prompt fission neutrons induced by thermal neutrons in the moderating system [3, 4, 5]. The current prototype is transportable and able to detect 2 g of  $^{235}\text{U}$  contained in a sample with maximum dimension of about 7.5 cm x 31 cm x 28 cm in about 2 minutes. The device is remotely manageable through a software for the neutron generator control, data acquisition and analysis.

In order to upgrade the NAI prototype to a near-industrial production level the following objectives have been identified:

1. to enlarge the neutron interrogation cavity for detecting the presence of fissile material in an object comparable in size to that of a large luggage;
2. to reduce the minimum detectable quantity of fissile material;
3. to reduce the weight of the device to improve its transportability;
4. to reduce the time needed for the detection of the dangerous material;
5. to link the recorded signal to the fissile mass independently from the matrix in which it is incorporated.

### 3 Monte Carlo feasibility study

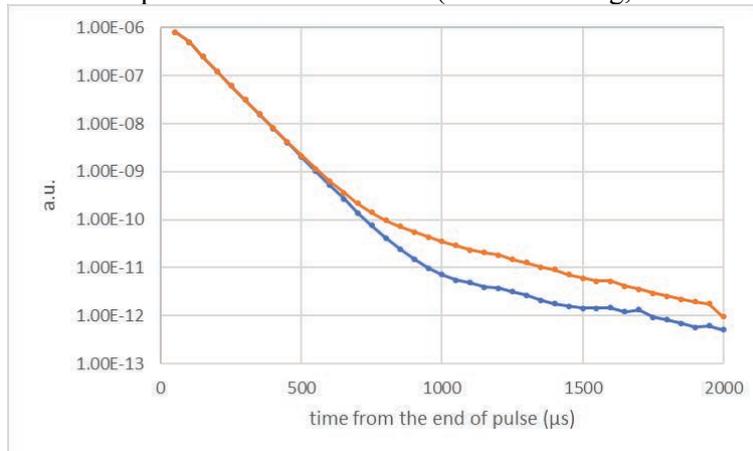
The NAI prototype upgrade has been based on a preliminary feasibility study through MCNPX (Monte Carlo N-Particle eXtended) simulations by means of a large use of HPC (High Parallel Computing) resources. In order to reach the objectives above, about 60 simulations have been carried out. Each simulation, involving  $10^{10}$  particle sources, has been performed through a parallel job executed by 256 processors with a CPU time of about  $2 \times 10^5$  s. The starting point NAI layout and the final updated layout resulting from the simulation study are shown in Fig. 1.



**Figure 1** – Monte Carlo feasibility study: starting point layout (A) and final updated layout (B).

- The current prototype, with the characteristics described in section 2, is the starting point for the feasibility study (Fig.1A). DDAA curves are reported in Fig.2, the blue curve corresponds to the signal recorded without the sample (background), the orange curve is related to the signal in presence of a sample containing about 6 g of  $^{235}\text{U}$ . The signal to

background ratio (S/B) in the time interval 1000  $\mu\text{s}$  – 1500  $\mu\text{s}$  from the end of the interrogation neutron pulse in ideal conditions (without ceiling, floor and walls) is S/B=4.93.



**Figure 2** – DDAA curves related to the starting point layout. The blue curve corresponds to the background, the orange curve to the signal in presence of a sample containing about 6 g of  $^{235}\text{U}$ . S/B=4.93 in the time interval 1000  $\mu\text{s}$  – 1500  $\mu\text{s}$ .

- The upgraded device resulting from the Monte Carlo simulation study is shown in Fig.1B. The following changes have been implemented:
  - ✓ addition of a detection block embedding 5 He-3 proportional counters. The number of detectors consequently increases from 10 to 15;
  - ✓ selection of a different material and reduction of the volume of the moderator surrounding the sample. The volume of the sample cavity increases by a factor of about 20, from 7.5 cm x 31.4 cm x 28 cm to 32.6 cm x 47 cm x 90 cm;
  - ✓ optimization of cadmium sheets thickness and position.

The signal to background ratio in the time interval 1000  $\mu\text{s}$  – 1500  $\mu\text{s}$  from the end of the interrogation neutron pulse in a real environmental scenario (a bunker with 20 cm thick concrete walls) is S/B=19.75, about 4 times higher respect to the original layout.

## 4 Conclusions

This paper has described the identification through Monte Carlo simulations of the best ENEA NAI device upgrade reachable with the fund made available by the PoC Program. Several Monte Carlo simulations have been carried out and the following improvements have been identified:

- ✓ addition of a detection block embedding 5 He-3 proportional counters;
- ✓ selection of a different material and reduction of the volume of the moderator surrounding the sample;
- ✓ optimization of cadmium sheets thickness and position.

Such improvements will be implemented to upgrade the NAI prototype to a near-industrial production level, in particular they will allow to enlarge the neutron interrogation cavity dimensions (in order to analyse a large luggage), to reduce the minimum detectable quantity of fissile material and the time needed for the detection and to improve the transportability.

The computing resources and the related technical support used for this work have been provided by CRESCO/ENEAGRID High Performance Computing infrastructure and its staff [6]. CRESCO/ENEAGRID High Performance Computing infrastructure is funded by ENEA, the Italian National Agency for New Technologies, Energy and Sustainable Economic Development and by Italian and European research programmes, see <http://www.cresco.enea.it/english> for information.

## References

- [1] N. Cherubini, A. Dodaro, G. Gandolfo, L. Lepore, G.A. Marzo, E. Piccinelli, R. Remetti, Field Prototype of the ENEA Neutron Active Interrogation Device for the Detection of Dirty Bombs. *Challenges* 2016, 7(2), 17; <http://dx.doi.org/10.3390/challe7020017>
- [2] N. Cherubini, A. Dodaro, G. Gandolfo, L. Lepore, G.A. Marzo, E. Piccinelli, R. Remetti, The neutron active interrogation system for in-field detection of transuranic-based radioactive dispersal devices for security applications, 26th ICONE, London, UK. Published in *ASME Proceedings, Nuclear Safety, Security, and Cyber Security (Vol. 4)*, July 2018.
- [3] R. Remetti, G. Gandolfo, L. Lepore, N. Cherubini, In field application of differential Die-Away time technique for detecting gram quantities of fissile materials, *Nuclear Instruments & Methods in physics research. section A*, vol 870, pages 123-130, 2017
- [4] Jordan, K.A., Gozani, T., Pulsed neutron differential die away analysis for detection of nuclear materials, *Nucl. Instrum. Methods Phys. Res. B*, 2007, 261, 365-368
- [5] Jordan, K.A.; Gozani, T.; Vujic, J. Differential die-away analysis system response modeling and detector design. *Nucl. Instrum. Methods Phys. Res. A* 2008, 589, 436–444
- [6] F. Iannone et al., "CRESCO ENEA HPC clusters: a working example of a multifabric GPFS Spectrum Scale layout," 2019 International Conference on High Performance Computing & Simulation (HPCS), Dublin, Ireland, 2019, pp. 1051-1052, doi: 10.1109/HPCS48598.2019.9188135.

# APPLICATIONS OF HYBRID SIMULATION TO FAST FREQUENCY CHIRPING OF REVERSED SHEAR ALFVÉN EIGENMODE

Tao Wang<sup>1\*</sup>

<sup>1</sup>*Advanced Energy Research Center, Shenzhen University, Shenzhen 518060, China*

**ABSTRACT.** The main research activities in the year 2019 are reported. The hybrid MHD-gyrokinetic code has been applied to analyse the resonant interactions between the reversed shear Alfvén eigenmode and fast ions. The results agree qualitatively with general experimental observations and theoretical understanding.

## 1 Introduction

The nuclear fusion research in the tokamak configuration requires external power input such as neutral beam injection for plasma heating and current drive. Fast ions are routinely produced in this process, and could drive global plasma instabilities in the Alfvénic frequency range [1]. Understanding the behaviour of the fast ions in fusion plasmas is crucial with both academic and realistic interests, where numerical simulations play a key role to connect theory and experiment.

When high external power is applied in the plasma current ramp-up stage, the safety factor  $q$  profile often includes an off-axis minimum  $q_{\min}$ . The reversed shear Alfvén eigenmode (RSAE) exhibits interesting and unique behaviours. It is well-known that the RSAE frequency predominantly sweeps upward following the relatively slow (characteristic timescale of the order of 10-100 ms) decrease of  $q$  profile. In addition, with high external power injection, fast frequency chirping (typically within 1 ms) is also observed. The intriguing fast frequency chirping is investigated by hybrid MHD-gyrokinetic code (HMGC) [2] simulations using realistic parameters [3]. It is suggested that fast chirping could take place in the non-adiabatic regime, associated with the “non-perturbative” effect of fast ions [1,4].

## 2 Simulation setup

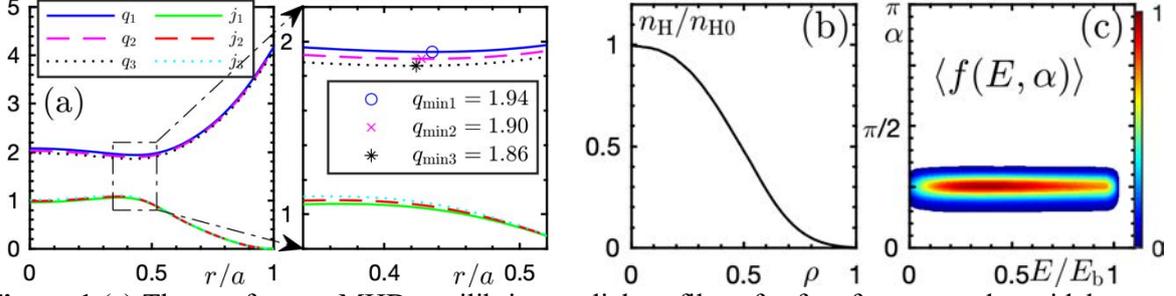
The physics model of HMGC uses reduced ideal MHD equations coupled with the fast ion pressure tensor [2]. Thus, the non-perturbative effect of fast ions is fully retained, namely, the dispersive properties of the MHD waves including the mode structure and real frequency, are determined by both the fluid response via the equilibrium profiles, as well as the fast ions.

Compared with the corresponding sections on RSAE analyses in the previous annual reports (2016-2018) using a numerically generated equilibrium, the simulation scenario in this work is designed to be closer to experimental conditions. The HL-2A tokamak with circular poloidal cross sections is used as a reference, with  $a = 0.40$  m,  $R_0 = 1.65$  m, on-axis magnetic field  $B_0 = 1.3$  T. The non-uniform MHD equilibrium is represented by the radial profiles of  $q$  and thermal ion density  $n_i$ . Since the RSAE frequency could significantly change with a small variation of  $q_{\min}$ , it is necessary to consider the slowly varying  $q$  profiles in the simulation setup, so as to capture the broad RSAE spectrum. This is achieved by controlling the plasma current profiles, as shown in figure 1(a). The increment of the total plasma current corresponds to a global decrease of  $q$  profile, which directs the RSAE frequency up-

---

<sup>1</sup> E-mail: t\_wang@zju.edu.cn

sweeping. Meanwhile, thermal ions are assumed to be Deuterium with a parabolic profile  $n_i(r)/n_{i0} = 1 - 0.7(r/a)^2$ , with  $n_{i0} = 2 \times 10^{19} \text{ m}^{-3}$  on the axis. The thermal plasma compressibility and diamagnetic effects are not taken into account in this work, i.e., thermal plasmas are assumed to be cold.



**Figure 1** (a) Three reference MHD equilibrium radial profiles of safety factor  $q$  and toroidal current density  $j$  (normalized to  $B_0/R_0$ ), labelled by the  $q_{\min}$  values. (b) Fast ion density profile as a function of a radial-like flux coordinate  $\rho$ . (c) Fast ion velocity space distribution function versus energy  $E$  and pitch angle  $\alpha = \cos^{-1}(v_{\parallel}/v)$ .

The fast (“hot”) ions are modelled as tangentially co-injected Deuterium beam ions, with a birth energy  $E_b = 45 \text{ keV}$ . Their initial distribution function is shown in figures 1(b) and (c). They are numerically fitted to a Monte Carlo code calculation, however, the pitch angle scattering effect in the low energy range is mostly neglected by assuming a narrow pitch angle distribution. This is restricted by the initialization procedure of kinetically treated species allowed by HMGC. The fast ion on-axis density is taken to be  $n_{H0} = 10^{18} \text{ m}^{-3}$ , which is usually an overestimation for neglecting the significant internal redistribution by the Alfvénic modes.

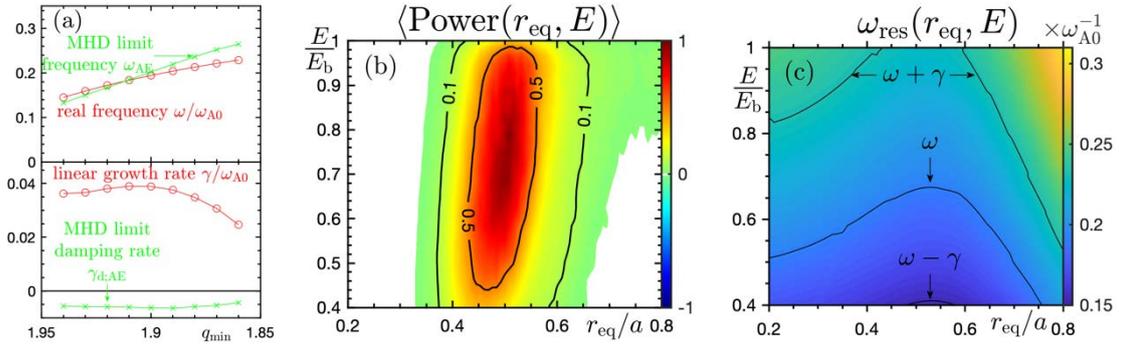
In principle, HMGC considers only the large aspect ratio approximation and fixed MHD equilibrium. For the present work using a realistic tokamak configuration, exponentially growing numerical instability is known to be problematic in the late nonlinear stage of a simulation case. The numerical diffusion parameters, numbers of grids and kinetic particle markers have been tested extensively in order to minimize such numerical artifact and to ensure the convergence. In addition, to verify the robustness of the conclusions made in this work, we have also performed parameter sensitivity scans over  $q_{\min} = 1.99 - 1.83$ , toroidal mode number  $n = 1 - 5$ ,  $B_0 = 1.0 - 1.3 \text{ T}$ ,  $n_{H0}/n_{i0} = 0.03 - 0.05$ ,  $E_b = 45 - 90 \text{ keV}$ , central pitch angle  $\cos(\alpha_{inj}) = 0.70 - 0.85$ . The following section reports the most representative results [3] with  $q_{\min}$  in the range of  $1.94 - 1.86$  (9 cases with  $\Delta q_{\min} = 0.01$ ),  $n = 3$  and other parameters as described above.

### 3 Simulation results

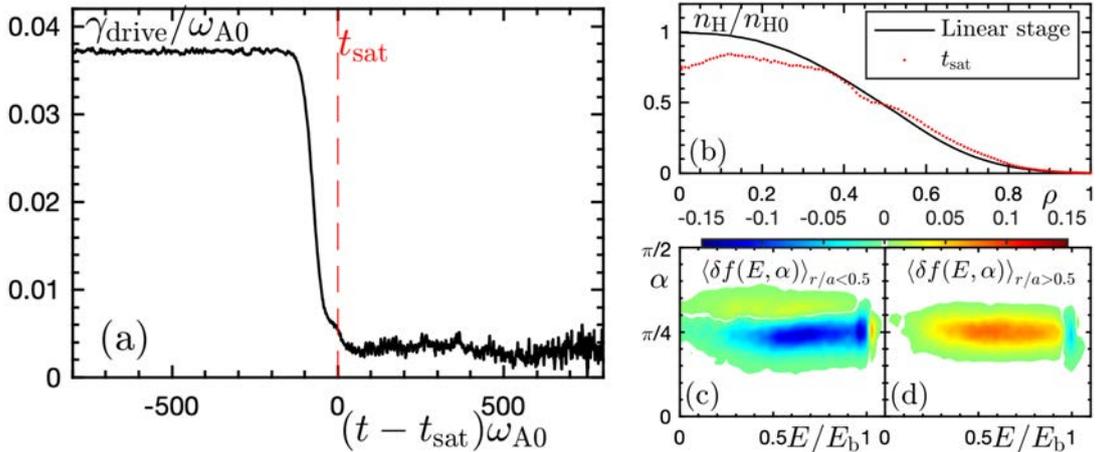
For the 9 cases with the  $q$  profile being the only different input parameter, figure 2 (a) shows the linear RSAE spectrum with and without fast ions, the simulations without fast ions are performed by the antenna excitation technique. One immediately notices the significant effect of the fast ion induced non-perturbative frequency shift, which can be understood from the perspective of the underlying wave-particle resonance condition. Due to the strong fast ion drive and the resonance broadening effect of finite  $\gamma/\omega$ , fast ions in a broad phase space range are resonant with the RSAE, as shown in figures 2 (b) and (c) for the case with  $q_{\min} = 1.90$  as an example. The linear mode frequency is weighted by the contribution of all resonant phase space portions so as to maximize the total power transfer. Thus, for the sub-Alfvénic fast ions with a relatively low energy, their resonance frequencies for most efficient power transfer restrict the linear RSAE frequency into a narrow band. In addition, the non-

perturbative effect is also apparent in the radial mode structure. The characteristic radial mode width is extended in the presence of fast ions with large drift orbit widths (a characteristic value  $\rho_d/a \sim 0.15$ ).

It is also interesting to observe in figure 2 (c) that the resonance frequency  $\omega_{\text{res}}$  is only weakly dependent on the radial coordinate. This implies that in the early nonlinear stage, a linearly resonant ion with  $\omega_{\text{res;lin}} \sim \omega$  will stay roughly resonant with the wave despite a local orbit displacement. Thus, for the nonlinear saturation with a significant drop of the total drive intensity (see figure 3 (a)), the resonant fast ions need to be non-locally convected, such that their drive intensity is reduced by them experiencing the radially non-uniform mode amplitude, rather than being shifted out of phase with the wave. In other words, the RSAE saturates predominantly via the “radial decoupling” mechanism [1,4], and the global fast ion transport is convective and coherent. The coherent fast ion flux induced by the RSAE saturation is shown in figure 3 (b)-(d), where the distortions to the integrated fast ion profiles are quite significant, since a broad range of the phase space is involved (figure 2 (b)).

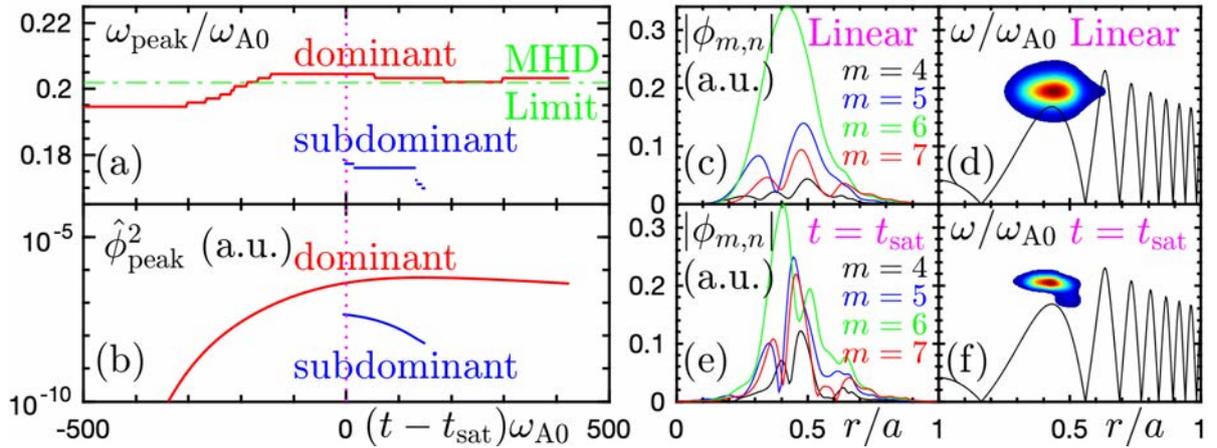


**Figure 2** (a) Linear RSAE frequencies (upper) and growth/damping rates (lower) with (red) and without (green) fast ions. (b) For the case with  $q_{\text{min}} = 1.90$  in the linear stage, the power transfer from the fast ions to the RSAE in  $(r_{\text{eq}}, E)$  space, where  $r_{\text{eq}}$  refers to a particle’s radial coordinate at the low field side midplane in the equilibrium orbit. (c) For the same case, numerically calculated fast ion resonance frequency, with the contour line of the peak mode frequency and the range of frequency spectrum  $\omega \pm \gamma$  indicated.



**Figure 3** (a) Time evolution of total fast ion drive intensity  $\gamma_{\text{drive}}$ , calculated from the integrated and renormalized power transfer. The saturation time  $t_{\text{sat}}$  is indicated. (b) Comparison the fast ion radial profiles in the linear stage and  $t_{\text{sat}}$ . (c) and (d) The velocity space fast ion redistributions at  $t_{\text{sat}}$  in, the inner and outer halves of the torus, respectively, normalized to the peak value in figure 1 (c).

The RSAE dynamics during the saturation stage is crucially related with the fast ion transport described above. It is interesting to observe a splitting of mode structure in the  $(r, \omega)$  space, as shown in figure 4. The dominant branch remains at essentially the same radial position, with up-chirping frequency and reduced radial mode width. All these characteristics agree with the spectrum obtained in the simulation without fast ions. Thus, this branch is interpreted as the “relaxation” branch, since it results from the significant fast ion transport shown in figure 3 and the weakening of their non-perturbative effect. The relaxation branch dominates the post saturation stage, in excellent agreement with the theoretical understanding of the radial localized (around  $q_{\min}$ ) character of the RSAE, and the long timescale frequency spectrum in close proximity to the MHD prediction as observed experimentally (often applied as the MHD spectroscopy). Meanwhile, a subdominant branch emerges below the relaxation one, it instead goes downwards in frequency and outwards in radial direction. This can be understood from the resonant fast ion convective transport, as the non-perturbatively excited wave tends to modify its frequency and radial structure so as to stay resonant. However, this “convective” branch is strongly damped by coupling with the shear Alfvén continuous spectrum, and becomes rather unimportant. Note that for both branches, the fast frequency chirping during the saturation stage is in the non-adiabatic regime [4]. That is, the phase space resonance structure sweeps in a rate comparable with its rotation rate, which can be estimated from the timescale of fast ion transport. Similar behaviors are observed in all other simulation cases, where the nonlinear RSAE dynamics is dominated by the relaxation branch, with fast frequency chirping upwards or downwards depending on the direction of fast ion induced frequency shift in the linear stage.



**Figure 4:** For two branches emerging during saturation stage, time evolutions of their frequencies (a) and intensities (b). Frames (c)-(f) compare the mode structures in the linear stage and saturation time.

## 4 Summary

The simulations could serve as a qualitative explanation of the fast RSAE frequency chirping observed experimentally, with the dominant physics mechanism in line with the theoretical understanding of non-perturbative wave-particle interaction in the non-adiabatic regime [1,4].

## Acknowledgment

The computing resources and the related technical support used for this work have been provided by CRESCO/ENEAGRID High Performance Computing infrastructure and its staff [5].

## References

- [1] Liu Chen and Fulvio Zonca. Physics of Alfvén waves and energetic particles in burning plasmas. *Rev. Mod. Phys.* **88**, 015008 (2016).
- [2] S. Briguglio *et al.* Hybrid magnetohydrodynamic-gyrokinetic simulation of toroidal Alfvén modes. *Phys. Plasmas* **2**, 3711 (1995).
- [3] Tao Wang *et al.* Dynamics of reversed shear Alfvén eigenmode and energetic particles during current ramp-up. *Nucl. Fusion* **60**, 126032 (2020).
- [4] F. Zonca *et al.* Nonlinear dynamics of phase space zonal structures and energetic particle physics in fusion plasmas. *New J. Phys.* **17**, 013052 (2015).
- [5] F. Iannone *et al.*, CRESCO ENEA HPC clusters: a working example of a multifabric GPFS Spectrum Scale layout. 2019 International Conference on High Performance Computing & Simulation (HPCS), Dublin, Ireland, 2019, pp. 1051-1052, doi: 10.1109/HPCS48598.2019.9188135.

# MONTE CARLO CALCULATION OF THE $k_{Q,Q_0}$ QUALITY CORRECTION FACTORS FOR A MICRO IONIZATION CHAMBER FOR RADIOTHERAPY DOSIMETRY

Alonzo Massimo<sup>1\*</sup>, Pimpinella Maria<sup>1</sup>

<sup>1</sup> ENEA, National Institute of Ionizing Radiation Metrology, ENEA-FSN-INMRI'

**ABSTRACT.** We report on Monte Carlo simulations for a Semiflex 3D type PTW 31021 ionization chamber. Such chamber is constituted by an air sensitive volume of  $0.07 \text{ cm}^3$  so it is especially suitable for dosimetry requiring high spatial resolution. We calculate the beam quality (Q) correction factors  $k_{Q,Q_0}$  for radiotherapy photon beams with respect to the  $^{60}\text{Co}$  reference gamma beam (quality  $Q_0$ ) and according to the ICRU 90 improvements.  $k_{Q,Q_0}$  factors are determined in the  $TPR_{20,10}$  range 0.66-0.80, emphasizing the effects on the simulation results of using different models of the radiation source as input for the simulations.

## 1 Introduction

Ionization chambers are used to measure the output of clinical accelerators and thus to determine the absorbed dose that will be delivered to a patient during a radiotherapy treatment. Therefore accurate determination of correction factors needed for ionization chamber dosimetry is extremely important for successful radiotherapy treatments.

Since radiotherapy ionization chambers are currently calibrated in a  $^{60}\text{Co}$  reference gamma beam (reference quality  $Q_0$ ), their use in beams of different energy (quality Q), such as beams produced by clinical accelerators, would imply the introduction of a systematic bias on the measured absorbed dose. To compensate for it, the quality correction factor  $k_{Q,Q_0}$  accounting for the change of the chamber response with beam energy must be determined and applied.

The Monte Carlo approach is nowadays robust enough to grant accurate simulation of ionization chamber response in different radiation beams, and then accurate  $k_{Q,Q_0}$  determination, as long as reliable models of both detector and radiation source are developed. This requirement is even more important for very small detectors, such as the ionization chamber considered in this work, for which material and shape of all detector components can significantly affect the measured signal.

The  $k_{Q,Q_0}$  definition is included in the IAEA Technical Report TRS 398 [1] which is intended to harmonize worldwide some key parameters definitions and to provide the best policy to measure (or calculate) them. It is a code of practice for dosimetry based on standards of absorbed dose to water ( $D_w$ ). This publication, firstly reported in 2000 and then revised in 2006, provides also tables reporting the  $k_{Q,Q_0}$  factors for the most common chambers used in radiotherapy dosimetry.

The chamber considered in this report for the  $k_{Q,Q_0}$  determination in accelerator photon beams has been introduced on the market in recent years and therefore it is not reported in [1], so, at present, well established  $k_{Q,Q_0}$  are still not available. However, some experimental and numerical determinations are reported in recent publications [2, 3]. These data will be used as a comparison pattern in the present report.

In the following, it will be proposed the  $k_{Q,Q_0}$  calculation based on Monte Carlo simulations with two different models of the input radiation source for both the Q and  $Q_0$  qualities. The first type of input source is a phase-space file derived from a previous Monte Carlo simulation of the radiation transport through the accelerator head (or the  $^{60}\text{Co}$  irradiator) and the collimation system. The phase-space file contains information about energy, direction and position of all particles in the useful beam, therefore it mimics the real beam characteristics, with an accuracy which depends on the number of recorded particles. The second type of input source consists of a point radiation source that emits particles isotropically and is collimated to irradiate only

---

<sup>1</sup>Corresponding author. E-mail: massimo.alonzo@enea.it

an area of 100 cm<sup>2</sup> at a distance of 100 cm. Energy of each simulated particle is sampled according to a given spectral distribution.

## 2 Quality factor $k_{Q,Q_0}$ calculation

According to [4], the quality factor  $k_{Q,Q_0}$  can be determined by Monte Carlo calculation as:

$$k_{Q,Q_0} = \frac{D_{w,Q}/D_{ch,Q}}{D_{w,Q_0}/D_{ch,Q_0}}$$

$D_w$  and  $D_{ch}$  are the absorbed dose to water at the measurement depth and the average absorbed dose in the active volume of the ionization chamber, respectively. The subscripts  $Q$  and  $Q_0$  stand for the generic quality (energy) and the reference one.

In this work  $D_w$  and  $D_{ch}$  are obtained simulating by Monte Carlo method the irradiation of the PTW 31021 chamber in a water phantom (a cube of water with side 30 cm) with photon beams in the energy range from <sup>60</sup>Co (1.25 MeV) to 25 MV.

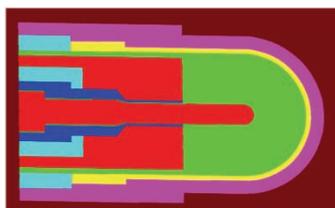
## 3 Monte Carlo and calculation details

In this report we refer to the Monte Carlo code system EGSnrc (<https://nrc-cnrc.github.io/EGSnrc/>) [5] which has been specifically designed and optimized to simulate the ionization chamber response. Moreover, we refer to the 2019 version that we have installed to keep into account the whole of the latest code improvements. It introduces also several updates regarding, mainly, data on photon cross sections and stopping powers. These updates, reported in the ICRU 90 [6], have been derived from measurements and it is expected that they affect the quality factors calculation to a certain extent [4].

The `egs_chamber` application [7] has been used to calculate the deposited energy in the air cavity of the PTW31021 chamber modelled with the `egs++` geometry package and reported in Fig.1 as longitudinal section. Chamber modelling stage is quite critical as it is necessary to introduce some geometrical approximations due to the lack of details in the provided schemes. Furthermore, EGSnrc offers a wide library of materials but often detectors contain non standard compounds for which it is necessary to create new, and approximated, models to establish the basic data for radiation interactions. Both issues can produce significant effects on the calculated absorbed dose.

Absorbed dose to water is scored inside a water disk having dimensions: radius  $R=0.25$  cm and height  $h=0.025$  cm, while absorbed dose in the ionization chamber is scored in an active volume of about 0.07 cm<sup>3</sup>. This volume (radius 2.4 mm and length 4.8 mm) is internally limited by an aluminium electrode and externally by a graphite sheet ( $\rho=1.85$  g cm<sup>-3</sup>) surrounded by a PMMA envelope (see Fig.1).

Both the water disk and the ionization chamber, are placed inside a water phantom at depth of 5 cm and 10 cm for the <sup>60</sup>Co and accelerator beams, respectively.



*Fig. 1 Longitudinal section of the chamber scheme simulated by the package `egs++`. The active volume (green) has radius 2.4 mm and length 4.8 mm.*

To avoid using a huge number of histories to achieve a small statistical uncertainty  $\sigma$ , and to reduce the simulation time  $T$ , Variance Reduction Techniques have been set up. Specifically, photon cross section enhancement (XCSE) and Russian Roulette (RR). Cross section enhancement factors (CS) of 128 and 64 were

applied, below and upper 10 MV, respectively. Using a single processor, the simulation time for the absorbed dose ranges from about  $T = 176$  hours for the water disk to about  $T = 647$  hours for the chamber. These values are referred to the case of the highest photon energy, a number of histories equal to  $5E9$  and an uncertainty target below 0.1%. These simulations have been more conveniently realized using 50 processors in parallel, on CRESCO 6/ENEAGRID HPC infrastructure [8], reaching a global uncertainty of about 0.04% and 0.05% for the water disk and the air camera respectively.

Regarding the photon sources employed, so the qualities we want to correct for, they are given in terms of phase-space files for beams produced by clinical accelerators and they are placed to design the correct field size dimension on the phantom surface. Each available phase-space file was used as input radiation source into the simulation as well as to derive the spectral distribution to be used for the simulation with the collimated point source. According to the TRS 398, except for the  $^{60}\text{Co}$  beam, the photon beam quality is specified in terms of the Tissue-Phantom-Ratio  $TPR_{20,10}$ .

## 4 Results

Exponential fit of the calculated values of  $k_{Q,Q_0}$  for both the cases, collimated point-source and phase-space file, are shown in Fig. 2. They are compared with the published data in [2] also obtained by means of Monte Carlo simulations with the EGSnrc code system. In Fig. 2 the set of values from [2] refers to conventional filtered photon beams as those used in the present work (i.e. data for flattening filter free beams are not considered). Each of the three series show a monotonic decreasing trend, that is well described by an exponential fit.

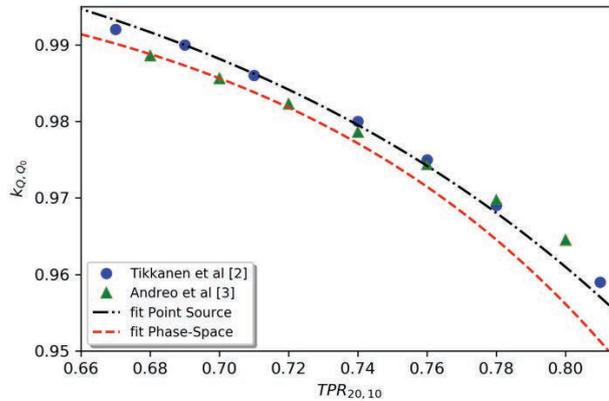


Fig. 2 Monte Carlo quality correction factors  $k_{Q,Q_0}$  for high energy photon beam qualities, described by the corresponding  $TPR_{20,10}$  parameters, with respect to the reference  $^{60}\text{Co}$  quality. Comparison is made considering data referring to collimated point-source and spectra, phase-space files and published data.

Regarding the point and the phase-space sources, black and red respectively in Fig. 2, the calculated values are associated with comparable statistical uncertainties that are in the order of 0.1%. In spite of a common trend,  $k_{Q,Q_0}$  values exhibit a systematic difference ranging between a minimum of about 0.19% corresponding to a  $TPR_{20,10}$  of about 0.74 and a maximum in the order of 0.5% for the highest energy.

The comparison with data in [2], blue circles in Fig. 2, reveals a very good agreement with the point-source  $k_{Q,Q_0}$  values, with differences within 0.2%. For the phase-space data a systematic underestimation of about 0.3%, with a maximum of about 0.8% at the highest energy, is observed in the whole considered quality range. A comparison with data in [3], triangles in Fig. 2, confirms agreement for the point-source data and underestimations for the phase-space data only at the highest beam energies. It is worth mentioning that the data in [3] represents a fit of combined Monte Carlo and experimental values referring to different accelerators from different research groups and also includes data of FFF beams. In [3] the level of agreement between individual values and fit is reported to be 0.5%. The same level of agreement exists with the point-source data

of this work, while differences are up to about 1% for the phase-space data. However, at lower  $TPR_{20,10}$  values the agreement with the Andreo et al' fit [3] is better for the phase-space data than for the point-source data, but this can be explained considering that in this region the contribution of FFF beam data tends to lower the  $k_{Q,Q_0}$  fitting curve.

A possible explanation for the underestimation associated to the phase-space files is linked to the files employed as radiation sources in the present work. Actually, density of particles in the phase space file represents a major issue to achieve a good statistics when small ionization chambers are simulated. Small active volumes, mean that only a small fraction of particles in the phase-space file is directly involved in the absorbed dose scoring. In addition, it must be considered also that by increasing the photon energy a cross section reduction is registered making the number of particles still more critical for the simulation results. In this condition the statistical uncertainty of the scored dose can be improved re-using the same particles to start simulations of new histories, but independence of results is not ensured. Then, depending on the particle density, the number of independent simulated histories could not be sufficient to optimally reproduce the properties of the radiation beam across the scoring volumes so producing less accurate results. Such a problem does not affect the simulations with the collimated point source. In this case the number of particles can be increased without restrictions, except for the simulation time, always producing independent simulated histories.

## 5 Conclusions

We have calculated the  $k_{Q,Q_0}$  quality correction factors for the micro ionization chamber PTW 31021 for radiotherapy photon beams in the  $TPR_{20,10}$  range from 0.66 to 0.80 with respect to the reference  $^{60}\text{Co}$  quality. These values have been determined both for ideal beams produced by a collimated point source and for phase-space files mimicking realistic clinical beams. We have compared them with a set of published values for the same beam qualities but referring to different accelerators. We have demonstrated the existence of a very good agreement with the point source results. The comparison with the phase-space file values shows a more scattered trend which can be addressed to the limited number of particles recorded in the phase-space files.

- [1] P. Andreo, D.T. Burns, K. Hohlfeld, M. Huq, T. Kanai, R. F. Laitano, V. Smyth, S. Vynckier. Absorbed dose determination in external beam cancer therapy; An International Code of Practice for Dosimetry Based on Standards of Absorbed Dose to Water. IAEA-TRS 398 (2000) (Vienna: IAEA).
- [2] J. Tikkanen, K. Zink, M. Pimpinella, P. Teles, J. Borbinha, J. Ojala, T. Siiskonen, C. Gomà and M. Pinto. Calculated beam quality correction factors for ionization chambers in MV photon beams. *Phys. Med. Biol.* **65**, 075003, (2020)
- [3] P Andreo 1 , D T Burns, R P Kapsch, M McEwen, S Vatnitsky, C E Andersen, F Ballester, J Borbinha, F Delaunay, P Francescon, M D Hanlon, L Mirzakhanian, B Muir, J Ojala, C P Oliver, M Pimpinella, M Pinto, L A de Prez, J Seuntjens, L Sommer, P Teles, J Tikkanen, J Vijande, K Zink . Determination of consensus  $k_Q$  values for megavoltage photon beams for the update of IAEA TRS-398, *Phys. Med. Biol.* **65**(9), 095011, (2020)
- [4] Maria Pimpinella, Luca Silvi, Massimo Pinto. Calculation of  $k_Q$  factors for Farmer-type ionization chambers following the recent recommendations on new key dosimetry data. *Phys. Med. EJMP* **57**, pp. 221-230, (2019)
- [5] I. Kawrakow, E. Mainegra-Hing, D.W.O. Rogers, F. Tessier, B.R.B. Walters. The EGSnrc Code System: Monte Carlo simulation of electron and photon transport. NRCC Report PIRS-701 (2018).
- [6] ICRU. Report 90. Key Data for Ionizing-Radiation Dosimetry: Measurement Standards and Applications. Journal of the International Commission on Radiation Units and Measurements 2016;Vol 14, Issue 1.
- [7] J. Wulff, K. Zink, I. Kawrakow. Efficiency improvements for ion chamber calculations in high energy photon beams. *Med. Phys.* **35**, pp 1328-36, (2008)
- [8] F. Iannone et al., "CRESCO ENEA HPC clusters: a working example of a multifabric GPFS Spectrum Scale layout," 2019 International Conference on High Performance Computing & Simulation (HPCS), Dublin, Ireland, 2019, pp. 1051-1052, doi: 10.1109/HPCS48598.2019.9188135.

# ABF SIMULATIONS FOR THE RECONSTRUCTION OF FREE ENERGY PROFILES DESCRIBING THE SWITCHING MECHANISM OF A pH-DEPENDENT DNA NANODEVICE

Federico Iacovelli<sup>1\*</sup>, Alice Romeo<sup>1</sup>, Mattia Falconi<sup>1\*</sup>

<sup>1</sup>*Department of Biology, Interuniversity Consortium, National Institute Biostructure and Biosystem (INBB), University of Rome Tor Vergata, Via della Ricerca Scientifica 1, 00133 Rome, Italy*

**ABSTRACT.** The pH-responsive behaviour of a triple helix DNA nanoswitch family sharing the same hydrogen bond forming domains and differing in the length of the linker connecting the double helical region to the triplex-forming region, has been characterized at the atomistic level through Adaptive Biasing Force Molecular Dynamics simulations. Reconstruction of the free energy profiles of triplex forming oligonucleotide unbinding from the double helix identifies different minimum energy paths for three nanoswitches, depending on the presence of 5, 15 or 25-bases in the connecting linker. The simulation data provide an atomistic explanation for previously published experimental results showing a two units increase in the  $pK_a$  switching mechanism decreasing the linker length from 25 to 5 bases, endorsing the validity of computational methods for the design and refinement of functional DNA nanodevices.

## 1 Introduction

The possibility to synthesize DNA sequences of any length, coupled to the specificity of Watson-Crick base pairing, has led to the current use of DNA as an efficient nanoscale building material. DNA has been widely used for building different three-dimensional structures, based on polyhedral geometries [1,2], as well as objects with complex shapes, mainly based on the DNA origami assembly technique [3]. DNA nanotechnology is now applied to solve real-world problems, developing functional and dynamical structures such as nanodevices or nanomachines. One important application is the biological sensing, where nanosensors have the potential to allow for a fast and straightforward detection of specific biological materials. DNA triplexes are an excellent example of sensing device for their pH-dependent conformational changes, which can be integrated into complex nanomachines [4]. In this context, the understanding of their atomistic behavior represents an important standpoint for a fine prediction of the rules governing their structure/function relationship. We already demonstrated the importance of Molecular Dynamics (MD) simulations in describing and predicting the atomistic behavior of DNA nanoswitches [5] integrated into complex nanostructures [6]. Recently, it has been experimentally shown that the pH-responsive behavior of a nucleic acid nanoswitch, which can form an intramolecular triplex structure through hydrogen bonds (Hoogsteen interactions) between a hairpin double helix (DH) and a single-strand triplex-forming oligo (TFO) having two protonation centers, strongly depends on the length of the linker connecting the two domains (Figure 1A) [7]. In this work, we apply adaptive biasing force molecular dynamics (ABF-MD) simulations to describe the effect of varying the length of the linker from 5 to 25 bases in regulating the pH-dependent conformational unbinding of the TFO from the DH of this diprotic nanoswitch. The results are in accordance with experimental evidences indicating that enhanced sampling techniques can be a valuable tool to rationally design the function of these devices.

\* Corresponding author. E-mail: [federico.iacovelli@uniroma2.it](mailto:federico.iacovelli@uniroma2.it); [falconi@uniroma2.it](mailto:falconi@uniroma2.it)

## 2 Simulation Methods

### 2.1 DNA Nanoswitch Modelling

The scheme of the three simulated diprotic, triplex-based nanoswitches is shown in Figure 1A. The systems share the same three-dimensional structure, apart the loop connecting the DH to the TFO, which can be composed by 5, 15 or 25 bases for the DIPRO5 (Figure 1B), DIPRO10 (Figure 1C) and DIPRO25 (Figure 1D) nanoswitches, respectively. The sequences used for the model building are those reported in the experimental work [7]:

**DIPRO5:** AAGAAAAGAA**TTTTATTCTTTCTTCTTTGTTCTTTCTT**

**DIPRO15:** AAGAAAAGAA**TTTTATTCTTTCTTCTTTGGTTGGTTGTTCTTTCTT**

**DIPRO25:** AAGAAAAGAA**TTTTATTCTTTCTTCTTTGGTTGGTTGGTTGGTTGGTTGTTCTTTCTT**

where the bases in bold represent the duplex forming regions (red and blue ribbons in Fig. 1B-D), in italic represent the triplex forming region (green ribbon in Fig. 1B-D) and the underlined bases represent the identical 5-bases loop and the variable loop (black and grey ribbons respectively in Fig. 1B-D). The coordinates of the three structures were generated as previously described [5].

### 2.2 MD and ABMD Simulations.

The topologies and the coordinate's files of the structures were generated as described in a previous work [5]. For each structure, a minimization run of 500 steps using the steepest descent followed by 1500 steps of conjugate gradient algorithm was performed to remove any unfavorable interaction. The systems were gradually heated from 0 to 300 K in the NVT ensemble over a period of 500 ps using the Langevin thermostat and with a restraint of  $0.5 \text{ kcal} \cdot \text{mol}^{-1} \cdot \text{\AA}^{-2}$  on each nucleotide to relax the solvent. Through 1.0 ns long equilibration runs the restraint forces were gradually decreased to  $0.1 \text{ kcal} \cdot \text{mol}^{-1} \cdot \text{\AA}^{-2}$ . The systems were simulated using an isobaric-isothermal (NPT) ensemble for 1.0 ns, fixing the temperature at 300 K and the pressure at 1.0 atm using the Langevin barostat. The SHAKE algorithm [8] was used to constrain covalent bonds involving hydrogen atoms. The systems were then subjected to a 10.0-ns equilibration run before starting the ABF-MD simulations, with a time step of 2.0 fs, using the PME method [9] for long-range interactions and a cut-off of 10.0 Å for the short-range interactions. The unbinding of the triplex-forming region equilibrated systems was simulated by means of 50.0-ns long ABF-MD simulations using the NAMD 2.13 MD engine [10]. ABF-MD enhances the sampling of high-energy regions of the FES by adding to the total energy of the system a biasing history dependent term, obtained by the sum of Gaussian hills laying on the subspace described by a set of user-defined collective variables (CVs). The FES can be reconstructed at the end of the simulation as the sum of the added Gaussian hills. In the simulation, two CVs were used to describe the triplex-forming oligo unbinding from the duplex region: 1) the distance between the center of mass of the C2' atoms of the nucleotides belonging to the duplex and the C2' atoms belonging to the triplex forming region; 2) the number of hydrogen bonds between the triplex forming region and the nucleotides belonging to the duplex using the AMBER implemented CV coordination number (CN):

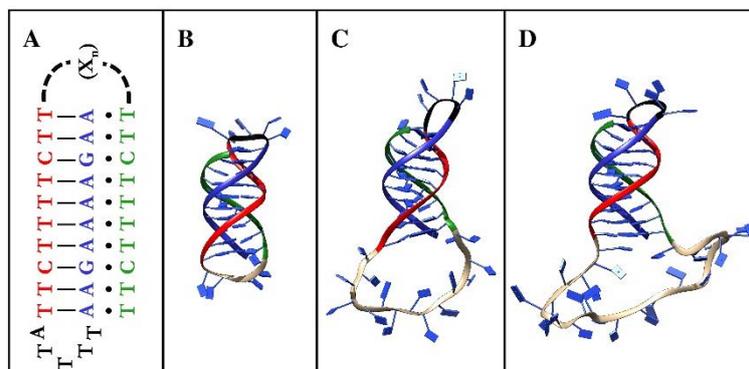
$$CN = \sum_{ij} \frac{1 - \left(\frac{r_{ij}}{r_0}\right)^6}{1 - \left(\frac{r_{ij}}{r_0}\right)^{12}}$$

where  $r_i$  and  $r_j$  are the coordinates of the atoms involved in the hydrogen bond interactions. The value of the parameter  $r_0$  was set to 3.2 Å, calculated in the 10.0 ns MD equilibration.

### 2.3 Trajectory analyses.

Principal component analysis (PCA) was performed over the entire 50-ns ABF-MD trajectories using the GROMACS 2019.1 analysis tools [11]. The *gmx cluster* module of GROMACS was used to perform a clustering analysis using the *gromos* algorithm on all the saved configurations. Contour plots

in Figures 2 were generated with *Matplotlib*, plots in Figures 3 was made with R and *ggplot2*, while the structures represented in Figures 1 and 2 have been drawn with UCSF Chimera [12]. The analyses and the simulations were performed by using 2 node, for a total of 96 CPUs, of the ENEA CRESCO6 HPC infrastructure [13].



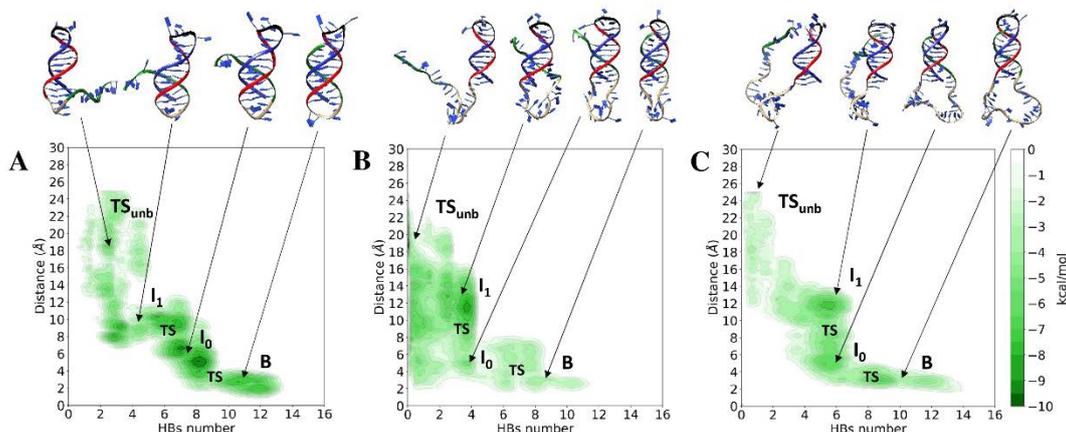
**Figure 1.** Schematic (A) and cartoon (B-D) representations of the simulated DNA nanoswitches. The red and blue colors indicate the two strands forming the double helix region, while the black indicates the conserved five bases loop and the green the TFO. The black dashed line in A and the tan colored ribbon in B-D represent the 5 (B), 15 (C) and 25 (D) bases variable loop. The pictures were produced using the UCSF Chimera 1.12 program [12].

## 3 Results

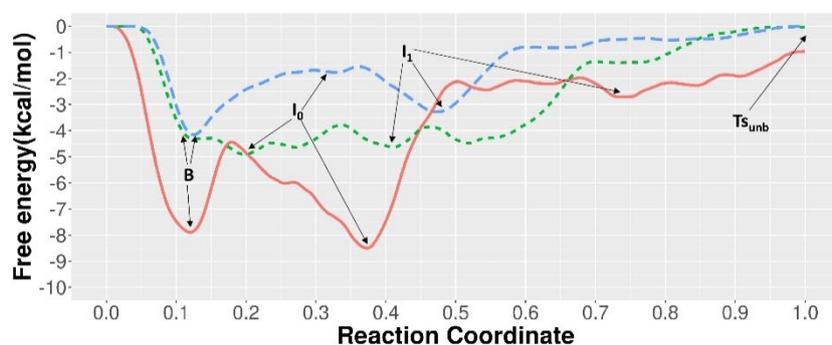
### 3.1 Unbinding of the TFO from the double helix.

The TFO unbinding profile from the double helix was analyzed in terms of free energy surface (FES) plotted as a function of the number of hydrogen bonds (HB) established between the TFO and the DH and of the distance between their centers of mass (Fig. 2). For all the three systems, the surface shows the presence of three minima corresponding to the bound (B) and two intermediate ( $I_0$  and  $I_1$ ) states separated by two transition states (TS), with  $T_{S_{unb}}$  representing the final step of the ABMD simulations (Fig. 2A-C). The coordinates corresponding to the conformations belonging to the states identified in the FES were clustered and the representative structures identifying the B, I and  $T_{S_{unb}}$  states are displayed in the upper part of Fig. 2A-C. For the DIPRO5 switch, the B state is stable, being characterized by the presence of 13-14 HBs and a distance between TFO and DH centers of mass close to 2.0 Å (Fig. 2A). The  $I_0$  state has a lower number of HBs and an average TFO-DH distance of 6 Å (Fig. 2A), due to the unbinding of three nucleotides at the 3' end of the TFO. In the  $I_1$  state, the number of HB interactions is reduced to 4 and in the final  $T_{S_{unb}}$  state just 2-3 HBs are present. In fact, in the  $T_{S_{unb}}$  state the TFO is not fully detached since the first three bases at the TFO 5' end are still interacting with the DH. The DIPRO15 system shows the presence of minima less stable than those observed in the DIPRO5 one (Fig. 2B). In detail, the B state shows the presence of 8-10 HBs, although the TFO is still quite close to the DH. The  $I_0$  state is similar to that of the DIPRO5 switch, showing the unbinding of the first three nucleotides at the TFO 3' end. The  $I_1$  state samples several loose structures showing a concurrent loss of three base pair interactions at both the 3' and 5' ends of the TFO (Fig. 2B). Starting from this configuration, the  $T_{S_{unb}}$  state is reached by an extensive sampling of a high-energy conformational basin, characterized by a TFO-DH distance of about 20 Å and the loss of all HBs (Fig. 2B), indicative of a full unbinding of the TFO from the DH. The DIPRO25 switch is the less stable one, being characterized by high-energy minima, with the B state easily evolving towards the  $I_0$  state (Fig. 2C). The  $T_{S_{unb}}$  state, identified by structures having a TFO-DH distance of 24-25 Å and

characterized by a complete loss of HBs interactions, is reached through the sampling of a high-energy conformational basin (Fig. 2C). A comparative free energy profile, along a minimum energy path, is plotted in Fig. 3 as a function of a generalized reaction coordinate. The energy of the B state for the DIPRO5 switch (red line), is two times lower than those observed for the DIPRO15 and DIPRO25 systems (green and blue lines, respectively). In line, the DIPRO5 B and  $I_1$  states are separated by a relatively high activation energy (Fig. 4, red line), again almost twice than that required for the TFO unbinding in the DIPRO15 and DIPRO25 systems (Fig. 3, green and blue lines, respectively).



**Figure 2.** Iso-energetic contour of the free energy surface plotted as function of the HBs number established between the TFO and the DH and of the distance between their centers of mass for the DIPRO5 (A), DIPRO15 (B) and DIPRO25 (C) system. The upper structures are the representative configurations of the most populated cluster of the Bound, Intermediate and Transition states extracted from the trajectories. The colour scheme follows that described in Figure 1.



**Figure 3.** Free-energy profile of TFO unbinding from the DH, calculated along a minimum energy path, as a function of a generalized ‘Reaction Coordinate’ for the DIPRO5 (red line), DIPRO15 (green line) and DIPRO25 (blue line) switches.

Overall, these data indicate that the length of the linker connecting DH with TFO has a crucial role in stabilizing the interaction of the TFO to the DH. In detail the interaction energy decreases upon increasing the loop length indicating that modulation of the entropic contribution related to the loop length has a direct effect on the TFO unfolding energy. This work demonstrates the importance of computational approaches, such as enhanced sampling MD simulations, for the design and tuning the functional properties of DNA nanoswitches.

## Acknowledgment

The computing resources and the related technical support used for this work have been provided by CRESCO/ENEAGRID High Performance Computing infrastructure and its staff [13]. CRESCO/ENEAGRID High Performance Computing infrastructure is funded by ENEA, the Italian National Agency for New Technologies, Energy and Sustainable Economic Development and by Italian and European research programmes, see <http://www.cresco.enea.it/english> for information.

## References

- [1] Alves C, Iacovelli F, Falconi M, Cardamone F, Morozzo della Rocca B, de Oliveira CLP, et al. A Simple and Fast Semiautomatic Procedure for the Atomistic Modeling of Complex DNA Polyhedra. *J Chem Inf Model*. 2016 May 23;**56**(5):941–9.
- [2] Zimmermann J, Cebulla MPJ, Mönninghoff S, von Kiedrowski G. Self-Assembly of a DNA Dodecahedron from 20 Trisoligonucleotides with C3h Linkers. *Angew Chemie Int Ed*. 2008 Apr 28;**47**(19):3626–30.
- [3] Fan S, Wang D, Kanaan A, Cheng J, Cui D, Song J. Create Nanoscale Patterns with DNA Origami. *Small*. 2019 Jun 24;**15**(26):1805554.
- [4] Hu Y, Ceconello A, Idili A, Ricci F, Willner I. Triplex DNA Nanostructures: From Basic Properties to Applications. *Angew Chemie Int Ed*. 2017 Nov 27;**56**(48):15210–33.
- [5] Iacovelli F, Idili A, Benincasa A, Mariottini D, Ottaviani A, Falconi M, et al. Simulative and Experimental Characterization of a pH-Dependent Clamp-like DNA Triple-Helix Nanoswitch. *J Am Chem Soc*. 2017 Apr 19;**139**(15):5321–9.
- [6] Ottaviani A, Iacovelli F, Idili A, Falconi M, Ricci F, Desideri A. Engineering a responsive DNA triple helix into an octahedral DNA nanostructure for a reversible opening/closing switching mechanism: a computational and experimental integrated study. *Nucleic Acids Res*. 2018 Nov 2;**46**(19):9951–9.
- [7] Mariottini D, Idili A, Nijenhuis MAD, Ercolani G, Ricci F. Entropy-Based Rational Modulation of the p K a of a Synthetic pH-Dependent Nanoswitch. *J Am Chem Soc*. 2019 Jul 24;**141**(29):11367–71.
- [8] Ryckaert JP, Ciccotti G, Berendsen HJC. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J Comput Phys* [Internet]. 1977;**23**(3):327–41.
- [9] Darden T, York D, Pedersen L. Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *J Chem Phys*. 1993;**98**(12):10089.
- [10] Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, et al. Scalable molecular dynamics with NAMD. *J Comput Chem* [Internet]. 2005;**26**(16):1781–802.
- [11] Abraham MJ, Murtola T, Schulz R, Páll S, Smith JC, Hess B, et al. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*. 2015 Sep;**1–2**:19–25.
- [12] Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, et al. UCSF Chimera—A Visualization System for Exploratory Research and Analysis. *J Comput Chem*. 2004;**25**:1605–12.
- [13] Iannone F, Ambrosino F, Bracco G, De Rosa M, Funel A, Guarnieri G, et al. CRESCO ENEA HPC clusters: a working example of a multifabric GPFS Spectrum Scale layout. In: 2019 International Conference on High Performance Computing & Simulation (HPCS). IEEE; 2019. p. 1051–2.

# THE MOLECULAR WAY TO QUANTUM BITS AND SPINTRONICS: AN *AB INITIO* STUDY

Federico Totti<sup>1\*</sup>, Francesco Buonocore<sup>2</sup>, Massimo Celino<sup>2</sup>, Matteo Briganti<sup>3</sup>, Andrea Albino<sup>1</sup>

<sup>1</sup>*Università degli Studi di Firenze, Chemistry Department “U. Schiff”, 50019, Sesto F.no, Italy*

<sup>2</sup>*ENEA, Casaccia Research Centre, 00123 Rome, Italy.*

<sup>3</sup>*Federal University of Parana, Department of Chemistry, 81530-900 Curitiba, Parana, Brazil*

**ABSTRACT.** The challenge of developing and implementing quantum devices is increasingly topical. Indeed, quantum computing, spintronics, and ultra-high density storages are experiencing a real boost since they appeared both as immature implementations or as next-future promising technological applications. In such a framework, the “molecular way” is trying get a place in the sun in this landscape as a more versatile alternative to the “solid state” one. In this context the season of hunting for the most promising magnetic molecules has begun. Here we present computational studies on several systems as potential molecular QuBits, permanent nano-magnets, and spintronics devices.

## 1 Introduction

The implementation of magnetic molecules as storage and logical devices would represent a real revolution in the field of information technology, due to their properties at the interface between classic and quantum realms. Several applications have been proposed for open-shell metal complexes: molecular units for memory storage at the sub-nanometric levels,[1] spin filters for the molecular control of spin currents[2] and, very recently, molecular quantum bits (QuBits) for quantum computation [3]. Specifically, this last achievement would allow to exploit the rules of quantum mechanics, such as superposition and entanglement, to process information to an increased speed and quantity per unit of time with respect to classical logic. The class of molecules mainly used in these regards are coordination complexes, i.e. systems containing open shells transition metal ions and/or lanthanide ones coordinated by organic ligands.

Among the several appointed candidates to build a real quantum computer (for example nitrogen vacancies in diamond, ionic traps, photons, semiconductors) from the storage to the logical processor, molecular compounds have a great advantage: their structural versatility. Synthetic chemistry can here play a fundamental role in modulating specific interactions between microscopic constituents in order to achieve the desired molecular properties. A significant part of such synthetic efforts have been driven by theoretical results and they provided outstanding results in the last few years, by designing molecules behaving as magnets, i.e. presenting quantum properties, above the liquid nitrogen temperatures.

For example, the first main challenge to face for the use magnetic molecules as QuBits is the preservation of the magnetic state for an enough amount of time (*i.e.* high relaxation time) to realize logical operations. This can be achieved only by a careful design of the coordination geometry and engineering of the local and lattice vibrational modes.

At the same time, the safe adsorption and organization of single molecule magnets or other magnetic molecules as spin-crossover systems on different surfaces and 2D materials is the other main bottleneck for the realization of end-user devices. This goal is necessary to address and directly control the single molecule by external stimuli. Even if a wide variety of experimental techniques is available in surface science, it is common that the explanation of crucial local properties remains elusive. Even in such framework, the computational support resulted decisive for a detailed description at the microscopic level. However, due to the large dimensions of the systems to be handled, a relatively limited number of works are appeared when single molecule magnets are used. [4] In this context, the present work introduces a first principles investigation of different promising systems for the advancement in the area of spintronic molecular devices: molecular crystals, surfaces and hybrid materials (molecules grafted on surface).

**1.1 Hydrogenated graphene deposition on Au(111).** Graphene is a honeycomb lattice constituted by a monolayer of carbon atoms that shows fascinating properties, such as its conducting electronic ground state or mechanical properties. The nature of this material settles today an increasing interest in exploration of the plethora of its technologic and scientific applications.

The purpose of this analysis winks at spintronics devices as magnetic molecules chemisorbed or physisorbed on conducting surfaces or electrodes. [4] The surface-molecule interaction in many cases lead to the loss of magnetic features of the molecule and hence an extensive work of characterization of this interaction as a function of the molecule and the substrates [5] is needed.

The resilient C-C bonds allow chemical modification of graphene without breaking the structure, offering the possibility to synthesize different substrates from the same starting material. The effect of hydrogenation of graphene and deposition on Au(111) (see Fig. 1) on the electronic structure is hereafter depicted.

**1.2 Magnetic molecules as potential molecular Qubits.** Among the open challenges, the possibility to retain the magnetization of molecular spins for a time long enough to apply a series of gates represents a fundamental one. A wide variety of experimental investigations is available in this field, [6] on the other hand the theoretical description is still at an early stage. [7-8] The present work introduces a comparative first principles investigation in  $S=1/2$  molecular complexes based on vanadium-(IV) and copper-(II), [VO(acac)<sub>2</sub>] (acac = acetylacetonate), [VO(TPP)] (TPP = tetraphenylporphyrin), [Cu(Pc)] (Pc = Phtalocyanine), [Cu(TTDPz)] (Fig. 2) (TTDPz = tetrakis-thiadiazoleporphyrazine).

The presented systems were chosen as suitable candidates for deposition on surface: neutral charge, evaporability, thermal stability are the main features required to the thin layer goal.

**1.3 Single Molecule Magnets grafted on Au(111).** Single molecule magnets are molecules that under a certain temperature, called the blocking temperature, show slow relaxation of the magnetization and eventually the opening of a hysteresis loop. Such a behaviour arises from the electronic structure of the single molecule itself and not from long range interactions in space as in classic bulk magnets. Such bistability made this class of compound very appealing candidates in the area of quantum information storage as magnetic memories [1,4].

One of the open questions in the studies of adsorbed magnetic molecules is the control and the tuning of the exchange coupling strength between magnetic units and the role played by the surface, even if non-magnetic. Indeed, the spin-injection in the surface can generate interesting phenomena by generating bound states which are spatially inhomogeneous and, consequently, can differently interact with other magnetic units [3]. In such a framework, the simulation by periodic DFT of an adsorbed monolayer of a functionalized Fe<sub>4</sub> molecule, a milestone in the area of SMM grafted on surface [1,4,5], will provide essential insights about the eventual presence and strength of magnetic coupling between molecules to understand and how to tune and control them.

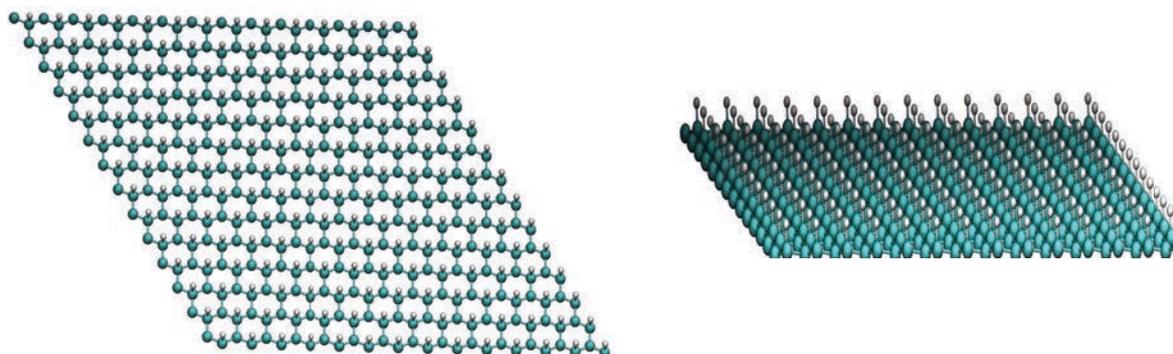
**1.4 Surface effects on a photochromic spin-crossover iron(II) molecular switch grafted on HOPG.** Spin-Crossover (SCO) metal complexes are among the most attractive systems as building blocks for spintronics, data storage, and sensing devices. These species may commonly exist in two electronic states with different magnetic, optical, and structural properties, and can be reversibly switched by various external stimuli (pressure, temperature, light-irradiation). Their surface organization by sublimation processes onto different substrates, with controlled thickness from micrometres down to sub-monolayer coverage, is crucial for forthcoming technological and spintronics applications, becoming a highly attractive and flourishing research field during recent years. A sub monolayer deposit of  $[\text{Fe}^{\text{II}}(\text{H}_2\text{B}(\text{pz})_2)_2(\text{phen}^*)]$ , (where pz is 1-pyrazolyl and phen\* is a diarylethene functionalized phenanthroline ligand, see Fig. 4) has been sublimated in ultra-high vacuum (UHV) on HOPG substrates and characterized by a multi-technique protocol. The photochromic ligand allows a ligand-driven light-induced spin change (LD-LISC) effect: reversible photocyclization of the photoactive diarylethene-based ligand which remotely triggers a high-spin to low-spin crossover at the coordinated iron(II) ion. Accordingly to earlier reports, this photocyclization can be observed only when the ligand maintains a specific photoactive conformation with the two aryl rings in an antiparallel orientation. Experimental findings showed, for the first time, a surface-induced inhibition of the LD-LISC mechanism. Periodic-DFT/GPW calculations has provided a sound rationalization of such a result.

## 2 Methods

Modeling molecules in a more realistic environment in an *ab initio* computational framework has become possible only in the past few years, enabled by an extensive work of the integration of density functional theory (DFT) in both gas-phase or periodic boundary conditions (PBC), post-Hartree-Fock (post-HF) methods, and spin dynamics for calculation of the dynamical magnetic properties of multispin systems. Using DFT with generalized gradient approximation for exchange and correlation energy, we analyse the details of the electronic properties as a function of chemical structure and inter-system interactions. CP2K [9] software has been employed for periodic DFT calculations, while post-HF have been performed with ORCA 4.2.1 [10] quantum chemistry package. All the calculations have been performed on CRESCO6 cluster.

## 3 Results

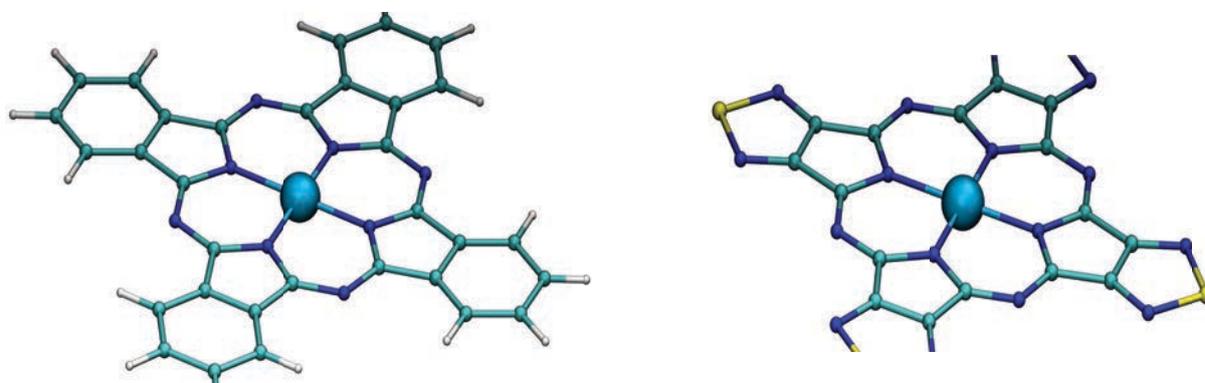
### 3.1 Hydrogenated graphene deposition on Au(111)



**Fig.1:** 196 unit cells of graphene, top (left) and side (right) views.

The electronic structure is changed by partial hydrogenation. The carbon atoms in graphene break their  $\pi$ -bonding network and the p-electrons associated with the non-hydrogenated carbon atoms are localized and unpaired. [11,12] These are interesting substrates to study their effect when depositing magnetic molecules such as  $\text{Tb}(\text{Pc})_2$  (Pc = Phtalocyanine). Graphone (see Fig. 1), indeed, becomes a semiconductor with a small band gap, showing substantial differences with graphene (with zero band gap) and graphane (with large band gap). When half of the carbon atoms are hydrogenated, strong  $\sigma$ -bonds are formed between C and H atoms and the  $\pi$ -bonding network is broken. In order to study the preferred coupling of the magnetic moments inside the material, we considered the following two magnetic configurations: ferromagnetic (FM) coupling; antiferromagnetic (AF) coupling. Our calculations are based on a supercell that consists of 196 unit cells, this choice is mandatory to study the effect of deposition of large molecules on the surface. *The typical CP2K run employed 8 nodes (384 cores).*

### 3.2 Magnetic molecules characterization



**Fig.2:** Left:  $\text{Cu}(\text{Pc})$  molecule; right:  $\text{Cu}(\text{TTDPz})$  molecule.

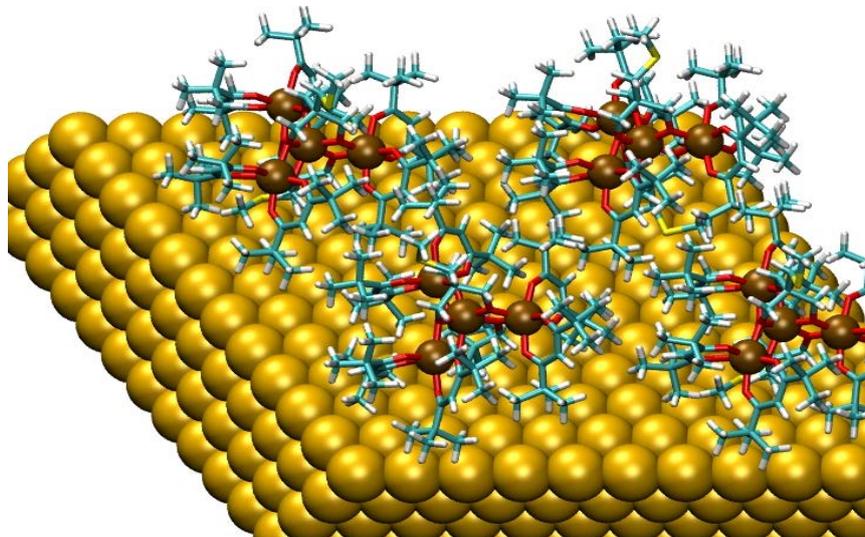
Our approach aims at addressing the spin-lattice relaxation time  $T_1$  that is the primary limiting parameter for the development of higher working temperature molecular Qubits. It consists of modeling the magnetic properties by first principles, when perturbing the molecular structures along the normal modes of vibrations, following a strategy adopted in previous works. [7,8]

The development of a rationale for the chemical design of new molecular Qubits must proceed through the understanding of the correlations between the spin-phonon dynamics and the chemical identity of the molecular units. Our first-principles study made it possible to disentangle different contributions to the spin-phonon coupling and connect them to chemical features. The results show that the vibrational modes are coupled to spin depending on several features. The main effects are due to the chemical nature of the first coordination shell and are modulated by the lattice environment. The spin-orbit coupling, peculiar of each element, the coordination geometry of metal atom, the symmetry group of the molecule and the space group of crystal are the main fingerprints to deal with to optimize the spin-lattice relaxation. [13] *The typical CP2K run employed 8 nodes (384 cores), while the typical ORCA only 1 node (48 cores).*

### 3.3 Single Molecule Magnets grafted on Au(111).

The  $\text{Fe}_4$  full monolayer adsorbed on gold have been modelled by placing four molecules on an hexagonal cell of Au(111) made of four slabs of gold atoms (see Fig. 3). The geometry has been fully

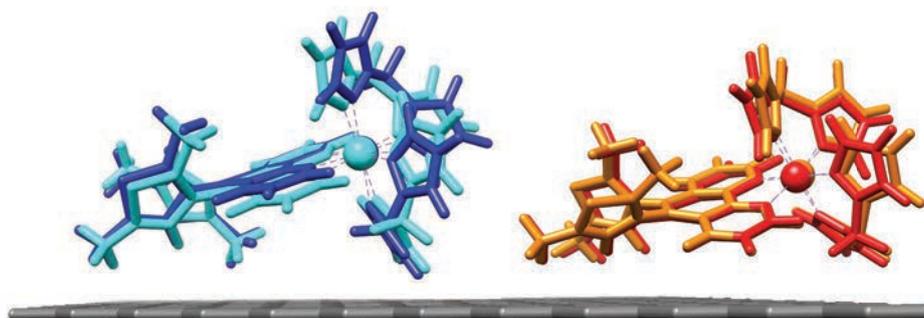
relaxed. DFT Broken Symmetry calculations have been performed on the final optimized structure [4]. The results provided the magnetic exchange coupling strength between the four Fe<sub>4</sub> molecules.



**Fig.3:** Four Fe<sub>4</sub> molecules on an hexagonal cell of Au(111) .

The obtained spin topology is complex, and it shows both antiferro- and ferro-magnetic interactions between molecules. Such interactions have been proven to happen through-surface super-exchange interactions due to the overlap between Fe<sub>4</sub> magnetic orbitals with valence surface bands. A more detailed analysis shows that the strength and the sign of the magnetic interaction depends from both the angle between the C<sub>3</sub> axis of the molecule and the normal to the surface along with the reciprocal orientation among Fe<sub>4</sub> complexes. *The typical CP2K run employed 16 nodes (768 cores).*

### 3.4 Surface effects on a photochromic spin-crossover iron(II) molecular switch grafted on HOPG



**Fig.4:** Superimposed computed geometries for low spin [Fe<sup>II</sup>(H<sub>2</sub>B(pz)<sub>2</sub>)<sub>2</sub>(phen\*)]-anti (blue), High spin [Fe<sup>II</sup>(H<sub>2</sub>B(pz)<sub>2</sub>)<sub>2</sub>(phen\*)]-anti (cyan), low spin [Fe<sup>II</sup>(H<sub>2</sub>B(pz)<sub>2</sub>)<sub>2</sub>(phen\*)]-par (red), and high spin [Fe<sup>II</sup>(H<sub>2</sub>B(pz)<sub>2</sub>)<sub>2</sub>(phen\*)]-par (orange). Only the topmost HOPG layer was left for the sake of clarity.

The HOPG (001) surface was “standalone” optimized as a four-layer slabs, each layer consisting of 288 carbon atoms, according to the graphite ABAB crystalline structure. The size of the hexagonal simulation cell was set to (29.568 × 29.568 × 60.000) Å<sup>3</sup>. Periodic boundary conditions were applied in the three directions, taking care of use a high z value to avoid spurious interactions between replicas. During the optimizations, the HOPG bottom layer was kept fixed to bulk positions while the top-

most ones were left free to relax in order to reproduce their surface-like behaviour. The pDFT/GPW calculations showed that expected room temperature photoconversion via the LD-LISC effect is not occurring in the case of submonolayer deposits because of a selective surface-driven destabilization (~11 kcal/mol) of the photo-active antiparallel conformer of the molecular switch leading to the loss of photoactivity of the material. The soundness of such result was corroborated by the nice agreement of computed density of states (DOS) for each conformer with the experimental ones.[14] *The typical CP2K run employed 8 nodes (384 cores).*

### 3 Conclusions

We have presented on going and published computational studies on different systems that present potential nano-technological applications as potential molecular QuBits, permanent nano-magnets, and spintronics devices. In this regards, we have exploited the state-of-art computational protocols to accurately describe highly correlated electronic structures and the spin-phonon interactions in potential Qubits and in SMMs. The developed protocols demonstrated to be fine enough to grasp the surface effects on the selective choice of conformational photo-active functional group of spin-crossover systems and to evidence an unprecedented mediated super-exchange interactions among Fe<sub>4</sub> SMMs by gold surface. Moreover, it would be also possible to verify that Au(111) surface can only slightly affect the electronic properties of graphene and its hydrogenated derivatives making them suitable to be exploited as potential hybrid surfaces where Qubits or single ion magnets might be adsorbed.

### 4 Acknowledgements

The computing resources and the related technical support used for this work have been provided by CRESCO/ENEAGRID High Performance Computing infrastructure[15] and its staff.

### References

- [1] M. Mannini, F. Pineider, C. Danieli, F. Totti, L. Sorace, P. Sainctavit, M. Arrio, E. Otero, L. Joly, J. Cezar, A. Cornia, R. Sessoli. Quantum tunnelling of the magnetization in a monolayer of oriented single-molecule magnets. *Nature* **468**, 7322, 417-421, (2010).
- [2] L. Bogani, W. Wernsdorfer. Molecular spintronics using single molecule magnets. *Nature Materials* **7**, 1, 179-186, (2008).
- [3] R. Sessoli. Toward the Quantum Computer: Magnetic Molecules Back in the Race. *ACS Central Science* **1**, 9, pp. 473-474, (2015).
- [4] A. Lunghi, M. Iannuzzi, R. Sessoli, F. Totti. Single molecule magnets grafted on gold: magnetic properties from ab initio molecular dynamics. *Journal of Materials Chemistry C* **3**, 28, pp. 7294-7304, (2015).
- [5] A. Caneschi, D. Gatteschi, F. Totti. Molecular magnets and surfaces: A promising marriage. A DFT insight. *Coordination Chemistry Reviews* **289-290**, 1, pp. 357-378, (2015).
- [6] M. Atzori, A. Chiesa, E. Morra, M. Chiesa, L. Sorace, S. Carretta, R. Sessoli. A two-qubit molecular architecture for electron-mediated nuclear quantum simulation. *Chemical Science*, **9**, 29, 6183-6192, (2018).
- [7] A. Lunghi, F. Totti, S. Sanvito, and R. Sessoli. Intra-molecular origin of the spin-phonon coupling in slow-relaxing molecular magnets. *Chemical Science*, **8**, 9, 6051-6059, (2017).
- [8] A. Albino, S. Benci, L. Tesi, M. Atzori, R. Torre, S. Sanvito, R. Sessoli, A. Lunghi. First-Principles Investigation of Spin-Phonon Coupling in Vanadium-Based Molecular Spin Quantum Bits. *Inorganic Chemistry*, **58**, 15, 10260-10268, (2019).
- [9] F. Neese. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* The ORCA Program System. **2**, 1, 73-78, (2012)

- [10] J. Hutter, M. Iannuzzi, F. Schiffmann, J. VandeVondele. Cp2k: Atomistic Simulations of Condensed Matter Systems. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **4**, 1, 15-25, (2014).
- [11] J. Zhou, Q. Wang, Q. Sun, X. S. Chen, Y. Kawazoe, P. Jena. Ferromagnetism in semihydrogenated graphene sheet. *Nano Letters*, **9**, 11, 3867-3870, (2009).
- [12] F. Buonocore, A. Mosca Conte, N. Lisi. Effects of the substrate on graphone magnetism: A density functional theory study, *Physica E: Low-Dimensional Systems and Nanostructures*, **78**, 65-72, (2016).
- [13] F. Santanni, A. Albino, M. Atzori, D. Ranieri, E. Salvadori, A. Lunghi, M. Chiesa, F. Totti, L. Sorace, A. Bencini and R. Sessoli. Nuclear-Spin-Economy and vibrational analysis in Macrocyclic-based Molecular Spin Qubits. *Inorganic Chemistry*, submitted
- [14] L. Poggini, G. Londi, M. Milek, A. Naim, V. Lanzilotto, B. Cortigiani, F. Bondino, E. Magnano, E. Otero, P. Sainctavit, M.-A. Arrio, A. Juhin, M. Marchivie, Marat M. Khusniyarov, F. Totti, Patrick Rosa, M. Mannini Sub-monolayer deposit of a photochromic spin-crossover iron(II) molecular switch on highly oriented pyrolytic graphite, *Nanoscale* **11**, 20006-20014 (2019).
- [15] F. Iannone *et al.*, "CRESCO ENEA HPC clusters: a working example of a multifabric GPFS Spectrum Scale layout," *International Conference on High Performance Computing & Simulation (HPCS)*, Dublin, Ireland, 2019, pp. 1051-1052, doi: 10.1109/HPCS48598.2019.9188135 (2019)

# MONTE CARLO SIMULATIONS SUPPORTING STUDIES OF PWR'S AND OF EXPERIMENTAL n\_TOF FACILITY AT CERN

Patrizio Console Camprini<sup>1\*</sup>

<sup>1</sup>*ENEA, Fusion and Technology for Nuclear Safety and Security Department, 40129, Bologna, Italy*

**ABSTRACT.** In the framework of the collaborations between ENEA and IRSN (French Institute for Radioprotection, Nuclear Safety and Security) studies are carried out concerning nuclear safety evaluations of PWR cores and some irradiation aspects within the reactor pit. Monte Carlo simulations are performed by means of the MCNP6.1 code for both eigenvalue calculations and fixed source problems. In particular, Gen-II reactor - like the Tihange model - and innovative Gen-III PWRs are evaluated as benchmarks for deep penetration problems. In addition, ex-core responses are used to employ and test variance reduction techniques. Within the collaborations with the INFN (Italian Institute for Nuclear Physics) Monte Carlo simulations are conducted to analyse the outcome of an experimental campaign performed at the n\_TOF facility at CERN (European Center for Nuclear Research).

## 1 Reactor core analysis for Gen-II and Gen-III PWRs

The present research activity concerned Monte Carlo simulations of Gen-II and Gen-III nuclear power plants and followed what has been carried out in the previous years. In fact, Monte Carlo MCNP6.1[1] code has been used to model and analyse benchmarks of a Gen-II system like the nuclear reactor at Tihange [2] and a reference Gen-III nuclear reactor core. The simulation regarding Gen-II nuclear core was about the determination of the detector response placed at 2 different positions outside the core: the first was at 45° angle and the second was at 315° (Fig. 1). Both were boron detectors and placed at about 250 cm from the axis of the core, 80 cm outside the core cylindrical region. Detector material was <sup>10</sup>B dispersed in natural boron, and the considered nuclear reaction was alpha particle production induced by a neutron absorption as (n,α).

Conversely, a Gen-III innovative nuclear system has been studied through a reference benchmark [3]. Deep penetration problems considered here were analysed considering the impact of the prompt fission spectrum according to the specific nuclide that underwent fission. Some minor and major actinides are considered in preliminary studies: <sup>235</sup>U, <sup>238</sup>U, <sup>239</sup>Pu, <sup>240</sup>Pu as well as neptunium and americium. Firstly, an eigenvalue problem has been run, and fission sites were sampled: at the same time an average fission spectrum was sampled. In addition, some important nuclides have been accounted for at the time of fission, to enrich statistics and knowledge of the source. A fixed source was prepared in this way and it was then used to characterize flux and reaction rates on an assembly-by-assembly basis and concerning ex-core detectors.

As explained, the approach to the calculation followed different models of the neutron source inside the core. A fixed source has been firstly defined by sampling the fundamental mode, treating a previous eigenvalue problem, and scoring neutron production on a particular axial binning on a pin-

by-pin pattern. The ex-core responses were then studied from this fundamental mode in a fixed source problem, aiming at all the specific tallies requested.

In the fixed source calculation, a volumetric source has been implemented with an isotropic angular distribution.

Different configurations of the fixed source were sampled to study the impact on responses:

- homogeneous source throughout the core
- source with an axial distribution assembly by assembly
- source with a pin-by-pin resolution and a subdivision of the pin in 2 cylindrical shells

In addition, an eigenvalue direct problem was solved together with the responses of interest.

Fixed source simulations allowed the implementation of the DSA variance reduction technique – developed at ENEA for fixed source simulations and thus extended to eigenvalue problems [4].

The DSA method produces variance reduction parameters through a minimization of the quality function – namely the maximization of the figure of merit. In fact, the adaptive multi-step iterative algorithm starts from initial sampling of several tracks, and thus increases knowledge of the system acting on different contribution to tally in phase space. Maximization of the figure of merit for a single or a multiple response produces the variance reduction parameters that can be expressed in a weight window form to easily utilise the MCNP calculation facility.

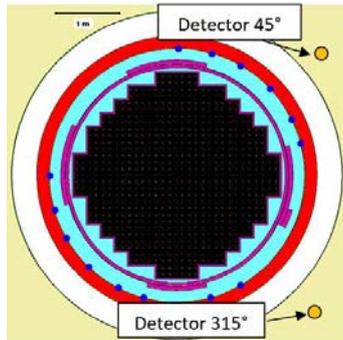


Fig.1: Gen-II model core cross-section

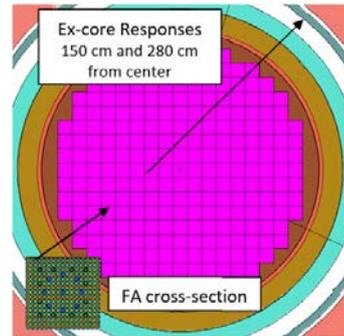


Fig.2: Gen-III PWR core cross-section (Fuel assembly – FA – cross-section)

## 2 Simulation of Experiment for the n\_TOF Facility at CERN

Within the framework of nuclear physics and specifically basic research, n\_TOF facility at CERN (European Center for Nuclear Research) is an important facility utilized to study nuclear reaction cross sections at different energies, thanks to a particularly powerful neutron beam extracted from the large accelerator current, through spallation reactions on a dedicated target.

The present study has been carried out in collaboration with Italian National Institute for Nuclear Physics (INFN), with the Bologna Section.

Monte Carlo simulations have been oriented to a neutron beam impinging on a uranium target since  $^{235}\text{U}(n, f)$  cross section is one of the most important cross section standards at thermal neutron energy between 0.15 MeV and 200 MeV, for measurements of neutron-induced reaction cross-sections [5]. It is used in a wide range of applications. In addition, as in most other neutron facilities, the  $^{235}\text{U}(n, f)$  cross section is used as a reference for fission cross section measurements (see [6] and references therein).

Despite its widespread use, however, the recommended  $^{235}\text{U}(n,f)$  cross section data at energies above 20 MeV are based on a small set of measurements [7,8]. Hence, there is a clear and long-standing demand from the International Atomic Energy Agency (IAEA) to improve this situation.

At energies above 200 MeV, the  $^{235}\text{U}(n,f)$  cross section plays an important role for several applications, as well as for fundamental nuclear physics.

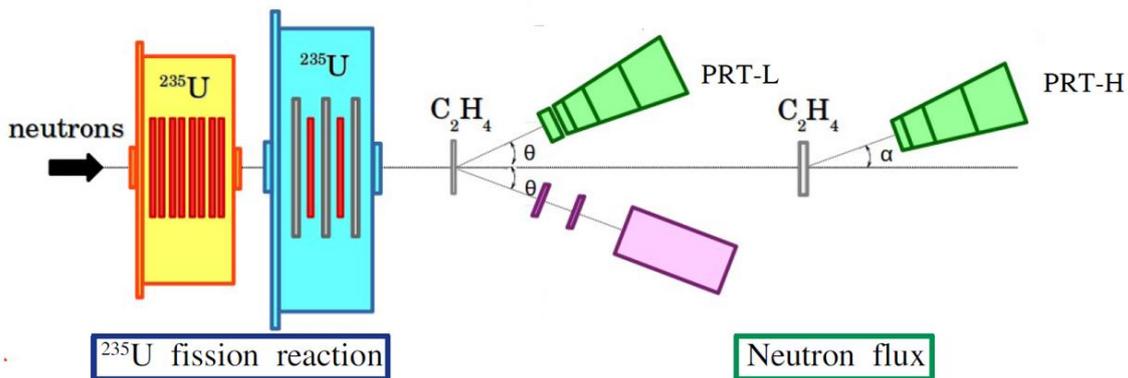
Despite the importance of the high-energy region, at present no data exist on neutron induced fission above 200 MeV, and it is necessary to rely on theoretical estimates.

However, new theoretical calculations [9] have indicated that the  $^{235}\text{U}(n,f)$  cross section at high energy may be substantially different from what was previously thought. Furthermore, in a new evaluation [10], recently released by the International Atomic Energy Agency (IAEA) Nuclear Data Section, it is stated that new absolute measurements of the neutron induced fission cross sections (e. g. relative to n-p scattering) on uranium, bismuth, lead and plutonium have the highest priority in establishing neutron induced fission reaction standards above 200 MeV.

The n\_TOF facility is then utilized in for the measurement of neutron-induced reaction cross-section as a function of the energy: the advantage of the time-of-flight technique with respect to monoenergetic neutron source is evident, in particular for the study of the fine structures in the cross section. Hence the n\_TOF facility offers a good opportunity to perform the measurement of  $^{235}\text{U}(n,f)$ , in particular where the neutron energy spectrum extends from thermal energies up to more than 1 GeV.

A set of uranium targets are prepared, and a neutron beam impinges on it, fissions are detected, and data are stored. In parallel with the fission detector, a series of 3 Proton Recoil Telescopes (PRT) will be used for measuring the number of neutrons impinging on the  $^{235}\text{U}$  samples, to characterize the incident beam itself. The simulated experimental setup is reported in Fig. 3.

As the PRT exploits the elastic scattering on hydrogen by detecting the recoil proton – coming from polyethylene targets downstream the uranium portion - these detectors are placed out of the neutron beam, and look at the hydrogen-rich samples irradiated by the beam. The drawback of the n-p scattering is the difficulty in handling high-density hydrogen samples. We plan to use a polyethylene  $\text{C}_2\text{H}_4$  sample since this material presents a favourable stoichiometric ratio between hydrogen and carbon, although the presence of carbon in the sample is an important source of background. In addition, because of the energy loss of protons in the sample itself, different thicknesses of  $\text{C}_2\text{H}_4$  are required for different energy ranges. In particular, the full energy region between  $E_n = 20$  MeV and  $E_n = 1$  GeV can be covered simultaneously by using 3 PRTs looking at a 2, 4 and 10 mm thick  $\text{C}_2\text{H}_4$  sample, respectively.



**Fig.3:** n\_TOF experimental setup: neutron telescopes (PRT-L, H) and  $\text{C}_2\text{H}_4$  intermediate targets.

Monte Carlo simulations have been carried out to support the analysis of this experiment through the code MCNP6.2 [11], to take advantage of the updated tools implemented in the latest version that allow to tag particular particles according to specific nuclear reactions.

Preliminary simulations were conducted to tune the optimal particle species to be transported – namely photon, neutron, electrons, deuterons, protons, and tritons.

Energy deposition tallies – as well as different available estimators – were compared. MCNP tool of energy deposition together with multi-particle tally type were considered as the ideal estimator for the problem in question.

Energy deposition in polyvinyl toluene composing the PRT telescopes has been the main objective of the study. Issues regarded the combined use of models and data tables: in fact, proton tables are present up to 150 MeV - thanks to the MCNP library LA150h – that have been used combined with Bertini model.

Experimental uncertainties concerning the beam centering and the radial shape have been verified. Small shift in centering has been retained and a Gaussian shape has been implemented for the radial distribution of the particles in the beam. Gauss parameters have been also considered as energy-dependent, according to the neutron source.

The main goal of the simulations has been the characterization of a test neutron beam impinging on the C<sub>2</sub>H<sub>4</sub> targets (Fig. 3), constituted of a series of sources homogeneous in energy and formed by many tiny intervals, from about 10 MeV up to 1 GeV.

Comparing the characterization for a given neutron beam downstream the uranium targets and the fission detector response it will be possible to completely understand the experiment outcome and enhance <sup>235</sup>U fission reaction cross-sections.

## Acknowledgement

The work on Gen-II and III PWR's was carried out in close collaboration with l'Institut de Radioprotection et de Sûreté Nucléaire (IRSN). IRSN supplied the MCNP models and the requested responses: the authors would like to thank Mariya Brovchenko.

The author thanks the INFN (National Institute for Nuclear Physics) – Bologna Section – for the collaboration concerning n\_TOF facility and all the information to prepare the calculations.

Simulations were performed through the CRESCO supercomputing facility since large HPC techniques and massive computing environment were necessary to achieve these calculations.

## References

- [1] Goorley J. et al., "Initial MCNP6 Release Overview - MCNP6 version 1.0", LA-UR-13-22934R.
- [2] Brovchenko M., et al. "Neutron-gamma flux and dose calculations in a Pressurized Water Reactor (PWR)", International Conference on Radiation Shielding (ICRS13), Paris, 3-6 Oct 2016
- [3] Burn K.W., et al. "Monte Carlo Methods for Core Safety Analysis: Design of Instrumentation for Monitoring PWR Core Degradation in Case of a Severe Accident". ENEA Report RdS 2014/063
- [4] Burn K.W., "Optimizing variance reduction in Monte Carlo eigenvalue calculations that employ the source iteration approach", Ann. Nucl. Energy 73, 218-240, 2014
- [5] A.D. Carlson, et al., Nuclear Data Sheets 110 (2009) 3215
- [6] D. Tarrío, et al. (The n\_TOF Collaboration), Phys. Rev. C 83 (2011) 044620
- [7] A.D. Carlson, O.A. Wasson, P.W. Lisowski, J.L. Ullmann, N.W. Hill, in Proc. of the Int. Conf. on Nuclear Data for Science and Technology, Juelich 1991, p.518 (1991), Germany
- [8] R. Nolte, et al., Nuclear Science and Engineering, 156 (2007) 197
- [9] S. Lo Meo, D. Mancusi, C. Massimi, G. Vannini, A. Ventura, , Nucl. Phys. A 933 (2015) 43

- [10] INDC International Nuclear Data Committee, "209Bi(n, f) and natPb(n, f) Cross Sections as a New Reference and Extension of the 235U, 238U and 239Pu(n, f) Standards up to 1 GeV", INDC(NDS)-0681 Distr. ST/J/G/NM, IAEA 2015
- [11] C.J. Werner (editor), "MCNP Users Manual - Code Version 6.2", LA-UR-17-29981 (2017)

# COMPUTATION OF SOLAR PHOTON SAIL THRUST BY USING LIGHT SCATTERING MODELS

Danilo Zola<sup>1\*</sup>, Salvatore Scaglione<sup>1</sup>,

<sup>1</sup>*ENEA Dipartimento Tecnologie Energetiche e Fonti Rinnovabili (TERIN),  
Centro Ricerche Casaccia, via Anguillarese 301, IT-00123, Roma, Italy*

**ABSTRACT.** The computation of the solar photon sail thrust needs of complex optical models whose numerical solution requires large computation resources. By parallelizing the algorithm and thanks to CRESCO relevant resources, the accuracy is increased saving time

## 1 Introduction

The goal of research outlined in this work, is to calculate the thrust exerted by sunlight on a Solar Photon Sail (SPS) deployed in the space. The sunlight causes a pressure, (the radiation pressure), on the exposed surface. In particular, the radiation pressure on a surface that reflects 100% of the incident sunlight, is equal to  $9.56 \mu\text{Pa}$ . This pressure generates a force, (the thrust), which will be more intense the larger the sail surface. The spacecraft acceleration depends on the sail and payload' masses, ( $\mathbf{F} = m\mathbf{a}$ ). An acceleration, even very small, can produce over time a very high speed, such that a spacecraft can cover in a few years such distances that would not be reachable with current propulsion technologies. Thanks to SPS, space missions can be achieved on trajectories on retrograde orbit or out of the ecliptic plane, which are unfeasible by using chemical or ionic propulsion. Furthermore, the spacecraft does not need to have fuel on board, especially if the attitude is operated by the same solar photon propulsion. The limits in today's use of this technology are essentially determined by the mass of the vehicle as a whole, (sail and payload), and by the effective optical properties of SPS. Nevertheless, the miniaturization of electronic instruments on board and the technological development in the field of materials makes promises to use this innovative technology of propulsion for space missions in the next decade. Since the thrust is due to the interaction between the incident light and the sail surface, the accuracy of the prediction of the trajectory travelled by a SPS spacecraft depends on having used a realistic optical model including the all optical features by its sail, because incident light of the Sun is actually specularly reflected and scattered, as well as partially absorbed and emitted. Several optical models are currently under investigation that include light scattering due to the roughness of the sail surface. Models including local pleats or curvatures of the surface have been considered not yet. This investigation leads to the elaboration of very detailed physical models that require substantial computing resources to obtain a solution with the desired precision. CRESCO allows to completely parallelize the algorithm with a considerable computation time reduction.

## 2 Computation of the solar photon sail thrust

Although the photons are massless they have a momentum (relativistic quantum theory). Due to the second and the third principle of dynamics, a change in the momentum of the photons generates a force

---

\*Corresponding author. E-mail: danilo.zola@enea.it.

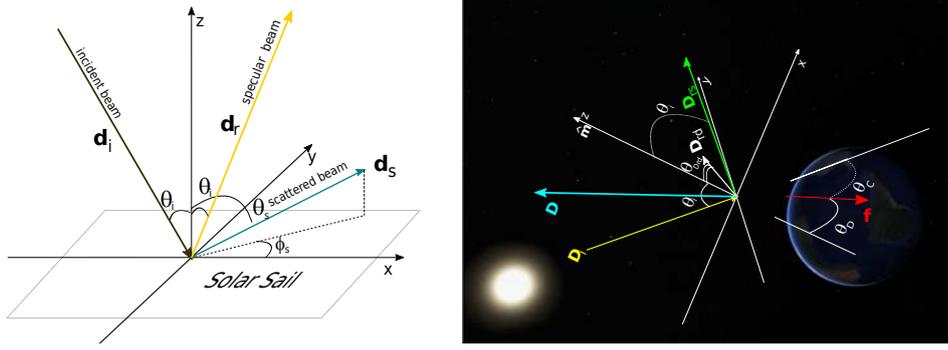


Figure 1: Left Panel): Optical geometry for defining the main angles and versors used in the text. (a) Right Panel): Illustration of the momentum vectors and the main angles discussed in this work.(b)

on the sail. The change of photon momentum  $\mathbf{D}$  per time and area unit is given by  $\mathbf{D} = \mathbf{D}_r - \mathbf{D}_i$  i.e. the difference between the momentum of the reflected photons  $\mathbf{D}_r$  and the momentum of incident photons  $\mathbf{D}_i$ . The momentum  $\mathbf{D}_i$  of the incoming radiation is given by:

$$\mathbf{D}_i = \frac{W(\rho_s)}{c} \hat{\mathbf{d}}_i \quad \text{with} \quad \hat{\mathbf{d}}_i = \begin{pmatrix} \sin \theta_i & 0 & -\cos \theta_i \end{pmatrix}^{\mathbb{T}} \quad (1)$$

$W$  is the energy flux at a distance  $\rho_s$  from the Sun.  $W$  is known experimentally at 1 A.U. and it can be estimated at any distance. The superscript  $\mathbb{T}$  in Equation 1 stands for transposed.

The momentum given by the reflected photons both specularly as well as scattered is given by:

$$\mathbf{D}_r(\theta_i) \equiv \frac{1}{c} \int_{\lambda_{min}}^{\lambda_{max}} I_d(\lambda) d\lambda \int_{\Omega_s} \cos \theta_s \text{BRDF}(\lambda, \theta_i, \theta_s, \phi_s) \hat{\mathbf{d}}_s d\Omega_s \quad (2)$$

In Equation (2),  $c$  is the light speed in vacuum,  $I_d(\lambda)$  is the photon flux (measured as  $\text{W nm}^{-1} \text{m}^{-2}$ ), which is a function of wavelength  $\lambda$  and depends on the distance from the Sun and crossing  $d\mathbf{S}_{\perp}$  which is the infinitesimal surface perpendicular to the photons flux. The wavelengths  $\lambda_{min}$  and  $\lambda_{max}$  limit the spectral range where the irradiance is not negligible. BRDF is the Bidirectional Reflection Distribution Function defined as the ratio between the differential radiance and the differential irradiance. Usually, BRDF is a function of  $\lambda$ , incident angle  $\theta_i$ , out of plane scattering angle  $\theta_s$  and in plane scattering angle  $\phi_s$ . The three angles are defined according to the Figure 1a. Moreover,  $\Omega_s$  is the scattering solid angle and  $\hat{\mathbf{d}}_s = \begin{pmatrix} \sin \theta_s \cos \phi_s & \sin \theta_s \sin \phi_s & \cos \theta_s \end{pmatrix}^{\mathbb{T}}$  is the unit vector of the scattered photons momentum. For  $\theta_s = \theta_i$  and  $\phi = 0$ , the versor  $\hat{\mathbf{d}}_s = \hat{\mathbf{d}}_r$  locates the specular direction.

The momentum  $\mathbf{D}_r$  induced by the Sunlight irradiance at  $\theta_i$  can be estimated if the BRDF of the SPS is known for any  $\lambda$ ,  $\theta_s$ ,  $\phi_s$ .

The force acting on the sail with a surface  $S$  when the solar radiation points at a pitch angle  $\theta_i$  is given by  $\mathbf{f} = -\mathbf{D}S \cos \theta_i$ .

The photon momentum due to the reflected radiation  $\mathbf{D}_r$  can be decomposed like  $\mathbf{D}_r = \mathbf{D}_{rs} + \mathbf{D}_{rd}$ , where  $\mathbf{D}_{rs}(\theta_i)$  is the photon momentum due to the specular beam given by:

$$\mathbf{D}_{rs} \equiv \frac{1}{c} \int_{\lambda_{min}}^{\lambda_{max}} I_d(\lambda) d\lambda \int_{\Omega_s} \cos \theta_s \text{BRDF}_S \hat{\mathbf{d}}_s d\Omega_s \quad (3)$$

The BRDF of reflected photons in the specular direction is given by:

$$\text{BRDF}_S(\lambda, \theta_i; \theta_s, \phi_s) = \frac{R_S(\lambda, \theta_i)}{\cos \theta_i \sin \theta_i} \delta(\theta_i - \theta_s) \delta(\phi_s) \quad (4)$$

and  $R_S(\lambda, \theta_i)$  is the *directional* spectral specular reflectance which can be measured or indirectly determined experimentally. The momentum of scattered photons due to the surface roughness is:

$$\mathbf{D}_{rd}(\theta_i) \equiv \frac{1}{c} \int_{\lambda_{min}}^{\lambda_{max}} I_d(\lambda) d\lambda \int_{\Omega_s} \cos \theta_s \text{BRDF}_D \hat{\mathbf{d}}_s d\Omega_s \quad (5)$$

and  $\text{BRDF}_D$  is the part of the total BRDF due to the scattered light only.

In Figure 1b, there is an illustration of the momentum vectors  $\mathbf{D}$ ,  $\mathbf{D}_i$ ,  $\mathbf{D}_{rd}$  and  $\mathbf{D}_{rs}$ . Furthermore,  $\theta_D$  and  $\theta_C$  are respectively the center line angle and the cone angle, whereas  $\theta_{rd}$  is the angle formed by  $\mathbf{D}_{rd}$  with the sail normal  $\hat{\mathbf{m}}$ .

For evaluating the photons momentum due to scattered radiation, the  $\text{BRDF}_D$  for any  $\lambda$  has to be computed. Several scattering models could be in principle used, (see ref. [1] for a critical review). In this work, the Rayleigh-Rice theory (RRT) [2] is used in order to explain the futures of the algorithm written for computing the SPS thrust. According to RRT, the BRDF of scattered photons is given by[3, 4]:

$$\text{BRDF}_D(\lambda, \theta_i, \theta_s, \phi_s) = \frac{16\pi^2}{\lambda^4} \cos \theta_i \cos \theta_s Q[\epsilon(\lambda), \theta_i, \theta_s, \phi_s] \text{PSD}[f_x(\lambda, \theta_i, \theta_s, \phi_s), f_y, \tilde{z}, \ell] \quad (6)$$

The polarization factor  $Q[\epsilon(\lambda), \theta_i, \theta_s, \phi_s]$  does not depend on the surface morphology but just on the intrinsic optical properties of the material constituting the surface of the sail, via the complex dielectric function  $\epsilon(\lambda)$ . PSD is a function in the space of the spatial frequencies  $f_x, f_y$  of the surface' topography and it is parameterized by the root mean square roughness  $\tilde{z}$  and the coherence length  $\ell$  which can be determined experimentally.[2] The spatial frequencies  $f_x$  and  $f_y$  are related to  $\theta_i, \theta_s$ , and  $\phi_s$  according to the following relation:

$$f_x = \frac{\sin \theta_s \cos \phi_s - \sin \theta_i}{\lambda} \quad (7)$$

$$f_y = \frac{\sin \theta_s \sin \phi_s}{\lambda} \quad (8)$$

In the following, the algorithm used to compute the three components of  $\mathbf{D}_{rd}$  is briefly outlined.

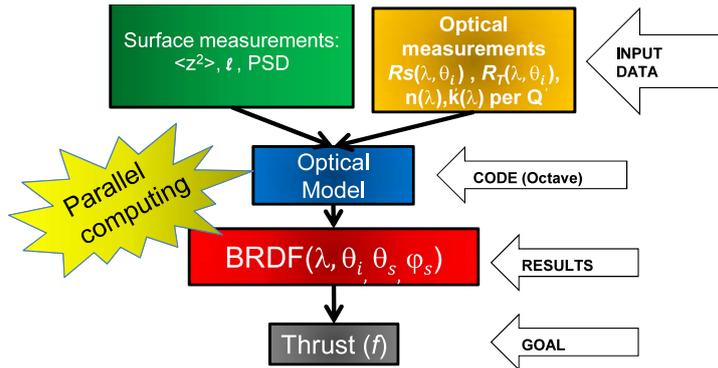


Figure 2: Illustration of the different steps followed for the computation of the SPS thrust.

For computing one point, i.e. a SPS momentum  $\mathbf{D}$  it is necessary to fix a value for  $\tilde{z}, \ell$  and  $\theta_i$ . In the following, the index of vectors and matrices will be displayed as superscripts for major clearness. The angle  $\theta_i$  is discretized by considering a step of 1 deg in the  $0 \leq \theta_i^m \leq \pi/2$  range within  $m = 1, 2, \dots, 91$ . The angles  $\theta_s$  and  $\phi_s$  are discretized with a vector of  $N = 180$  elements as respectively  $\phi_s^i = 0, \dots, i \frac{4}{N} \frac{\pi}{2}, \dots, 2\pi$  for  $i = 1, \dots, N$  and  $\theta_s^j = 0, \dots, j \frac{1}{N} \frac{\pi}{2}, \dots, \pi/2$  for  $j = 1, \dots, N$ . In this

way,  $\hat{\mathbf{d}}_s, \hat{\mathbf{d}}_i$  are also straightforward to discretize.

The Sun irradiance  $I(\lambda)$  is discretized on a vector  $\lambda^n$  considering a 1 nm step within  $200 \text{ nm} \leq \lambda^n \leq 20000 \text{ nm}$  spectral range where  $n = 1, 2, \dots, 19801$ . The vector  $I^n$  is the discretized AM0 experimental irradiance and in the same manner, the experimental complex dielectric function  $\epsilon(\lambda)$  becomes the complex vector  $\epsilon^n$  with 19801 components.

Discretizing  $f_x$  and  $f_y$  by means of Equations. 7 and 8, a  $N \times N$  matrix for PSD is obtained and in the same way the matrix for  $Q$  is also built. Finally, for a fixed  $\lambda^l$ ,  $\text{BRDF}_D$  is discretized via (Eq. 6)

$$\text{BRDF}_D^{i,j}(\lambda^l, \theta_i, \phi_s^i, \theta_s^j) = 16(\pi/\lambda^4) \cos(\theta_i) \cos(\theta_s^j) Q^{i,j} \text{PSD}^{i,j}$$

By performing a triple numerical integration in  $d\lambda, d\theta_s, d\phi_s$ , the SPS momentum  $\mathbf{D}_{rd}$  due to the scattered photons is computed, (see Eq. 5). By using the resulting  $\mathbf{D}_i$  of Eq. 1 and  $\mathbf{D}_{rs}$  of Eqs. 3 and 4, the SPS thrust can be finally estimated.

The algorithm is written in a GNU Octave language by defining a function

```
function [dDrd,x dDrd,y dDrd,z] = f(λ, ε, θi, z̃, ℓ),
```

which compute for a fixed  $\lambda, \theta_i, \tilde{z}, \ell$ , the components of the three integrands:

$$d\mathbf{D}_{rd}(\theta_i, \lambda) \equiv \int_{\Omega_s} \cos \theta_s \text{BRDF}_D \hat{\mathbf{d}}_s d\Omega_s \quad (9)$$

A main program calls the function  $f$  cycling the function for different  $\lambda^n$  values.

In CRESCO, the function  $f(\lambda, \epsilon, \theta_i, \tilde{z}, \ell)$  is parallelized by using the package *parallel* of GNU OCTAVE and calling `pararrayfun`:

```
[ dDrd,x dDrd,y dDrd,z ] = pararrayfun(Np, @f, λ, ε, θi, z̃, ℓ)
```

where  $N_p$  is the number of processors which could in theory be 19800 for a fully parallelization of the computation. The last numerical integration of Eq. 5 is performed in the main program. In this way, the SPS momentum is obtained for a fixed  $\theta_i$  and  $\tilde{z}$  and  $\ell$ .

In CRESCO is available to run the different indexed Octave code, by using the LSF command `bsub` and indexing, for example, the 91th different  $\theta_i$  values:

```
bsub -J "Index[1-91]" -q "queue-name" -o "output-filename.txt"
-i "...input. %I" octave -q -f "program-name.m".
```

In this way, the computation time has been significantly reduced if compared with a pc desktop equipped with eight processors.

## References

- [1] T. M. Elfouhaily and C-A Guérin. A critical survey of approximate scattering wave theories from random rough surfaces. *Waves Random Media*, 14(4):R1–R40, 2004.
- [2] John C. Stover. *Optical Scattering: Measurement and Analysis*, volume PM224. SPIE, iii edition, 2012.
- [3] Giovanni Vulpetti. Applying vector scattering theory to solar-photon sail thrust modeling. In Malcolm Macdonald, editor, *Advances in Solar Sailing*, Springer Praxis Books, pages 489–508. Springer Berlin Heidelberg, 2014.
- [4] D. Zola, C. Circi, G. Vulpetti, and S. Scaglione. Photon momentum change of quasi-smooth solar sails. *J. Opt. Soc. Am. A*, 35(8):1261–1271, Aug 2018.

# AB INITIO CALCULATIONS APPROACH TO STUDY A-Si:H/C-Si INTERFACES

Francesco Buonocore<sup>1</sup>, Pablo Luis Garcia-Muller<sup>2</sup>, Rafael Mayo-Garcia<sup>2</sup>, Simone Giusepponi<sup>1</sup>, Massimo Celino<sup>1</sup>

<sup>1</sup>*ENEA-TERIN-ICT Division, C.R. Casaccia, via Anguillarese 301, 00123 Rome, Italy*

<sup>2</sup>*CIEMAT, Avda. Complutense, 40, 28040 Madrid, Spain*

**ABSTRACT.** In the silicon heterojunction solar cells, intrinsic hydrogenated amorphous silicon (a-Si:H) is used to passivate the crystal silicon (c-Si) surface to suppress the electrical losses at interfaces and to keep ultralow contact resistivity for the selective transport of one type of carrier only. We generated snapshots of the equilibrated c-Si/a-Si:H interface atom configurations at room temperature. The ab initio characterization of molecular dynamics snapshots has been executed on selected configurations to monitor the electronic properties of the c-Si/a-Si:H interface. The evolution of the intragap states is monitored by analyzing density of states and charge density. This all will allow to design more efficient silicon solar cells belonging to the silicon heterojunction technology.

## 1 Introduction

In recent years, the silicon heterojunction (SHJ) solar cells reached the highest efficiency of 26.6% [1], mainly due to the passivation contacts. In these devices, intrinsic hydrogenated amorphous silicon (a-Si:H) was used to passivate the Si surface and the p/n-type doped hydrogenated amorphous silicon was employed to select the transport carriers. The application of SHJs offers several advantages: first, a-Si:H provides efficient passivation of Si dangling bonds at the interface; second, field-effect passivation can produce a significant inversion effect at the a-Si:H/c-Si interface that is able to increase the carrier lifetime. Moreover, fewer process steps are required to build-up the solar cells (SCs), and the low-temperature ( $< 200$  °C) processing allows the use of very thin wafers without any substrate damage.

The amorphous-crystalline heterointerfaces play a crucial role in the photovoltaic operation of SHJ technology, but the microscopic mechanisms of transport and recombination mechanisms at the interface are still poorly understood. The purpose of the present work is to simulate at atomistic resolution a large scale amorphous-crystalline heterointerfaces and to investigate the electronic properties (defects and intragap states) that are preliminary to the study of the transport mechanisms underlying photovoltaic devices based on SHJ technology.

In order to design a predictive ab initio simulation of the a-Si:H/c-Si interface, it is essential to build up an atomic-scale model of the interface that exhibits the experimentally observed features and at the same time it is almost free of defects. Indeed, due to the limited number of atoms in an ab initio calculation, already few defects lead to a strong overestimation of the gap-state density, possibly even resulting in metallic behavior. In the present work, ab initio calculations to investigate the time evolution of the intra-gap states of an a-Si:H/c-Si interface structures constituted by 1,152 atoms whose atomic configurations have been simulated from ReaxFF (Reactive Force Field) molecular dynamics (MD) [2]. Therefore, we are considering a hetero-interface system larger than those of similar studies so that the periodic cell size effects are reduced. This allows us to model more realistic hetero-interfaces. The electronic structure is calculated and analyzed with a focus on the

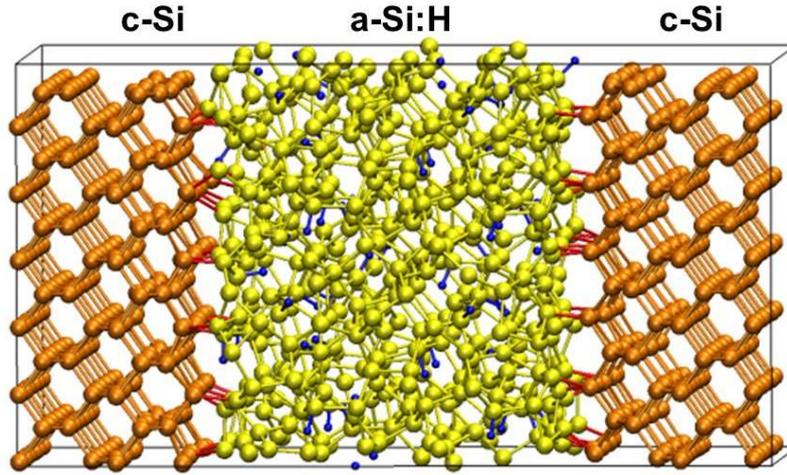
identification and characterization of the intragap states at the interface, which have a crucial impact on the device performance due to their role as recombination centers. We monitor the evolution of the relevant structural and electronic properties, such as the defect distribution, the density of states and the potential barriers. In this way, insights are gained on how and why these properties change.

## 2 Methods

We used first principles calculations based the density functional theory (DFT) to characterize the change of the electronic properties during the equilibration process simulated by MD simulations using the ReaxFF. Snapshots of the thermodynamic properties of the system as well as per atom dynamic values (coordinates, velocities, forces, etc.) are taken at 1 ps intervals for DFT DOS post-processing. The high value of first principles calculations is to derive the physical properties directly from the basic interactions without introducing adjusting parameters.

The two a-Si:H/c-Si interfaces are made of hydrogenated amorphous silicon (a-Si:H) between two crystalline silicon (c-Si) slabs. The relaxed  $p(2\times 1)$  symmetric reconstruction of the Si(001) surface constitutes the c-Si side of the interface. It is formed by 576 Si atoms, 16 layers of silicon with 36 atoms each. The a-Si:H side of the system is generated by cutting the a-Si:H structure, built as in reference [3], such that the surface area is equal to the c-Si side and the thickness is about 16 Å. It is composed of 512 Si atoms and 64 H atoms. The total length of the periodic cell is  $L_z = 46.44$  Å, while in the x- and y-direction the system has  $L_x = L_y = 23.22$  Å. Periodic boundary conditions (PBC) are imposed in all directions.

The geometry of the ab initio relaxed structure is shown in Figure 1. We use the PWscf (Plane-Wave Self-Consistent Field) code of the Quantum ESPRESSO suite [4,5] to relax the a-Si:H/c-Si system. Si and H ultrasoft pseudopotentials with Perdew-Burke-Ernzerhof (PBE) [6] approximant GGA exchange-correlation potential, available in the Quantum ESPRESSO library [7]. The electronic wave functions were expanded in a plane-wave basis set with a kinetic energy cut-off equal to 40 Ry (the charge density cut-off was 240 Ry). The Brillouin zone integration for the self-consistent calculation is restricted to the  $\Gamma$ -point, which is justified by the sufficiently large super cell. Gaussian smearing of 0.08 Ry is needed to reach convergence due to defect states at the Fermi level. All the parameters are chosen by checking the convergence of the total energy of the system. The energy minimization is performed by using conjugate gradient (CG) minimization energy method, with the convergence threshold for self-consistency equal to  $10^{-6}$  Ry. Broyden-Fletcher-Goldfarb-Shanno (BFGS) quasi-newton algorithm is used to perform geometry optimization. Ionic relaxation is stopped when both the following conditions is satisfied: energy changes less than  $10^{-4}$  Ry between two consecutive self-consistent field (SCF) steps and all components of all forces were smaller than  $5 \times 10^{-4}$  Ry/Bohr. Then, this relaxed system is used as starting configuration for energy minimization and classical MD simulations. The non-self-consistent calculation of the electronic states is performed on a  $2\times 2\times 1$  k-point grid, which was found to yield a sufficiently accurate representation of the relevant quantities (density of states, electron localization function, and charge density).

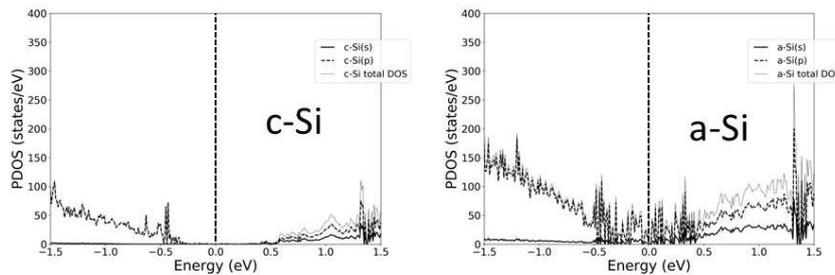


**Fig. 1:** The ab initio relaxed a-Si:H/c-Si interfaces. The silicon atoms and their bonds are in orange in the c-Si side and are in yellow in the a-Si:H side, hydrogen atoms and bonds with silicon atoms are in blue. Bonds between c-Si and a-Si are in red.

### 3 Results and discussion

In this section, the results provided by the aforementioned simulations are reported. These results have been obtained using the Cresco6 cluster of the CRESCO/ENEAGRID High Performance Computing infrastructure [8,9] where around 200,000 CPU hours have been used.

In Figure 1 we show the projected density of states (PDOS) of c-Si and a-Si:H ab initio relaxed, representing the PDOS of the a-Si:H/c-Si interfaces at  $T=0$  K that will be used as comparison. We see that c-Si has a gap of around 0.8 eV, below the experimental value of 1.1 eV. It is well known that standard DFT [10] underestimates band gaps, due to the incomplete description of many-body effects. However, in this study the focus is on the formation of the intragap states related to the defects rather than on the evaluation of the band gap itself. From the PDOS in Figure 2b we see that broad peaks are induced from defects in a-Si:H bulk and at the interfaces.

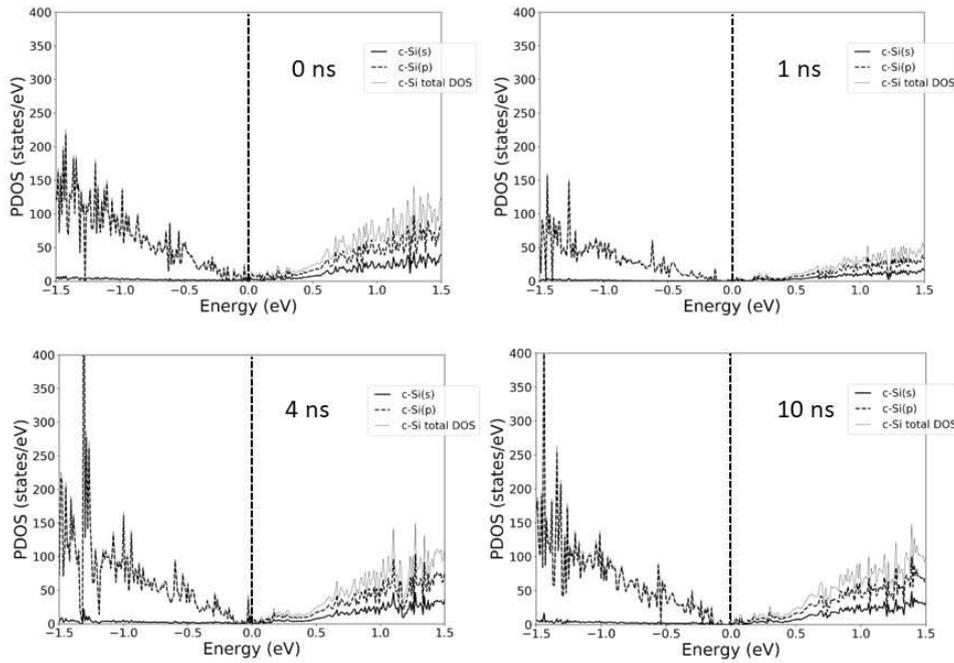


**Fig. 2:** Projected density of states of crystalline and amorphous silicon of the ab initio relaxed a-Si:H/c-Si interfaces. The vertical dashed line at 0 eV evidences the Fermi energy.

We analyze the time evolution of the electronic properties of the a-Si:H/c-Si interfaces to monitor how intragap states change during the equilibration process at 300 K after the quenching. We follow the time evolution of the projected density of states (PDOS) during the equilibration process starting

from  $t=0$  ns, when the process begins until  $t=10$  ns. In Figures 3 and 4, the PDOS of c-Si and a-Si:H at  $t=0, 1, 4,$  and  $10$  ns is shown, respectively. We observe at the start of the equilibration ( $t=0$  ns) that a dense concentration of peaks is found in the energy gap for the a-Si PDOS. Those peaks are related to defects at both the interface and in the a-Si:H bulk. In particular, one intense peak is found nearby the Fermi energy level at  $0$  eV. As the equilibration progresses, the energy of the PDOS peaks changes in energy and the corresponding intensity changes too, both in the intragap range and outside of it. Overall, after the ReaxFF MD equilibration we have a decrease of the density of the defects compared to that of the starting DFT relaxed a-Si:H/c-Si interfaces.

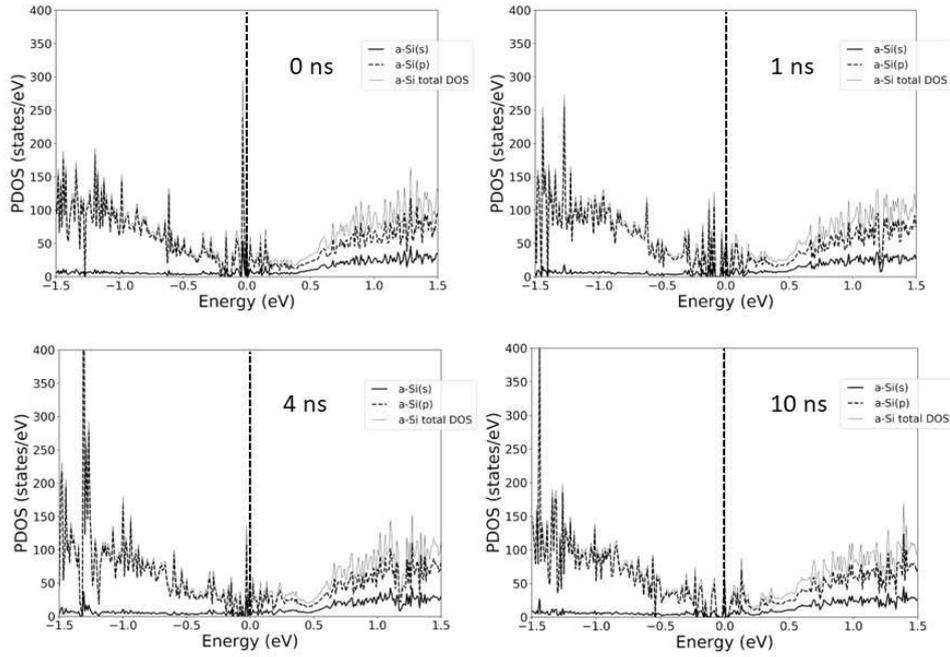
### c-Si, T = 300 K



**Fig.3:** Projected density of states of crystalline silicon at 0, 1, 4 and 10 ns for  $T=300$  K. The vertical dashed line at  $0$  eV evidences the Fermi energy.

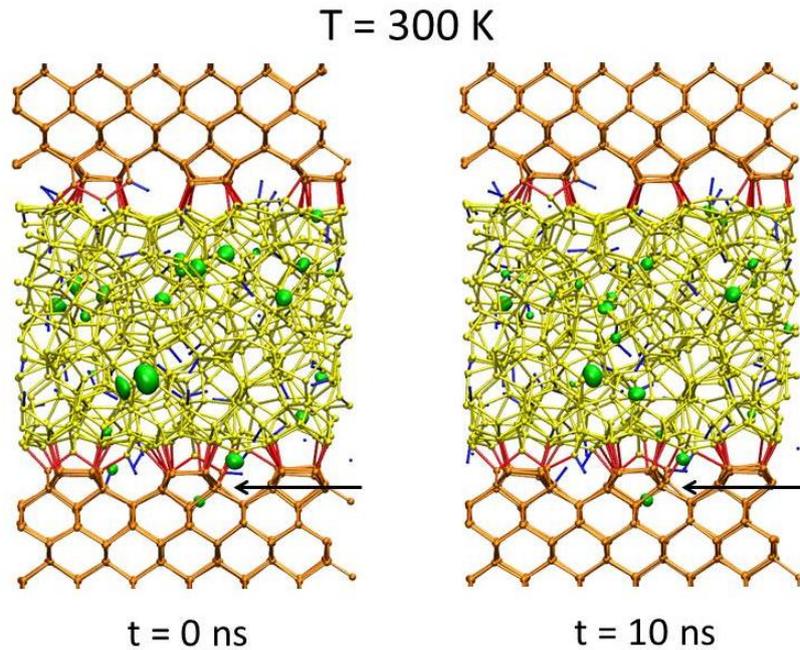
To gain a deeper understanding of the structural properties, a coordination analysis of the Si atoms is performed. A geometrical criterion is used to identify the nearest neighbors in the coordination analysis, applying a distance cutoff of  $2.85$  Å and  $1.7$  Å for Si-Si pairs and for Si-H pairs, respectively. Concerning the  $t=10$  ns snapshot, it is observed that the average number of neighbors of Si atoms is  $4.01$ . In detail,  $16$  Si atoms have threefold coordination ( $1.4\%$ ),  $1044$  Si atoms have fourfold coordination ( $96.0\%$ ) and the remaining  $28$  Si atoms have fivefold coordination ( $2.6\%$ ). Just small variations have been found during the equilibration process taking into account that the average coordination number equals to  $4.01$  for the  $t=10$  ns snapshot too. In conclusion, after the quenching process the system is quite ordered with a high percentage of fourfold coordinated Si atoms, and it keeps this condition during thermalisation at room temperature.

## a-Si, T = 300 K



**Fig. 4:** Projected density of states of amorphous silicon at 0, 1, 4 and 10 ns for  $T = 300$  K. The vertical dashed line at 0 eV evidences the Fermi energy.

In order to elucidate the spatial localization, we calculated the local DOS (LDOS) of the intragap energy levels. In Figure 5, we compare the LDOS at  $t = 0$  and 10 ns. We found that defect states are localized both in the bulk of a-Si and at the a-Si:H/c-Si interface. Defects can be formed in the few c-Si layers nearest to a-Si:H. However, we see a change in the distribution of defects in the interval of time given to our simulation. Indeed, the number and the intensity of the intragap states in the PDOS at the end of the equilibration process are lower than at the start. We investigated the coordination of the atoms nearby the LDOS isosurface at the interface indicated by the arrows in Figure 5. Following the above criteria for atomic distances, they are defects three-fold coordinated.



**Fig.5:** Local density of states of the intragap states of the a-Si:H/c-Si interfaces at the beginning a) and at the end b) of the thermalisation at T= 300 K.

## 4 Conclusions

In conclusion we used ab initio calculations to investigate the time evolution by ReaxFF MD simulations of the intra-gap states of a large a-Si:H/c-Si interface system. Therefore, the electronic structure is calculated and analyzed with a focus on the identification and characterization of the intragap states, which have a crucial impact on the device performance due to their role as recombination centers. Throughout the annealing process, we monitor the evolution of the relevant structural and electronic properties.

An ab-initio relaxed system with a double a-Si:H/c-Si interface has been used as starting configuration for MD simulations. The minimization at T= 0 K produces a shrinking of the system along the z direction of about 6 Å. Then, in the quenching process, in which the temperature raised up to 1100 K and cooled down to 300 K, it has been observed a displacement of hydrogen atoms towards the interfaces. This trend is maintained during the subsequent thermalisation at T= 300 K for 10 ns, in fact, the concentration of hydrogen atoms near the interfaces remains almost constant.

We have found that at the end of the equilibration process of 10 ns at room temperature the intensity of the PDOS related to intragap states is decreased as well as the number of the electronic states into the gap. Nonetheless, the defects states are still localized both in the bulk of a-Si that at the interface with c-Si, until to be formed in the few c-Si layers nearest to a-Si:H. However, the system is quite ordered after the quenching process with a high percentage of fourfold coordinated Si atoms, and it keeps this condition during the final thermalisation. Further studies are in progress to investigate the high temperatures effects on the electronic properties of realistic a-Si:H/c-Si interfaces. This study paves the way to the investigation of the transport mechanisms in order to design more efficient silicon solar cells based on the SHJ technology.

## Acknowledgements

The computing resources and the related technical support used for this work have been provided by CRESCO/ENEAGRID High Performance Computing infrastructure and its staff [8,9]. CRESCO/ENEAGRID High Performance Computing infrastructure is funded by ENEA, the Italian National Agency for New Technologies, Energy and Sustainable Economic Development and by Italian and European research programs. The Spanish Ministry of Science, Innovation, and Universities (CODEC-OSE, RTI2018-096006-B-I00), the Comunidad de Madrid (CABAHLA project, S2018/TCS-4423), and the CYTED Network RICAP (517RT0529), with ERDF funds are acknowledged. The authors acknowledge financial funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824158 (EoCoE-II).

## References

- [1] K. W. Yoshikawa *et al.*, Exceeding conversion efficiency of 26% by heterojunction interdigitated back contact solar cell with thin film Si technology. *Sol. Energy Mater. Sol. Cells* 173 (2017) 37.
- [2] F. Buonocore, P. L. Garcia-Muller, S. Giusepponi, M. Celino, R. Mayo-Garcia, A combined classical molecular dynamics simulations and ab initio calculations approach to study a-Si:H/c-Si interfaces. *Proceedings of the 2020 Winter Simulation Conference*, accepted for publication.
- [3] P. Czaja, M. Celino, S. Giusepponi, M. Gusso, U. Aeberhard, Ab initio study on localization and finite size effects in the structural, electronic, and optical properties of hydrogenated amorphous silicon. *Comp. Mat. Sci.* 155 (2018) 159–168.
- [4] P. Giannozzi *et al.*, Advanced capabilities for materials modelling with Quantum ESPRESSO. *J. Phys.: Condens. Matter* 29 (2017) 465901.
- [5] P. Giannozzi *et al.*, Quantum ESPRESSO: a modular and open-source software project for quantum simulations of materials. *J. Phys.: Condens. Matter* 21 (2009) 395502.
- [6] J. P. Perdew, K. Burke, M. Ernzerhof, Generalized Gradient Approximation made simple. *Phys. Rev. Lett.* 77 (1996) 3865.
- [7] <http://www.quantum-espresso.org>. Accessed 20th April 2020.
- [8] F. Iannone *et al.*, CRESCO ENEA HPC clusters: a working example of a multifabric GPFS Spectrum Scale layout. *2019 International Conference on High Performance Computing & Simulation (HPCS)*, Dublin, Ireland, 2019, pp. 1051-1052.
- [9] A. Mariano *et al.*, Fast Access to Remote Objects 2.0 a renewed gateway to ENEAGRID distributed computing resources. *FGCS* 94 (2019) 920-928.
- [10] J. P. Perdew, Density functional theory and the band gap problem. *Int. J. Quantum Chem.* 28(S19) (1985) 497.

# FIRST PRINCIPLES MOLECULAR DYNAMICS SIMULATIONS OF AMORPHOUS $\text{GeO}_2$ UNDER PRESSURE IN THE RANGE 0-30GPA

Giorgio Mancini<sup>1\*</sup> and Massimo Celino<sup>2</sup>

<sup>1</sup>*Università Di Camerino, Sezione di Fisica della Scuola di Scienze e Tecnologie, Via Madonna delle Carceri 9B, 62002 Camerino (MC), Italia*

<sup>2</sup>*ENEA, Ente per le Nuove Tecnologie, l'Energia e lo Sviluppo Economico Sostenibile, C. R. Casaccia, Via Anguillarese 301, 00123 Roma, Italia*

**ABSTRACT.** As part of a process aimed to get data for the widest possible high pressures range, a series of *First Principles Molecular Dynamics* simulations were carried out for solid state amorphous  $\text{GeO}_2$  subjected to hydrostatic pressures. The simulations were carried out by Carr-Parrinello Molecular Dynamics (CPMD) [1, 2] using Goedecker pseudo-potentials [3]. In this paper we report the preliminary results for the range 0-30 GPa we covered so far.

## 1 Introduction

A preliminary series of *First Principles Molecular Dynamics* simulations for  $\text{GeO}_2$  were carried out as an exploration probe in the range 0-20GPa [4] to get a first qualitative set of data to compare the changes in amorphous Ge and  $\text{GeO}_2$  under pressure. They suggested that although Ge and  $\text{GeO}_2$  had qualitatively the same density evolution up to 10GPa,  $\text{GeO}_2$  didn't undergo the steep LDA-HDA phase transition showed by Ge in the range 10-12GPa; the pressure-induced density changes in  $\text{GeO}_2$  kept linear in the entire range 0-20GPa and reversed to the initial state upon pressure release back to 0GPa. The probing range was eventually extended up, but serious numerical instabilities arose at about 28GPa. According to the first exploration indications, we switched to Goedecker pseudo-potentials for a second, longer and accurate series of simulations.

## 2 Simulations details

Our starting point was constituted by amorphous  $\text{GeO}_2$  at room temperature, obtained quenching down from 4000K a liquid  $\text{GeO}_2$  sample. It consisted of 240 atoms (80 for germanium and 160 for oxygen) contained in a 15.602Å-edged cubic simulation box to get a density  $\rho = 3.66 \text{ gr/cm}^3$ . The simulations were performed using CPMD software, describing the self consistent evolution of the electronic structure within the frame of density functional theory. A generalized gradient approximation (BLYP-GGA) was adopted for the exchange and correlation part of the total energy [5, 6] together with norm conserving Goedecker pseudo-potentials for the core-valence interactions. The electronic wave functions were expanded in plane waves up to the kinetic energy cutoff of 100 Ry. A value of 400a.u. was used for the fictitious electronic mass,  $2000\text{cm}^{-1}$  and  $10000\text{cm}^{-1}$  for the characteristic frequencies of ions and electrons, respectively, and 2a.u. (0.048 fs) for the time-step. The initial system was quenched down to

---

\*Corresponding author. E-mail: giorgio.mancini@unicam.it.

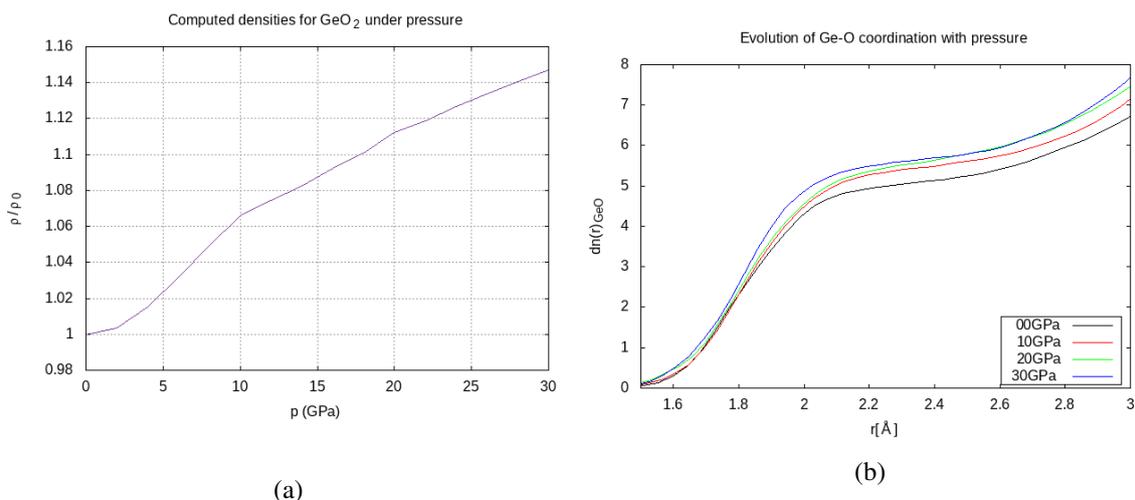


Figure 1: (a) Density slopes identify phase transition points (b) corresponding Ge-O coordination numbers

300K and relaxed at an initial pressure of 0Pa; it subsequently was submitted to increasing pressures up to 30GPa. Pressures were applied via the Parrinello-Rahman method [7], in steps of 2GPa. At each stage the atomic system was equilibrated for  $\sim 50$ ps according to the criterion that all the physical quantities converged. Thermostatting was obtained by Nosé-Hoover chain thermostats [8–10] set on ions and electronic degrees of freedom, a fictitious electronic kinetic energy of 0.05au was used throughout the entire pressure range.

### 3 Computational resources

The calculations were performed using the facilities and services available at the ENEAGRID infrastructure (Italy). Molecular Dynamics simulations have been carried out using CPMD v3.15.3 running on CRESCO4 cluster. 800GB of disk storage has been granted on the GPFS file system.

### 4 Results

A description of the density evolution with pressure is reported in Figure 1 (a) for  $\text{GeO}_2$ . It clearly shows that  $\text{GeO}_2$  undergoes no sudden LDA-HDA transformation (in turn Ge shows a marked slope change at  $\sim 10$ GPa). Four-fold $\rightarrow$ five-fold $\rightarrow$ six-fold coordinated Ge formation is depicted in Figure 1 (b). Density grows with pressure presenting regular, distinct rates in four intervals, in which Ge-O coordination numbers remain invariant, leading to a step-wise overall variation in the 0-30GPa range ( $dn(r)$  curves superpose in 12-20GPa and 22-30GPa). The system is formed by four-fold and five-fold coordinated Ge atoms up to 6GPa. Formation of six-fold coordinated Ge is observed in the interval 6-10GPa. In 10-20Gpa the system presents five-fold and six-fold coordinated Ge atoms: in this range the five-fold coordinate Ge atoms are the majority. The transition at 20GPa inverts the roles and the six-fold coordinated Ge atoms become predominant up to 30GPa.

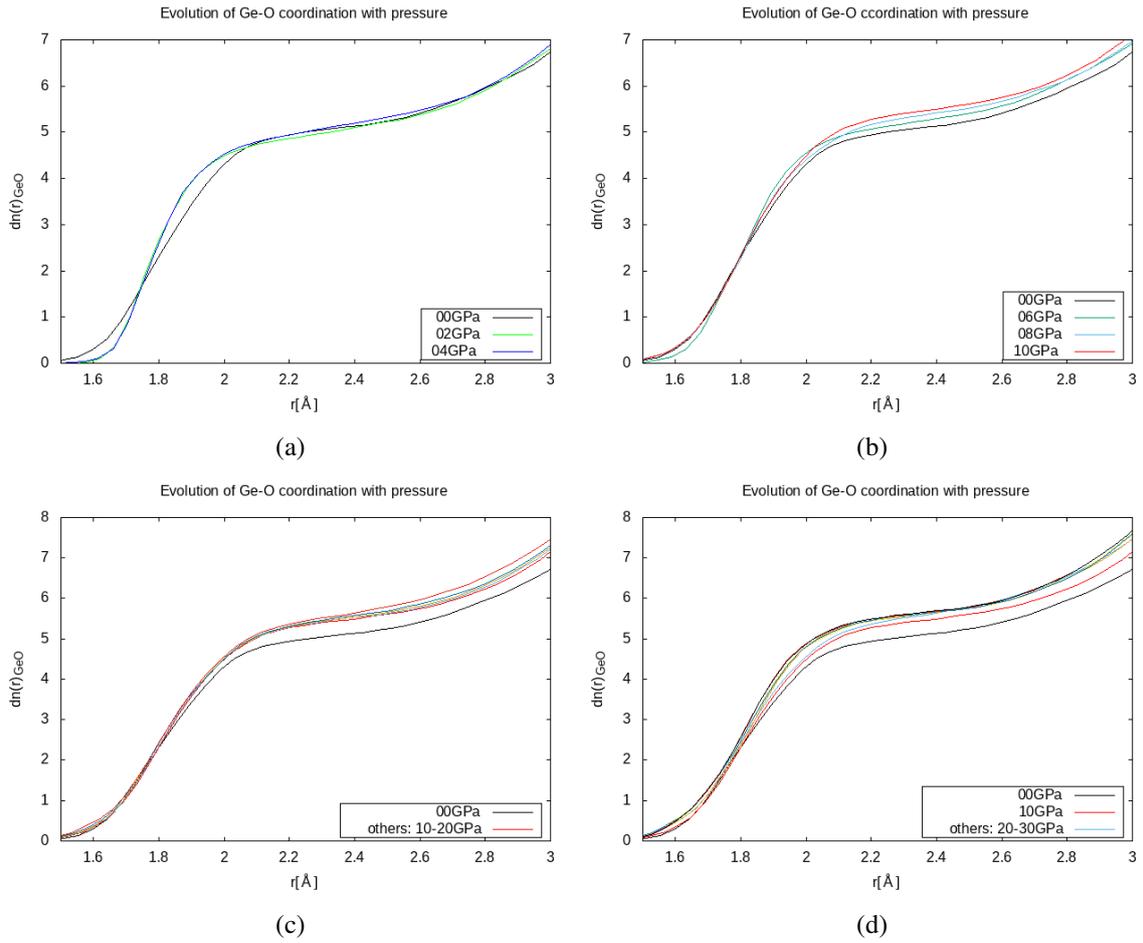


Figure 2: Stepwise changes of  $dn(r)_{GeO}$  in the intervals (a) 0-4GPa, (b) 4-10GPa, (c) 10-20GPa and (d) 20-30GPa

## 5 Conclusions

CPMD simulations for GeO<sub>2</sub> were performed in the interval 0-30GPa, in which Goedecker pseudopotentials showed a very good numerical stability. Although they determined a higher density than 3.66 gr/cm<sup>3</sup> for the quenched system initially relaxed at 0GPa (a fact deserving further investigations), the results were in good qualitative agreement with known computational and experimental findings (e.g. [11, 12] and references therein). A step-wise evolution of coordination numbers was observed.

## References

- [1] R. Car and M. Parrinello. Unified approach for molecular dynamics and density-functional theory. *Physical Review Letters*, 55(22):2471–2474, 1985.
- [2] CPMD v3.13.2 copyright ibm corp 1990-2008, copyright mpi für festkörperforschung stuttgart 1997–2001.
- [3] S. Goedecker, M. Teter, and J. Hutter. Separable dual-space gaussian pseudopotentials. *Physical Review B*, 54:1703–1710, 1996.
- [4] G. Mancini, M. Celino, A. Di Cicco, and E Covino. *Ab-Initio* simulations of amorphous geo<sub>2</sub> under pressure: A first search for phase transitions. *High Performance Computing on CRESCO infrastructure: research activities and results 2018*, pages 89–93, December 2019. ISBN: 978-88-8286-373-9.
- [5] A. D. Becke. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A*, 38:3098–3100, Sep 1988.
- [6] Chengteh Lee, Weitao Yang, and Robert G. Parr. Development of the colle-salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B*, 37:785–789, Jan 1988.
- [7] M. Parrinello and A. Rahman. Crystal structure and pair potentials: A molecular-dynamics study. *Phys. Rev. Lett.*, 45:1196–1199, Oct 1980.
- [8] The CPMD Consortium. Car-Parrinello Molecular Dynamics: An ab-initio Electronic Structure and Molecular Dynamics Program. Manual for CPMD version 3.15.1.
- [9] S Nosé. A molecular dynamics method for simulations in the canonical ensemble. *Molecular Physics*, 52:255–268, 1984.
- [10] S.A Nosé. A unified formulation of the constant temperature molecular dynamics methods. *The Journal Of Chemical Physics*, 81:511–519, 1984.
- [11] Philip S Salmon et al. Density-driven structural transformations in network forming glasses: a high-pressure neutron diffraction study of GeO<sub>2</sub>glass up to 17.5 GPa. *Journal of Physics: Condensed Matter*, 24(41):415102, sep 2012.
- [12] M. Micoulaut, Y. Guissani, and B. Guillot. Simulated structural and thermal properties of glassy and liquid germania. *Phys. Rev. E*, 73:031504, Mar 2006.

# LARGE EDDY SIMULATION OF H<sub>2</sub>/O<sub>2</sub> COMBUSTION WITH O<sub>2</sub> TRANSCRITICAL INJECTION

E. Giacomazzi<sup>1</sup>, D. Cecere<sup>1\*</sup>, N. Arcidiacono<sup>1</sup> and F.R. Picchia<sup>1</sup>

<sup>1</sup>*Process and Energy Systems Engineering Laboratory, ENEA, Rome, Italy*

**ABSTRACT.** The article is devoted to the analysis of supercritical hydrogen/oxygen combustion at 60 bar, with transcritical (liquid) injection of oxygen. The test case considered is the MASCOTTE C-60. The tapered injector is simulated to properly identify the flame stabilization mechanism. The case is simulated by means of the in-house parallel code HeaRT in the three-dimensional LES framework. The cubic Peng-Robinson equation of state is assumed. Diffusive mechanisms and transport properties are accurately modelled. A finite-rate detailed scheme (involving the main radicals) already validated for high-pressure H<sub>2</sub>/O<sub>2</sub> combustion is adopted.

## 1 Introduction

High-pressure combustion of reactants exhibiting real gas behaviour is becoming an important research topic for the growing interest in supercritical CO<sub>2</sub> and organic Rankine gas turbine cycles, diesel engines with higher and higher pressures, liquid oxygen rocket engines, cooling systems. In such applications, the flows can be far away from ideal thermodynamics, and the different fluid behaviour has to be accounted for by means of real gas equations of state and specific models for molecular transport properties. Experimental work at such high-pressure conditions may be prohibitive: the use of advanced laser diagnostics is not an easy task to achieve, and very few examples exist in literature; besides, facilities themselves may be very expensive. Hence, most of research and design in this area is based on numerical simulation. However, also the numerical approach is not an easy task, suffering problems of both modelling accuracy, computational efficiency and lack of robustness by numerical schemes.

## 2 Experimental Set-up

The case simulated in this work was experimentally investigated in the MASCOTTE cryogenic combustion test facility developed by ONERA for rocket applications. It consists of a combustor having a 50 mm × 50 mm square-section, 40 cm length with a final nozzle, and an axisymmetric coaxial injector for O<sub>2</sub>, and H<sub>2</sub> or CH<sub>4</sub>. The particular test case numerically studied in this article has a nominal pressure of 60 bar (case C-60) and considers injection of transcritical (liquid) O<sub>2</sub> and gaseous H<sub>2</sub> [8, 4]. The liquid oxygen is injected from the central pipe, that has a constant section of radius 1.8 mm before a linear divergent section with an angle of 8° and a length of 4.98 mm (the final radius is 2.5 mm); the end of the divergent is squared-off with a width of 0.3 mm. The hydrogen flows through the coaxial pipe having a radial width of 2.2 mm. The thickness of the wall separating the two jets is  $h = 1$  mm. The injection data are reported in Table 1. The reduced temperature for O<sub>2</sub> and H<sub>2</sub> are 0.54 and 8.29,

---

\*Corresponding author. E-mail: donato.cecere@enea.it.

Pressure	$T^{inj}$	$\rho_{PR}^{inj}$	$\mu^{inj}$	$\rho U \mathcal{A}$	$\mathcal{A}$	$U^{inj}$	$u^{inj}$	$\ell_t$	$Re_t^{inj}$	$\eta$
60 bar	K	kg/m <sup>3</sup>	kg/(m · s)	g/s	mm <sup>2</sup>	m/s	m/s	mm		$\mu m$
O <sub>2</sub>	83	1182.97	$2.46 \cdot 10^{-4}$	105	10.1736	8.72	0.436	1	2097	3.2
H <sub>2</sub>	275	5.125	$8.59 \cdot 10^{-6}$	42	53.8824	152.09	7.6	1	4534	1.8

Table 1: Injection data for the two reactants. The density values refer to the Peng-Robinson equation in its enhanced translated volume formulation (the error with respect to NIST reference data are 0.35% and 0.52%, respectively for O<sub>2</sub> and H<sub>2</sub>).

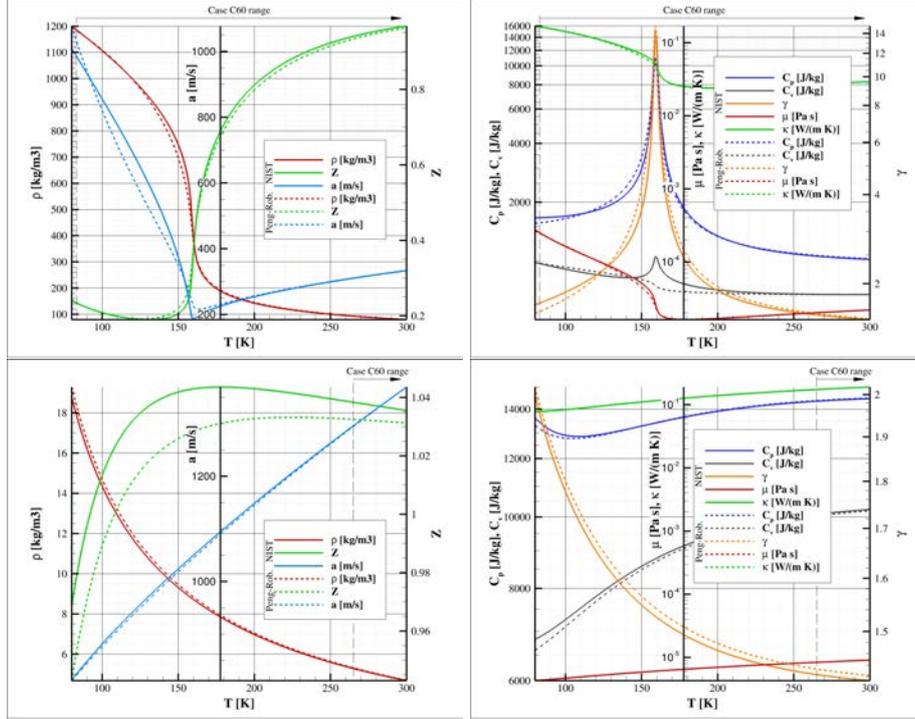


Figure 1: Comparison of some real gas properties estimated through the translated volume Peng-Robinson EoS and NIST reference data for O<sub>2</sub>: EoS and NIST reference data for O<sub>2</sub> (top) and H<sub>2</sub> (bottom): density  $\rho$ , sound speed  $a$ , compressibility factor  $Z$ , specific heat ratio  $\gamma = C_p/C_v$ , viscosity  $\mu$ , thermal conductivity  $\kappa$ .

respectively; their reduced pressure are 1.19 and 4.56. Some real gas properties are reported in Figure 1, where the NIST reference data are compared with the ones estimated through the translated volume Peng-Robinson equation of state. The tapered central pipe is modelled in the simulation in a stepwise manner. The cooling helium film injection along the combustor walls adopted in the experiments is neglected. The computational domain in the streamwise direction is shorter (200 mm) than the experimental combustor and its final nozzle is not considered. A turbulent velocity fluctuation of 5% is imposed at the inlet of both reactants; turbulent Reynolds numbers were estimated considering the thickness of the wall separating the two jets,  $h$ , as the integral turbulent length scale.

### 3 Physical and Numerical Models

In this work the compressible Navier-Stokes equations are solved for a reacting real gas flow. The Peng-Robinson cubic equation of state in its improved translated volume formulation is assumed [2, 9]: it requires only the acentric factor on top of the critical properties and is thus easy to implement for a wide

range of species. It provides slightly better accuracy than the Redlich-Kwong EoS because of a more complex temperature dependence. As a consequence, energies and specific heats predictions are more accurate, an important feature for reacting flows. The mathematical models adopted are derived for a Newtonian and Stokesian fluid of  $N_s$  chemical species. The mass diffusion flux has three contributions: the first due to concentration gradients (modelled through the Hirschfelder and Curtiss' law for multi-component mixtures), the second due to pressure gradients (the baro-diffusion mechanism), and the third one due to temperature gradients (the thermo-diffusion or Soret effect). The heat diffusion has also three contributions: the first due to temperature gradients (the Fourier diffusion), the second due to mass diffusion fluxes, and the third one is the Dufour effect (reciprocal of the Soret effect). The thermo-diffusion, or Soret effect, is the mass diffusion due to temperature gradients, driving light species towards hot regions of the flow. This effect, often neglected, is nevertheless important, in particular for hydrogen combustion, and in general when very light species play an important role. Usually the Dufour effect is negligible even when thermo-diffusion is not, and hence it is neglected. Molecular viscosity and thermal conductivity are accurately modelled through NIST models in REFPROP with an Extended Corresponding States method and fluid-specific correlations. The diffusion coefficient  $D_i$  of the  $i$ -th species into the rest of mixture is modelled according to the Hirschfelder and Curtiss expression, where the required binary diffusion coefficient is calculated by means of kinetic theory. The thermo-diffusion coefficient  $\mathcal{D}_i^T$  is estimated by means of the EGLIB routines. The combustion of hydrogen and oxygen is modelled by means of the finite-rate detailed chemical mechanism derived by Boivin from the San Diego mechanism, furtherly simplified by taking the high-pressure limit of the falloff reactions (due to the high pressure of the test case) [7]. The mechanism accounts for 8 species and 12 elementary reactions. Unclosed turbulent combustion subgrid terms of the filtered compressible Navier-Stokes equations are modelled through the Smagorinsky model and the authors' LTSM (Localised Turbulent Scale Model) turbulent combustion model.

The numerical simulations are performed by means of the in-house parallel code *HeaRT* and ENEA's supercomputing facility *CRESCO* [6]. The *HeaRT* code solves the compressible Navier-Stokes equations discretised through staggered finite difference schemes. A second-order accurate centered scheme is adopted for diffusive fluxes. Convective terms are discretized through the *AUSM<sup>+</sup>-up* method [3] coupled with a second-order accurate interpolation with a TVD, linear preserving limiter for non-uniform grids [1] to reduce spurious oscillations. The low-storage third-order accurate Runge-Kutta method of Shu-Osher is used for time integration. Non-reflecting boundary conditions are implemented at open boundaries in their extended form to take into account the effect of variable transport properties, local heat release and real gas effects [5]. A synthetic turbulence generator is adopted at flow inlets: this is especially important for the  $H_2$  jet to correctly capture its spreading angle and to avoid a laminar-to-turbulence transition region not experienced in the experiment.

## 4 Results

The combustor in the simulation is initially filled in with a mixture of hot products coming from a stoichiometric  $H_2/O_2$  premixed flame calculation at 60 bar. The cold reactants are injected and the mixing process is simulated in a two-dimensional axi-symmetric framework. Then, ignition is achieved by locating a spark in the mixing layer between the two reactants at  $\sim 1.3$  mm from the tip of the injector. The diffusion flame is developed in a 2D framework later copied and rotated to generate the initial three-dimensional field for the final LES simulation. The oxygen flows into the combustion chamber through an injector with a tapered final section, where the flow exhibits separation nearly  $\sim 3$  mm before the exit. The resulting recirculation zone captures some hydrogen and transports it within the tapered injector. Due to this recirculation, a quite stable flame establishes in it if the walls

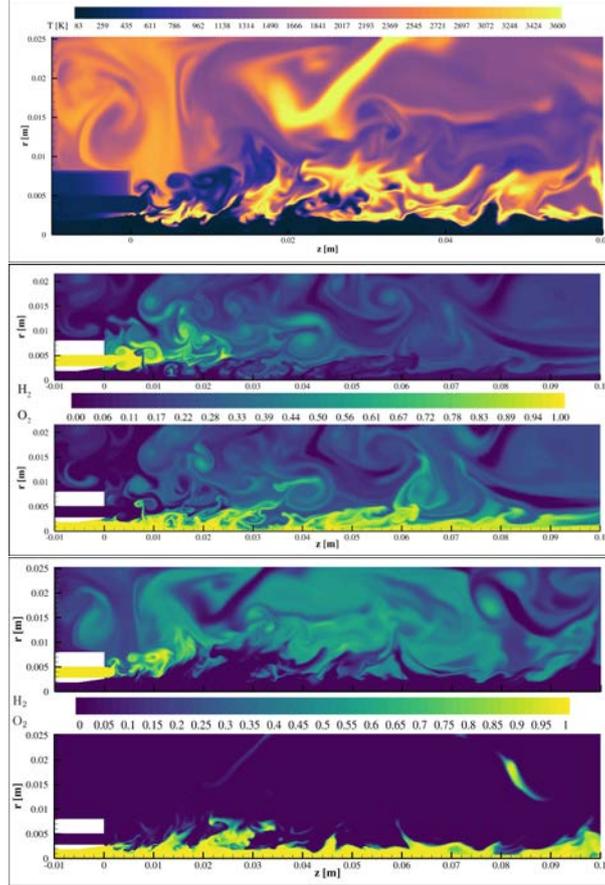


Figure 2: Instantaneous snapshots of the non-reacting (top) and reacting (bottom)  $H_2$  and  $O_2$  mass fractions.

are considered adiabatic. However, a more accurate solution was achieved by solving the heat transfer into the solid walls of the coaxial injector: in this case, the flame does not develop inside the  $O_2$  injector. It was observed that after  $\sim 15$  s from ignition the average temperature of the tip region of the  $O_2$  injector reaches 1569 K, explaining the maximum duration of the experiment reported in [4]. Temperature distribution in the solid zones after  $\sim 13$  s from ignition was assumed as initial distribution for the three-dimensional LES simulation.

Figure 2 shows an instantaneous field of hydrogen and oxygen mass fractions. The oxygen interface appears weakly turbulent up to  $\sim 2$  mm from the jet exit, in agreement with the visual transition length of the experiment, although exhibiting periodic growing of a single high-momentum ligament protruding into the surrounding  $H_2$  jet. Then, the high-density liquid  $O_2$  jet largely and slowly penetrates the combustion chamber along its axis, showing an interface richer of ligaments (more sharply defined in non-reacting condition) promoting intense turbulent mixing, as commonly exhibited by transcritical jets. When the flame is not already established, some ligaments are even able to pass through the surrounding hydrogen stream, thus reaching the external region of the combustor and potentially allow combustion of the outer hydrogen jet surface. Some oxygen pockets are produced and transported downstream within the the flame brush. The faster hydrogen jet exhibits high-frequency ( $\sim 31.5$  kHz) vortex shedding as it is injected into the chamber, producing an inner ( $O_2$  side) and an outer shear-layer. The coherent structures released in the outer shear-layer have the size of the thickness of the external wall of the  $H_2$  pipe ( $\sim 3$  mm); they grow moving downstream and quickly coalesce, evolving

in developed turbulence and feeding the recirculation regions of the combustor generated by the impingement of the hydrogen jet on the chamber walls. Such recirculations bring hot products back to the inlet side. The coherent structures released in the inner shear-layer are smaller, of the size of the tip of the O<sub>2</sub> injector, grow less and live shorter being damped by the flame. It is observed that the velocity fluctuations in the H<sub>2</sub> jet region in the combustion chamber are largely enhanced with respect to the inlet turbulence level of 5%, reaching levels as high as 70% in the shear-layers and as 40% along the central path of the H<sub>2</sub> stream.

Figure 2 also reports an instantaneous snapshot of the temperature field. The flame is anchored at the tip of the O<sub>2</sub> injector. The vortices shed at high frequency from the H<sub>2</sub> inner wall produce intense mixing of fresh reactants with hot products, providing a dynamic flame anchoring. At the same time, the synergic action of their flame stretching and of cold high-momentum oxygen ligaments can produce localized extinctions. The flame develops along the stoichiometric mixture fraction in the inner shear-layer, close but external to the maximum density gradient iso-surface. It is initially largely stretched and thin. Then, a flame brush develops. While rapidly expanding, largely wrinkled by coherent structures and turbulence with wavelengths increasing with the axial distance, the flame pushes the H<sub>2</sub> jet away from the axis. When the flame starts to interact with the main recirculation zones of the combustor, its expansion is enhanced.

Ignition of the flame is due to a stable hot spot established at the tip of the O<sub>2</sub> injector close to which the flame dynamics is complex. From this small reacting region, a primary flame develops around the stoichiometric line slowly creeping along the liquid oxygen interface; then, this flame detaches from the tip and is blown out, but another flame from the tip hot spot enters into the O<sub>2</sub> injector for a short length, before being transported downstream and generate another primary flame. At the same time of the primary flame initial development, part of the flame from the tip of the injector periodically enters into the H<sub>2</sub> pipe along its inner wall, due to the vortex shedding in the inner shear-layer. Then, it is released with the vortex at  $\sim 31.5$  kHz, rapidly transported downstream as a second richer flame within the fast H<sub>2</sub> stream and quickly extinguished due to intense stretching. The primary flame is blown out every 8 releases of the secondary flame.

## References

- [1] M. Berger, M.J. Aftosmis, and S.M. Murman. Analysis of slope limiters on irregular grids. In *43rd AIAA Aerospace Sciences Meeting*, number AIAA 2005-0490, NAS Technical Report NAS-05-007, 10-13 January 2005. Reno, Nevada.
- [2] K.G. Harstad, R.S. Miller, and J. Bellan. Efficient high-pressure state equations. *American Institute of Chemical Engineers Journal*, 43(6):1605–1610, 1997.
- [3] M.S. Liou. A sequel to AUSM, part II: AUSM<sup>+</sup>-up for All Speeds. *J. of Computational Physics*, 214:137–170, 2006.
- [4] Habiballah M., Orain M., Grisch F., Vingert L., and Gicquel P. Experimental studies of high-pressure cryogenic flames on the MASCOTTE facility. *Combustion Science and Technology*, 178:101–128, 2006.
- [5] Nora Okong'o and Josette Bellan. Consistent Boundary Conditions for Multicomponent Real Gas Mixtures Based on Characteristic Waves. *Journal of Computational Physics*, 176(2):330–344, mar 2002.
- [6] G. et al. Ponti. The role of medium size facilities in the HPC ecosystem: the case of the new

- CRESCO4 cluster integrated in the ENEAGRID infrastructure. In *Proceedings of the 2014 International Conference on High Performance Computing and Simulation*, volume HPCS 2014, pages 1030–1033, 2014.
- [7] A. Ruiz. *Unsteady Numerical Simulations of Transcritical Turbulent Combustion in Liquid Rocket Engines*. Ed megep: Énergétique et transferts - unité de recherche: Cerfacs, Institut National Polytechnique de Toulouse (INP Toulouse), Toulouse, France, February 2012.
- [8] L. Vingert, M. Habiballah, P. Vuillermoz, and S. Zurbach. Mascotte, a test facility for cryogenic research at high pressure. In *51st International Astronautical Congress*. AIAA, 2000. Rio de Janeiro, Brazil.
- [9] D.S.H. Wong and S.I. Sandler. A theoretically correct mixing rule for cubic equations of state. *American Institute of Chemical Engineers Journal*, 38(x):671–680, May 1992.

# INTEGRATIVE POLYMER MODELLING AND NUMERICAL SIMULATIONS TO PREDICT THE COMPLEX 3D SPATIAL ORGANIZATION OF CHROMOSOMES

Mattia Conte<sup>1\*</sup>, Andrea Esposito<sup>1</sup>, Luca Fiorillo<sup>1</sup>,  
Andrea Maria Chiariello<sup>1</sup>, Alex Abraham<sup>1</sup>, and Simona Bianco<sup>1\*</sup>

<sup>1</sup> *Dipartimento di Fisica, Università di Napoli Federico II, and INFN Napoli, Complesso Universitario di Monte Sant'Angelo, 80126, Naples, Italy*

*\*Corresponding authors. Email: matconte@na.infn.it, biancos@na.infn.it*

**ABSTRACT.** Innovative technologies and novel protocols from molecular biology are providing many insights on the three-dimensional (3D) spatial organization of chromosomes within the eukaryotic cell nucleus. Such experiments, based e.g. on conformation capture techniques, have led to the discovery of topological structural features of chromosomes, returning a picture where the genome is folded in complex conformations in 3D space so to bring in close spatial proximity genes and their regulators. However, the molecular mechanisms leading to precise and cell-specific chromosome conformations are not yet understood. Quantitative theoretical models from physics can thus be essential to dissect the mechanical rules of genome folding starting from first principles. Here, we review an integrative approach for chromosome modelling that combines polymer physics, machine learning strategies and massive molecular dynamics simulations. In particular, we show that our model, validated against recent super-resolution imaging experiments, is capable to predict the 3D structure of specific genomic regions at the single DNA molecule level. Efficient algorithms, parallel running strategies, and tremendous computational efforts are fundamental to achieve such deep model accuracy and to provide increasingly refined descriptions of chromosome large-scale 3D organization.

## 1. Introduction

In the last decade, diverse technologies from molecular biology, such as Hi-C [1], GAM [2] or SPRITE [3], have revealed that chromosomes are folded in the cell nucleus in complex not-random three-dimensional (3D) conformations. Hi-C methods, for instance, revealed in mammalian cells large, megabase-sized local chromosome domains with enriched inner interaction levels, called Topologically Associating Domains (TADs), as a key structural feature of genome organization [4, 5]. Yet, different TADs can also interact giving rise to higher order chromosome structures, forming a hierarchy of domains-within-domains (metaTADs) spanning across genomic scales up to the range of entire chromosomes [6]. Such an organization has been shown to be essential for transcriptional regulation, e.g. by bringing in close spatial proximity regulatory elements and their target genes [7], and its disruption has been linked to disease [8, 9]. However, the molecular and physical mechanisms driving genome folding are still largely mysterious. Quantitative models starting from first-principled approaches and based on massive numerical simulations are thus becoming an essential tool to interpret the huge amount of genomic structural data and unveil the mechanisms underlying the 3D spatial organization of real chromosomes [10–21]. Here, we focus on the textbook scenario where

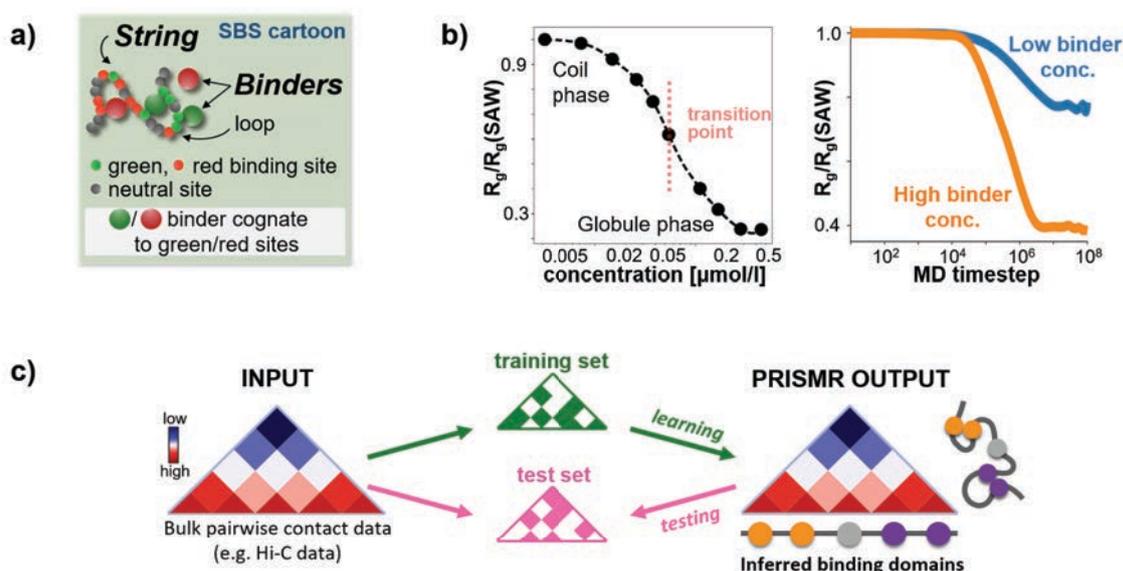
diffusing cognate binding factors mediate the contact between distal DNA binding sites, as described by the Strings and Binders Switch (SBS) chromosome polymer model [16, 17, 21]. As general pipeline, detailed in the next Section, the SBS polymer model is trained on available bulk Hi-C data of the genomic region of interest by using a machine learning procedure [12]. Once inferred the optimal model for the chromosome region to study, we derive by molecular dynamics (MD) simulations a thermodynamics ensemble of single-molecule 3D conformations for that region [22]. Parallel strategies at this stage, e.g. by domain decomposition, are fundamental to efficiently manage MD simulations, since the running times increase with the size of the input data, deeply linked to the complexity of the simulated system [23]. In the last Section, as case study, we use the SBS model to predict the ensemble of single-molecule 3D structures of a 2.5 Mb DNA region in human HCT116 cancer cells and we validate the theoretical predictions against the most recent super-resolution imaging dataset available to date in the literature [13, 24].

## 2. The computational pipeline of the SBS polymer model

The SBS model is a coarse-grained polymer approach where a genomic region is represented as a Self-Avoiding Walk (SAW) polymer chain having different specific binding sites for cognate diffusing molecular binders. In our notation, each different type of site is schematized by using a different color and a specific attractive homotypic interaction is only permitted between same coloured polymer sites and cognate binders (Fig. 1a). A standard Langevin motion, numerically solved by an optimized Verlet integrator, produces the system dynamics with interaction potentials taken by classical polymer physics studies [21, 22, 25]. As dictated by polymer physics [26, 27], our polymer system undergoes a phase transition from a coil, i.e. randomly open, to a phase separated compact globular state as the binder concentration (or affinity strength) grows above a threshold point (Fig. 1b, left). MD simulations are performed using the free available LAMMPS software [28], which employs the Message Passing Interface (MPI) protocol and achieves parallelization by domain decomposition, namely by partitioning the simulation box into sub-domains assigned to different processors. Starting from an initial SAW conformation, the folding of the polymer chain into a final compact globular conformation requires many hours of running time, depending on the polymer length. In MD units, the time required to reach the thermodynamics transition point is roughly  $10^5$  timesteps for a polymer made of nearly 800 sites (Fig. 1b, right). This number can easily grow upon increasing the system complexity, such as the total number of simulated particles, thus making crucial the role played by High-Performance-Computing (HPC) resources. Indeed, the computing resource for the model simulations have been provided by CRESCO/ENEAGRID High Performance Computing infrastructure and its staff [29]. All details about the model can be found in recent reviews [30–33].

To derive the SBS model of a real genomic region or a whole chromosome, i.e. the optimal number of distinct types of binding sites and their arrangement along the polymer chain, we developed a machine learning procedure based on the PRISMR algorithm [12]. The algorithm takes as input a bulk pairwise contact matrix, e.g., Hi-C or GAM, of the investigated genomic region and returns as output the polymer model that best recapitulates the input data. PRISMR is based on a Simulated Annealing Monte Carlo that minimizes a cost function defined as the mean squared error function between the input and model-derived contact matrix. The cost function is also corrected by a Bayesian term in order to avoid overfitting. In our latest improved implementation [13], by adopting an approach standard to supervised learning, we split Hi-C data into a training set and a complementary test set. Our optimization procedure is then run on the training dataset to infer the

SBS polymer model best describing the input contact data, whereas the test set is used to test the model predicted contact frequencies (Fig. 1c).



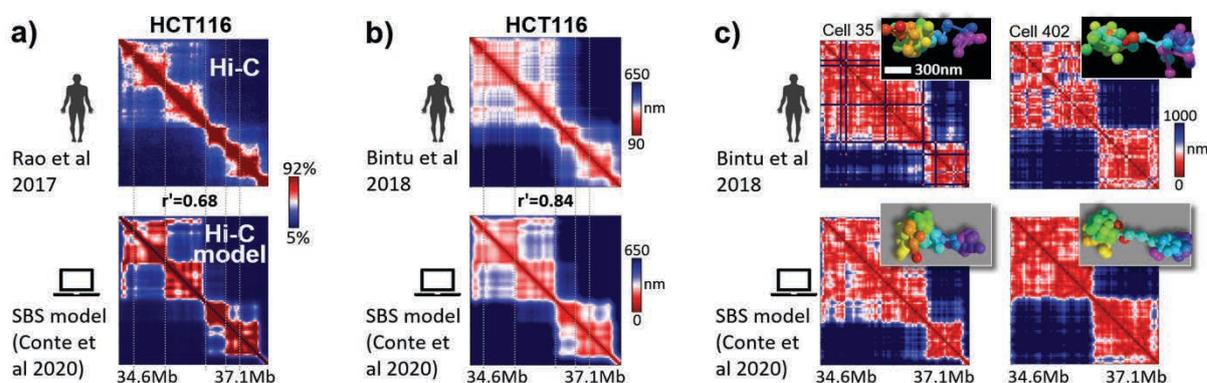
**Fig.1: The SBS polymer model of chromosomes.** **a)** Cartoon of the Strings and Binders Switch (SBS) model. Folding is driven by homotypic interactions between polymer binding sites and cognate diffusing binders. **b)** The gyration radius,  $R_g$ , is a classical order parameter used to monitor the polymer dynamics [26]. Left: Upon increasing the binder concentration,  $R_g$  sharply decreases signalling the coil-to-globule phase transition. Typically the critical concentration falls in the fractions of  $\mu\text{mol/l}$  range [21], values compatible with transcription factor concentrations. Right: Normalized gyration radius against MD timestep. The MD time required to reach the transition point depends on polymer complexity. In this case, it is roughly  $10^5$  MD timesteps for a polymer made of nearly 800 sites [13]. **c)** Schematic outline of the PRISM optimization algorithm [12, 13]. All items are adapted from [13].

### 3. SBS model single-molecule 3D conformations are validated against independent high-resolution imaging data

The SBS modelling approach has been successfully employed to explain the 3D organization of several genomic regions of interest in human and murine cells, such as the Sox9 [34, 35], EPHA4 [12], HoxD [36], Shh [37],  $\alpha$ - and  $\beta$ -globin [38], or Nkx2-5 [39] regions. It has also allowed investigating the complex architectural changes occurring during cell differentiation and spread light on their deep connection with transcription [36, 38, 40]. SBS model predictions have been widely validated against various independent experiments; e.g., notably, they were shown to accurately capture the effect of pathogenic genomic mutations, such as deletions or duplications, on architecture [12, 36, 41].

The SBS model has been used to understand the mechanisms of formation of architectural patterns also at the single-molecule level. For each binder concentration, in fact, a given SBS polymer model can fold in a variety of single-molecules 3D conformations, not just in a single, unique structure. We

have recently shown [13] that our model predicted single-molecules 3D structures are a bona fide representation of the actual chromosome conformations in single-cells as emerging from recent high-resolution imaging experiments [24] in many genomic regions and cell lines. We review here as case study the modeling of a 2.5 Mb DNA region (chr21:34.6–37.1 Mb) in human HCT116 cells. To infer the different types of the putative binding sites of the SBS polymer model of the studied genomic region, we employed the PRISMR algorithm taking as input only bulk Hi-C data from [42]. After setting the binder affinities and concentration [13, 21], we derived by MD simulations a thermodynamic ensemble of single-molecule 3D conformations of the model of the investigated DNA region. To check that the model derived 3D structures recapitulate the input bulk Hi-C data, we computed the model pairwise contact matrix and found high correlations between model and Hi-C contact data (Fig. 2a, genomic distance-corrected Pearson correlation coefficient,  $r' = 0.68$ ). As a first validation of our model, we computed the model median distance matrix of the studied region, namely the median of the predicted single-molecule distance matrices, and compared it against independent super-resolution imaging data from [24]. We found a distance-corrected Pearson correlation  $r' = 0.84$ , even higher than the one with Hi-C data, showing that our polymer model is consistent with both Hi-C and independent bulk imaging data (Fig. 2b). Next, to demonstrate that the 3D structures predicted by our model are similar to those observed in real single-cell experiments, we performed an all-against-all comparison between model predicted and experimental 3D structures. Using a criterion, fully detailed in [13], that finds the optimal rotation between two 3D structures to minimize the mean squared deviation (RMSD) of their coordinates, each imaged 3D structure was univocally associated to a corresponding model 3D conformation by searching for the minimum RMSD (Fig. 2c). By that RMSD mapping, we showed that single-molecules from the model are statistically indistinguishable from experimental single-cell structures [13], providing evidence in the studied genomic region that chromosome folding is explained at the single-cell level by a thermodynamics mechanism of globule phase separation.



**Fig.2: SBS predicted single-molecules are a bona fide representation of chromosome conformations in single-cells.** **a)** Comparison between Hi-C [42] (top) and model derived (bottom) pairwise contact matrices of the studied 2.5 Mb DNA region (chr21:34.6–37.1 Mb) in human HCT116 cells. **b)** Comparison between imaged [24] (top) and model predicted (bottom) median distance matrix of the studied 2.5 Mb DNA region in human HCT116 cells. **c)** Examples of experiment-model optimal matches based on RMSD criterion. Imaged single-cell distance matrices and relative 3D structures of the studied HCT116 region from [24] (top) and corresponding model single-molecule predictions. All items are adapted from [13].

## 4. Conclusions

We reviewed here a schematic polymer model for the large-scale genome spatial organization where specific contacts between distal DNA sites are established by diffusing Brownian molecular binders [16, 17]. We focused on the modelling of a 2.5 Mb DNA region in human HCT116 cells [13], showing that single-molecule predictions from the model are consistently validated against recent super-resolution imaging data [24], thus providing insights into the molecular mechanisms of genome folding at the single-cell level. To ensure such statistical robustness to model results, the simulated ensemble should contain at least thousand independent polymer conformations. Considering that the execution times increase as a power-law with the size of the input data [23], HPC is crucial, as well as a fast and efficient MD implementation based on appropriate parallel paradigms. Improving the efficiency and the fine details of the employed computational tools is a fundamental step to develop increasingly precise models of chromosome nuclear 3D organization.

## References

1. Lieberman-Aiden, E., Van Berkum, N.L., Williams, L., Imakaev, M., Ragozy, T., Telling, A., Amit, I., Lajoie, B.R., Sabo, P.J., Dorschner, M.O., Sandstrom, R., Bernstein, B., Bender, M.A., Groudine, M., Gnirke, A., Stamatoyannopoulos, J., Mirny, L.A., Lander, E.S., Dekker, J.: Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* (80-. ). 326, 289–293 (2009). <https://doi.org/10.1126/science.1181369>.
2. Beagrie, R.A., Scialdone, A., Schueler, M., Kraemer, D.C.A., Chotalia, M., Xie, S.Q., Barbieri, M., De Santiago, I., Lavitas, L.M., Branco, M.R., Fraser, J., Dostie, J., Game, L., Dillon, N., Edwards, P.A.W., Nicodemi, M., Pombo, A.: Complex multi-enhancer contacts captured by genome architecture mapping. *Nature*. (2017). <https://doi.org/10.1038/nature21411>.
3. Quinodoz, S.A., Ollikainen, N., Tabak, B., Palla, A., Schmidt, J.M., Detmar, E., Lai, M.M., Shishkin, A.A., Bhat, P., Takei, Y., Trinh, V., Aznauryan, E., Russell, P., Cheng, C., Jovanovic, M., Chow, A., Cai, L., McDonel, P., Garber, M., Guttman, M.: Higher-Order Inter-chromosomal Hubs Shape 3D Genome Organization in the Nucleus. *Cell*. 174, 744–757.e24 (2018). <https://doi.org/10.1016/j.cell.2018.05.024>.
4. Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., Ren, B.: Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature*. 485, 376–380 (2012). <https://doi.org/10.1038/nature11082>.
5. Nora, E.P., Lajoie, B.R., Schulz, E.G., Giorgetti, L., Okamoto, I., Servant, N., Piolot, T., Van Berkum, N.L., Meisig, J., Sedat, J., Gribnau, J., Barillot, E., Blüthgen, N., Dekker, J., Heard, E.: Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature*. 485, 381–385 (2012). <https://doi.org/10.1038/nature11049>.
6. Fraser, J., Ferrai, C., Chiariello, A.M., Schueler, M., Rito, T., Laudanno, G., Barbieri, M., Moore, B.L., Kraemer, D.C., Aitken, S., Xie, S.Q., Morris, K.J., Itoh, M., Kawaji, H., Jaeger, I., Hayashizaki, Y., Carninci, P., Forrest, A.R., Sempole, C.A., Dostie, J., Pombo, A., Nicodemi, M.: Hierarchical folding and reorganization of chromosomes are linked to transcriptional changes in cellular differentiation. *Mol. Syst. Biol.* 11, 852–852 (2015). <https://doi.org/10.15252/msb.20156492>.
7. Dekker, J., Mirny, L.: The 3D Genome as Moderator of Chromosomal Communication. *Cell*. 164, P1110–1121 (2016). <https://doi.org/10.1016/j.cell.2016.02.007>.
8. Lupiáñez, D.G., Kraft, K., Heinrich, V., Krawitz, P., Brancati, F., Klopocki, E., Horn, D., Kayserili, H., Opitz, J.M., Laxova, R., Santos-Simarro, F., Gilbert-Dussardier, B., Wittler, L., Borschiwer, M., Haas, S.A., Osterwalder, M., Franke, M., Timmermann, B., Hecht, J.,

- Spielmann, M., Visel, A., Mundlos, S.: Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. *Cell*. 161, 1012–1025 (2015). <https://doi.org/10.1016/j.cell.2015.04.004>.
9. Hnisz, D., Weintraub, A.S., Day, D.S., Valton, A.L., Bak, R.O., Li, C.H., Goldmann, J., Lajoie, B.R., Fan, Z.P., Sigova, A.A., Reddy, J., Borges-Rivera, D., Lee, T.I., Jaenisch, R., Porteus, M.H., Dekker, J., Young, R.A.: Activation of proto-oncogenes by disruption of chromosome neighborhoods. *Science* (80- ). (2016). <https://doi.org/10.1126/science.aad9024>.
  10. Sanborn, A.L., Rao, S.S.P.P., Huang, S.-C.C., Durand, N.C., Huntley, M.H., Jewett, A.I., Bochkov, I.D., Chinnappan, D., Cutkosky, A., Li, J., Geeting, K.P., Gnirke, A., Melnikov, A., McKenna, D., Stamenova, E.K., Lander, E.S., Aiden, E.L.: Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. *Proc. Natl. Acad. Sci.* 112, E6456–E6465 (2015). <https://doi.org/10.1073/pnas.1518552112>.
  11. Fudenberg, G., Imakaev, M., Lu, C., Goloborodko, A., Abdennur, N., Mirny, L.A.: Formation of Chromosomal Domains by Loop Extrusion. *Cell Rep.* 15, 2038–2049 (2016). <https://doi.org/10.1016/j.celrep.2016.04.085>.
  12. Bianco, S., Lupiáñez, D.G., Chiariello, A.M., Annunziatella, C., Kraft, K., Schöpflin, R., Wittler, L., Andrey, G., Vingron, M., Pombo, A., Mundlos, S., Nicodemi, M.: Polymer physics predicts the effects of structural variants on chromatin architecture. *Nat. Genet.* (2018). <https://doi.org/10.1038/s41588-018-0098-8>.
  13. Conte, M., Fiorillo, L., Bianco, S., Chiariello, A.M., Esposito, A., Nicodemi, M.: Polymer physics indicates chromatin folding variability across single-cells results from state degeneracy in phase separation. *Nat. Commun.* (2020). <https://doi.org/10.1038/s41467-020-17141-4>.
  14. Brackley, C.A., Johnson, J., Michieletto, D., Morozov, A.N., Nicodemi, M., Cook, P.R., Marenduzzo, D.: Nonequilibrium Chromosome Looping via Molecular Slip Links. *Phys. Rev. Lett.* (2017). <https://doi.org/10.1103/PhysRevLett.119.138101>.
  15. Buckle, A., Brackley, C.A., Boyle, S., Marenduzzo, D., Gilbert, N.: Polymer Simulations of Heteromorphic Chromatin Predict the 3D Folding of Complex Genomic Loci. *Mol. Cell.* 72, 786–797.e11 (2018). <https://doi.org/10.1016/j.molcel.2018.09.016>.
  16. Nicodemi, M., Prisco, A.: Thermodynamic pathways to genome spatial organization in the cell nucleus. *Biophys. J.* 96, 2168–2177 (2009). <https://doi.org/10.1016/j.bpj.2008.12.3919>.
  17. Barbieri, M., Chotalia, M., Fraser, J., Lavitas, L.-M., Dostie, J., Pombo, A., Nicodemi, M.: Complexity of chromatin folding is captured by the strings and binders switch model. *Proc. Natl. Acad. Sci.* (2012). <https://doi.org/10.1073/pnas.1204799109>.
  18. Barbieri, M., Xie, S.Q.S.Q., Torlai Triglia, E., Chiariello, A.M.A.M., Bianco, S., De Santiago, I., Branco, M.R.M.R., Rueda, D., Nicodemi, M., Pombo, A.: Active and poised promoter states drive folding of the extended HoxB locus in mouse embryonic stem cells. *Nat. Struct. Mol. Biol.* 24, 515–524 (2017). <https://doi.org/10.1038/nsmb.3402>.
  19. Jost, D., Carrivain, P., Cavalli, G., Vaillant, C.: Modeling epigenome folding: Formation and dynamics of topologically associated chromatin domains. *Nucleic Acids Res.* 42, 9553–9561 (2014). <https://doi.org/10.1093/nar/gku698>.
  20. Bohn, M., Heermann, D.W.: Diffusion-driven looping provides a consistent provides a consistent framework for chromatin organization. *PLoS One.* 5, e12218 (2010). <https://doi.org/10.1371/journal.pone.0012218>.
  21. Chiariello, A.M., Annunziatella, C., Bianco, S., Esposito, A., Nicodemi, M.: Polymer physics of chromosome large-scale 3D organisation. *Sci. Rep.* (2016). <https://doi.org/10.1038/srep29775>.
  22. Annunziatella, C., Chiariello, A.M.A.M., Esposito, A., Bianco, S., Fiorillo, L., Nicodemi, M.: Molecular Dynamics simulations of the Strings and Binders Switch model of chromatin. (2018). <https://doi.org/10.1016/j.ymeth.2018.02.024>.

23. Conte, M., Esposito, A., Fiorillo, L., Campanile, R., Annunziatella, C., Corrado, A., Chiariello, M.G., Bianco, S., Chiariello, A.M.: Efficient computational implementation of polymer physics models to explore chromatin structure. *Int. J. Parallel, Emergent Distrib. Syst.* (2019). <https://doi.org/10.1080/17445760.2019.1643020>.
24. Bintu, B., Mateo, L.J., Su, J.-H., Sinnott-Armstrong, N.A., Parker, M., Kinrot, S., Yamaya, K., Boettiger, A.N., Zhuang, X.: Super-resolution chromatin tracing reveals domains and cooperative interactions in single cells. *Science* (80-. ). 362, eaau1783 (2018). <https://doi.org/10.1126/science.aau1783>.
25. Kremer, K., Grest, G.S.: Dynamics of entangled linear polymer melts: A molecular-dynamics simulation. *J. Chem. Phys.* 92, 5057–5086 (1990). <https://doi.org/10.1063/1.458541>.
26. De Gennes, P.G.: *Scaling concepts in polymer physics*. Cornell university press. Ithaca N.Y., (1979). [https://doi.org/10.1163/\\_q3\\_SIM\\_00374](https://doi.org/10.1163/_q3_SIM_00374).
27. Chiariello, A.M., Corberi, F., Salerno, M.: The Interplay between Phase Separation and Gene-Enhancer Communication: A Theoretical Study. *Biophys. J.* 119, 873–883 (2020). <https://doi.org/10.1016/j.bpj.2020.07.007>.
28. Plimpton, S.: Fast parallel algorithms for short-range molecular dynamics. *J. Comput. Phys.* 117, 1–19 (1995). <https://doi.org/10.1006/jcph.1995.1039>.
29. Iannone, F. et al: CRESCO ENEA HPC clusters: a working example of a multifabric GPFS Spectrum Scale layout. 2019 International Conference on High Performance Computing & Simulation (HPCS), Dublin, Ireland, 2019, pp. 1051-1052, doi: 10.1109/HPCS48598.2019.9188135.
30. Annunziatella, C., Chiariello, A.M., Esposito, A., Bianco, S., Fiorillo, L., Nicodemi, M.: Molecular Dynamics simulations of the Strings and Binders Switch model of chromatin. *Methods.* 142, (2018). <https://doi.org/10.1016/j.ymeth.2018.02.024>.
31. Esposito, A., Chiariello, A.M., Conte, M., Fiorillo, L., Musella, F., Sciarretta, R., Bianco, S.: Higher-order Chromosome Structures Investigated by Polymer Physics in Cellular Morphogenesis and Differentiation, (2020). <https://doi.org/10.1016/j.jmb.2019.12.017>.
32. Fiorillo, L., Bianco, S., Esposito, A., Conte, M., Sciarretta, R., Musella, F., Chiariello, A.M.: A modern challenge of polymer physics: Novel ways to study, interpret, and reconstruct chromatin structure, (2020). <https://doi.org/10.1002/wcms.1454>.
33. Bianco, S., Chiariello, A.M., Conte, M., Esposito, A., Fiorillo, L., Musella, F., Nicodemi, M.: Computational approaches from polymer physics to investigate chromatin folding, (2020). <https://doi.org/10.1016/j.ceb.2020.01.002>.
34. Chiariello, A.M., Annunziatella, C., Bianco, S., Esposito, A., Nicodemi, M.: Polymer physics of chromosome large-scale 3D organisation. *Sci. Rep.* 6, (2016). <https://doi.org/10.1038/srep29775>.
35. Fiorillo, L., Bianco, S., Chiariello, A.M., Barbieri, M., Esposito, A., Annunziatella, C., Conte, M., Corrado, A., Prisco, A., Pombo, A., Nicodemi, M.: Inference of chromosome 3D structures from GAM data by a physics computational approach. *Methods.* (2019). <https://doi.org/10.1016/j.ymeth.2019.09.018>.
36. Bianco, S., Annunziatella, C., Andrey, G., Chiariello, A.M., Esposito, A., Fiorillo, L., Prisco, A., Conte, M., Campanile, R., Nicodemi, M.: Modeling Single-Molecule Conformations of the HoxD Region in Mouse Embryonic Stem and Cortical Neuronal Cells. *Cell Rep.* 28, (2019). <https://doi.org/10.1016/j.celrep.2019.07.013>.
37. Paliou, C., Guckelberger, P., Schöpflin, R., Heinrich, V., Esposito, A., Chiariello, A.M., Bianco, S., Annunziatella, C., Helmuth, J., Haas, S., Jerkovic, I., Brieske, N., Wittler, L., Timmermann, B., Nicodemi, M., Vingron, M., Mundlos, S., Andrey, G.: Prefomed chromatin topology assists transcriptional robustness of Shh during limb development. *Proc. Natl. Acad. Sci. U. S. A.* 116, (2019). <https://doi.org/10.1073/pnas.1900672116>.
38. Chiariello, A.M., Bianco, S., Oudelaar, A.M., Esposito, A., Annunziatella, C., Fiorillo, L.,

- Conte, M., Corrado, A., Prisco, A., Larke, M.S.C., Telenius, J.M., Sciarretta, R., Musella, F., Buckle, V.J., Higgs, D.R., Hughes, J.R., Nicodemi, M.: A Dynamic Folded Hairpin Conformation Is Associated with  $\alpha$ -Globin Activation in Erythroid Cells. *Cell Rep.* (2020). <https://doi.org/10.1016/j.celrep.2020.01.044>.
39. Salamon, I., Serio, S., Bianco, S., Pagiatakis, C., Crasto, S., Chiariello, A.M., Conte, M., Cattaneo, P., Fiorillo, L., Felicetta, A., di Pasquale, E., Kunderfranco, P., Nicodemi, M., Papait, R., Condorelli, G.: Divergent Transcription of the Nkx2-5 Locus Generates Two Enhancer RNAs with Opposing Functions. *iScience.* 23, (2020).
40. Oudelaar, A.M.M., Davies, J.O.J., Hanssen, L.L.P., Telenius, J.M., Schwessinger, R., Liu, Y., Brown, J.M., Downes, D.J., Chiariello, A.M., Bianco, S., Nicodemi, M., Buckle, V.J., Dekker, J., Higgs, D.R., Hughes, J.R.: Single-allele chromatin interactions identify regulatory hubs in dynamic compartmentalized domains. *Nat. Genet.* 50, (2018). <https://doi.org/10.1038/s41588-018-0253-2>.
41. Kragesteen, B.K., Spielmann, M., Paliou, C., Heinrich, V., Schöpflin, R., Esposito, A., Annunziatella, C., Bianco, S., Chiariello, A.M., Jerković, I., Harabula, I., Guckelberger, P., Pechstein, M., Wittler, L., Chan, W.L.W.-L., Franke, M., Lupiáñez, D.G.D., Kraft, K., Timmermann, B., Vingron, M., Visel, A., Nicodemi, M., Mundlos, S., Andrey, G.: Dynamic 3D chromatin architecture contributes to enhancer specificity and limb morphogenesis. *Nat. Genet.* 50, 1463–1473 (2018). <https://doi.org/10.1038/s41588-018-0221-x>.
42. Rao, S.S.P., Huang, S.C., Glenn St Hilaire, B., Engreitz, J.M., Perez, E.M., Kieffer-Kwon, K.R., Sanborn, A.L., Johnstone, S.E., Bascom, G.D., Bochkov, I.D., Huang, X., Shamim, M.S., Shin, J., Turner, D., Ye, Z., Omer, A.D., Robinson, J.T., Schlick, T., Bernstein, B.E., Casellas, R., Lander, E.S., Aiden, E.L.: Cohesin Loss Eliminates All Loop Domains. *Cell.* 171, 305-320.e24 (2017). <https://doi.org/10.1016/j.cell.2017.09.026>.

# AB INITIO STUDY OF NA INSERTION AT TiO<sub>2</sub> ANATASE SURFACES AS NEGATIVE ELECTRODE FOR SODIUM BATTERIES

Arianna Massaro<sup>1</sup>, Ana B. Muñoz García,<sup>2</sup> Michele Pavone<sup>1\*</sup>

<sup>1</sup>University of Naples “Federico II”, Department of Chemical Sciences, Comp. Univ. Monte Sant’Angelo Via Cintia 21, 80126, Naples, Italy

<sup>2</sup>University of Naples “Federico II”, Department of Physics “Ettore Pancini”, Comp. Univ. Monte Sant’Angelo Via Cintia 21, 80126, Naples, Italy

**ABSTRACT.** We provide a first-principles investigation on Na adsorption and insertion at (101), (100) and (001) surfaces of TiO<sub>2</sub> anatase as promising anode material for Na-ion batteries (NIB). The simple coverage of Na atoms on anatase surfaces is not a reliable model for representing ion-intercalation, whereas the Na migration into the lattice subsurface can explain the experimental evidence. We show that (001) termination is the most effective toward Na migration due to the peculiar structural features of the surface lattice window.

## 1 Introduction

The research carried out at University of Naples “Federico II” focused on the ab initio simulation of Na insertion at three different TiO<sub>2</sub> anatase surfaces, *i.e.* (101), (100) and (001). Ion intercalation represents one of the main processes in a NIB device, thus understanding the detailed mechanism is crucial to boost the development of electrode materials that are expected to be good Na-hosts. TiO<sub>2</sub> anatase has been proposed as promising anode material for the overall balance of performance, stability and cost. As the exposed crystal facets in different morphologies of nanostructured anatase can affect the electrochemical performances, here we report a theoretical investigation of Na adsorption and migration through (101), (100) and (001) surface terminations, thus explaining the different activities toward sodiation reported in the literature [1].

## 2.1 Methods and Computational details

Spin-polarized density functional theory (DFT)+U theory was employed to overcome the self-interaction error that affects DFT when applied to transition metal oxides with tightly localized d-electrons ( $U_{eff}=3.3\text{eV}$  for Ti atoms) [2]. We use the Perdew-Burke-Ernzerhof (PBE) exchange-correlation functional, projector-augmented wave (PAW) potentials and plane wave (PW) basis sets, as implemented in the Vienna ab-initio simulation package (VASP) code. A kinetic energy of 600 eV was used to converge the PW basis set. We build up three 6L-slab models containing 24 f.u. (72 atoms) and used  $2\times 2\times 1$  k-points sampling based on the Monkhorst-Pack scheme. Climbing image-nudge elastic band (CI-NEB) has been employed to locate minimum energy paths (MEP) of Na migration and to compute the corresponding barrier heights [3]. The projected density of states (PDOS) are computed via single-point energy calculations at the HSE06 level of theory [4].

## 2.2 Results and Discussion

We first model Na adsorption and insertion by performing the geometry optimization of one Na atom adsorbed on top of the pristine surfaces and then inserted in the subsurface layer, respectively. The corresponding energy variations for each surface have been calculated as follows:

$$E_{ads} = E_{Na_{OUT}} - E_{slab} - \frac{1}{2}E_{Na} \quad (1)$$

$$E_{ins} = E_{Na_{IN}} - E_{slab} - \frac{1}{2}E_{Na} \quad (2)$$

where  $E_{Na(OUT)}$  and  $E_{Na(IN)}$  are the total energies of the optimized slabs with, respectively, one adsorbed and one inserted sodium atom,  $E_{slab}$  is the total energy of the clean surface and  $E_{Na}$  is the total energy of metallic sodium (2-atoms cell), which is taken as reference. We also consider the distortion energy, *i.e.* the energy required to adapt the  $TiO_2$  lattice in the pristine surface structure to the one in the Na adsorbed state, that has been defined as:

$$E_{dist} = E_{slab}^{Na_{OUT}} - E_{slab}^{slab} \quad (3)$$

where  $E_{slab}^{Na(OUT)}$  and  $E_{slab}^{slab}$  are the energies of the slab at the geometry of Na adsorbed state and in its minimum-energy structure, respectively. The results are listed in Table 1.

**Table 1:** Adsorption,  $E_{ads}$ , insertion,  $E_{ins}$ , and distortion,  $E_{dist}$ , energies computed according to Equations 1, 2 and 3, respectively.  $E_{ins}$  are compared to the  $TiO_2$  anatase bulk value.

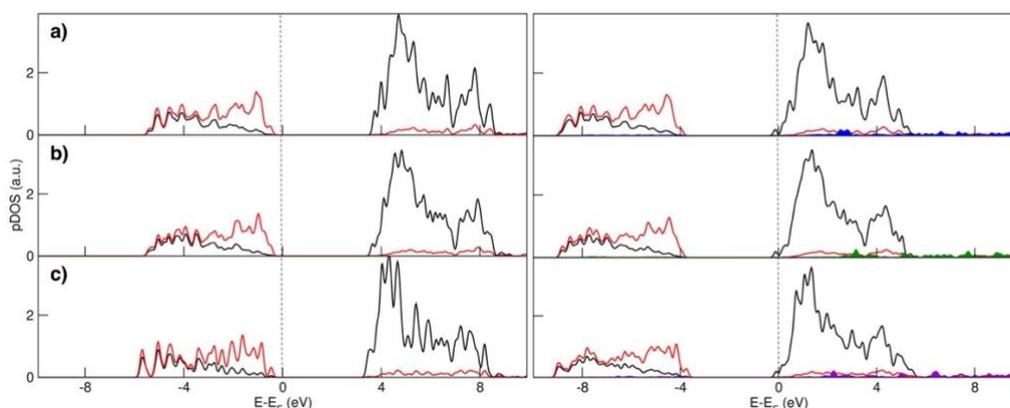
	<b>E<sub>ads</sub> (eV)</b>	<b>E<sub>ins</sub> (eV)</b>	<b>E<sub>dist</sub> (eV)</b>
<i>bulk</i>		0.531	
(101)	-1.120	0.266	0.577
(100)	-1.554	0.460	0.318
(001)	-0.801	0.364	0.726

Insertion energies slightly differ for  $\sim 0.1$  eV among each other and from the bulk reference value, while the large difference in adsorption energy can be explained through distortion effects: the smaller the distortion energy, the more favourable the adsorption process. Electronic structure analysis by PDOS in Figure 1 show that the adsorption of a Na atom leads to the reduction of Ti species and population of states at the edge of the  $TiO_2$  conduction band, which is consistent with the desired n-type conductivity for the NIB electrode. The simple Na coverage (*i.e.* adsorption) does not catch the enhanced efficiency of (001)-terminated nanoparticles (NPs) observed in experiments. For this reason, we provide a further investigation by taking into account the Na migration process. We consider the minimum-energy structures of Na-adsorbed ( $Na_{OUT}$ ) and Na-inserted ( $Na_{IN}$ ) systems as the initial and final equilibrium states, respectively, along the process coordinate, *i.e.* the migration path. We applied the CI-NEB method to compute the MEP between the two states of interest,  $Na_{OUT}$  and  $Na_{IN}$ , respectively. The results are shown in Figure 2. The energy barriers associated to both sodiation and desodiation processes can be determined as:

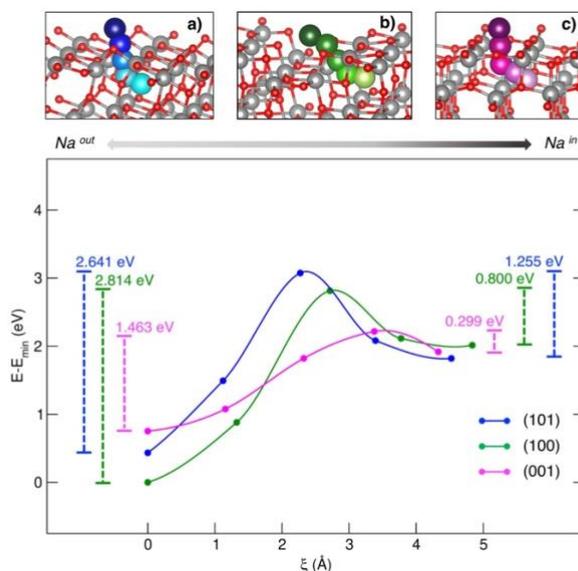
$$E_{migr}^{OUT} = E_{Na_{TS}} - E_{Na_{OUT}} \quad (4)$$

$$E_{migr}^{IN} = E_{Na_{IN}} - E_{Na_{TS}} \quad (5)$$

where  $E_{Na(OUT)}$ ,  $E_{Na(TS)}$  and  $E_{Na(IN)}$  are the total energies of the initial, transition, and final states, respectively, along the MEP.



**Fig.1:** PDOS for a) (101), b) (100) and c) (001) surfaces before (left) and after Na adsorption (right) at the HSE06 level of theory. Colour code: Ti d states (black), O p states (red), Na s states (blue, green and violet respectively for (101), (100) and (001) surfaces).



**Fig.2:** MEP (top) and energetics (bottom) for Na migration through a) (101), b) (100) and c) (001) surfaces. Energy barriers for the in/out migration are computed according to Equation 4 and 5 and are reported to the right- and left-hand sides of the graph, respectively.

In order to rationalize such trend, we focus on the structural features of the different surface terminations. The size of the lattice window at the transition state ( $\text{Na}_{\text{TS}}$ ) provides a direct measure of the steric hindrance that the migrating sodium encounters along its path. The lattice window is the area delimited by the Ti-O lattice and perpendicular to the path that the migrating sodium should cross in order to enter the  $\text{TiO}_2$  crystalline structure. The total area ( $\mathcal{A}$ ) of the lattice window can be calculated for each surface straightforwardly from two Ti-Ti distances. However, this descriptor does not capture the significant difference between the lowest barrier found in (001) and the two higher ones in (100) and (101), as it can be seen from Table 2. Indeed, the limiting factor in the sodium path is related to an effective lattice window delimited by the shortest Ti-Ti distance, thus defining the minimum accessible area,  $a_{\text{min}}$ , which indirectly correlates to the migration barrier heights: smaller lattice windows in (101) and (100) surfaces result in higher barriers, while the (001) surface shows the largest lattice window and, thus, the lowest barrier.

**Table 2:** Total,  $A_{tot}$ , and minimum accessible,  $a_{min}$ , areas of the lattice windows and out migration energy,  $E_{migr}^{OUT}$ , computed according to Equation 4.

	$A_{tot}$ ( $\text{\AA}^2$ )		$a_{min}$ ( $\text{\AA}^2$ )		$E_{migr}^{OUT}$ (eV)
	Na <sub>OUT</sub>	Na <sub>TS</sub>	Na <sub>OUT</sub>	Na <sub>TS</sub>	
(101)	11.4	12.3	8.7	9.6	2.641
(100)	11.4	12.6	8.7	9.9	2.814
(001)	15.2	14.9	15.1	14.8	1.463

### 2.3 Conclusions

The most favourable Na adsorption is found at the (100) termination, thanks to a convenient accommodation on the surface that does not perturb much the crystalline lattice. We find that charge/discharge processes in TiO<sub>2</sub> anatase can effectively occur via Na-insertion mechanism and prove that the (001)-surface is the most active toward Na insertion thanks to its peculiar structural features: the large surface lattice window allows for an easy diffusion of the large Na<sup>+</sup>, thus supporting the experimental observation of high performances with TiO<sub>2</sub> NPs exposing the (001) surface termination in NIBs. Analysis of MEPs demonstrates that the corresponding migration barrier heights depend only on structural modification of the lattice upon Na<sup>+</sup> insertion, thus unveiling the key role of surface lattice windows. The identification of this structural description can be very useful for the rational design of novel high-performing materials with efficient Na uptake mechanism, particularly nanostructured oxide-based materials having promising Na-host characteristics. These results have been gathered in a paper that has been recently published on *Nanoscale Advances* [5]. The computing resources and the related technical support used for this work have been provided by CRESCO/ENEAGRID High Performance Computing infrastructure and its staff; CRESCO/ENEAGRID High Performance Computing infrastructure is funded by ENEA, the Italian National Agency for New Technologies, Energy and Sustainable Economic Development and by Italian and European research programmes. See: <http://www.cresco.enea.it/english> for information [6].

### References

- [1] G. Longoni, *et al.*, Shape-Controlled TiO<sub>2</sub> Nanocrystals for Na-ion Battery Electrodes: The Role of Different Exposed Crystal Facets on the Electrochemical Properties. *Nano Lett.*, **17**, pp 992-1000 (2017).
- [2] V.I. Anisimov, *et al.*, Band theory and Mott Insulators: Hubbard  $U$  instead of Stoner  $I$ . *Phys. Rev. B*, **44**, p 943 (1991).
- [3] G. Henkelman, *et al.*, A Climbing Image Nudged Elastic Band Method for Finding Saddle Points and Minimum Energy Paths. *J. Chem. Phys.*, **22**, p 113 (2000).
- [4] J. Heyd, *et al.*, Hybrid functionals based on a screened Coulomb potential, *J. Chem. Phys.*, **118**, pp 8207-8215 (2003). J. Heyd, *et al.*, Erratum: Hybrid functionals based on a screened Coulomb potential [J. Chem. Phys. 118, 8207 (2003)], *J. Chem. Phys.*, **124**, p 219906 (2006).
- [5] A. Massaro, A.B. Munoz-Garcia, P. Maddalena, F. Bella, G. Meligrana, C. Gerbaldi, M. Pavone, First-principles study of Na insertion at TiO<sub>2</sub> anatase surfaces: new hints for Na-ion battery design, *Nanoscale Adv.*, **2**, pp 2745-2751, (2020).
- [6] G. Ponti, *et al.*, The role of medium size facilities in the HPC ecosystem: the case of the new CRESCO4 cluster integrated in the ENEAGRID infrastructure, *Int. Conf. on High Perform. Comp. & Simul. (HPCS)* pp 1030-1033 (2014).

# KNOWLEDGE EXTRACTION FROM SOCIAL MEDIA WEB SOURCES: ELEMENTS AFFECTING WEB CRAWLING AND DATA ANALYTICS TASKS IN ENEAGRID

Daniela Alderuccio<sup>1\*</sup>, Silvio Migliori<sup>2</sup> and ICT-HPC Team

<sup>1</sup>*ENEA – TERIN-ICT-HPC – Rome, Italy*

<sup>2</sup>*ENEA– TERIN-ICT – Rome, Italy*

**ABSTRACT.** Web Crawling and Data Analytics give access to Knowledge hidden in data extracted from web open sources (i.e. Internet, social network, electronic papers, etc.), providing information, real-time feedback and updates about users. Among these web open sources, Social Media are virtual deposit of information, where you can extract User Knowledge. ENEA research focuses on external web data source (on-line news, social media platform, etc.).

Social Media Analytics in ENEAGRID needs some preliminary considerations on Knowledge Extraction, enabling optimization of web crawling and data storage task. The question is twofold: (i) how open source information/data from different Social Media influences web crawling, data storage and analysis; (ii) how User's cultural identity affects user knowledge in social media content and mirrors in social media usage.

For their computational costs, these activities are performed in the ENEAGRID/CRESCO infrastructure by using data & task parallelism. Social Media data are processed on a specific "Text Mining & Analytics" Platform, hosted in ENEAGRID environment, for the management of research process in Economic-Social domains. This activity takes place within the ENEA research in: Parallel Text Mining, Web Crawling & Data Analytics of Social Big Data.

Keywords: Social Media Analytics, Big Data from Open Sources in Socio-Economic Sciences, Social Media Intelligence, Web Crawling, Parallel Text Mining.

## 1 Introduction

Social Media are virtual deposit of information where you can extract user knowledge (i.e. information, feedback and updates about users), finding user preferences, behavioural patterns, and trend topics over time. All these user data form User Knowledge, and can be accessed through Web Crawling and Data Analytics, getting an idea of real life scenarios.

Big companies, small enterprises and research teams conduct analytic-driven business using large amounts of data on social media in order to identify significant information and transform it into new knowledge. For this reason there is a growing need to analyze social media data.

This paper presents some considerations on Social Media Knowledge Extraction, with the goal of refining Data Analytics in ENEAGRID [5] and for an optimal allocation of the involved computational resources. Our considerations focus on: (i) how open source information/data from different Social Media influences web crawling, raw data storage and analysis; (ii) how User's cultural identity affects user knowledge in social media content and is reflected in social media usage.

## 2 Data Analytics of Social Big Data in ENEAGRID

ENEA research focuses on external web data source (on-line news, social media platform, etc.). Text mining task of social media data will be performed in the TM & Analytics Platform, thanks to the integration of Language Technologies in GRID environment for Parallel Text Mining of Large Corpora [1-2]. For their computational costs, these activities are performed in the ENEAGRID/CRESCO infrastructure [5] by using data & task parallelism.

### 2.1 ENEA Web Crawling & Data Storage

Internet is the world's largest data source and Web Crawling is the task of downloading a large amount of data from it.

In 2019 ENEA Web crawling task in social media crawler focuses on contents from Twitter and Instagram in real case scenarios: in financial domain [6] and industry 4.0 - Fashion/ Luxury domain (forthcoming). In these application domains, crawling tasks are related to: (i) text crawling, mostly from tweets, and partially from Instagram (tag, captions, etc.); (ii) image crawling, mostly from Instagram and partially from Twitter. Integrated in HPC ENEAGRID/CRESCO infrastructure, ENEA created a collaborative Web Crawling Project in the TM & Analytics Platform and implemented a social media crawler, used to retrieve and analyze data from the Web [6].

Raw data storage from Social Media crawling is at the base of Web Crawling. This task will interest: textual data storage (mostly from Twitter), and image data storage (mostly from Instagram).

### 2.2 Data Analysis

Text Analytics moved from research centers to real-world institutions, such as brands, financial and banking institutions. Analysing the social media landscape, we selected Twitter and Instagram features, to study real-case scenario in marketing and financial domain. Text Mining and Knowledge Extraction from Social Media will be then applied to textual data & relationships: lexical and semantic tagging of texts, hashtags, emoji, named entity, etc.

## 3 Social Media Analytics in ENEAGRID: case-studies

In this research Web crawling task in social media refers to contents from Twitter and Instagram in real case scenario: Industry 4.0, Fashion/Luxury [forthcoming] and in Financial domain [6].

### 3.1 Industry 4.0 : Extracting Knowledge from Social Networks in Fashion domain

Social media are increasingly being used as a customer service platform where User become Consumer and then potential Customers, asking for answers quickly and in real-time. Consumers established mobile shopping habit and 2019 was a booming year for social media shopping<sup>1</sup>. Social media platforms are becoming powerful tools for direct selling not just advertising. Social media shift from measuring user awareness to user engagement. They gather rich volumes of content, profile clients preferences, offer access to content tailored to client interests and location. In Fashion/ Luxury domain, Social media metrics are used to grow the brand, turn customer into advocates, drive leads and sales, improve customer retention, etc. We started some analysis and research are still in progress; results will be published in the Report of CRESCO Results 2020.

---

<sup>1</sup> State of Fashion 2019, McKinsey's & Company - .

<https://www.mckinsey.com/~/media/McKinsey/Industries/Retail/Our%20Insights/The%20State%20of%20Fashion%202019%20A%20year%20of%20awakening/The-State-of-Fashion-2019-final.ashx>

### 3.2 Finance 4.0: Extracting Knowledge from Social Networks in Financial domain

In 2009 Twitter modified its interface, so shifting to an informative function and modifying relationships inside the social network. The relevance of financial Twitter is recognized as well as the positive relation between social and financial importance<sup>2</sup> [7]. The impact of information from online data sources into the financial market is widely acknowledged by researchers and professionals, and there is also a huge consensus about Twitter specifically.

Knowledge extraction in Twitter can be performed mostly by crawling tweets textual contents.

ENEAGRID is equipped with a social media crawler downloading contents from Twitter.

In the first exploratory analysis [6] the ICT-HPC Laboratory use for web crawling and data storage:

(i) the ENEAGRID Computational Power with its High performance storage systems of 3 PB with backup for disaster recovery; (ii) and the Parallelization of crawling task with the use of parallel developer accounts, speeding up crawling tasks and minimizing the limitation imposed by Twitter Policy on the number of tweets downloaded per user.

In the Financial Domain we started to analyze Twitter data related to the Cryptocurrency domain. In order to get a preliminary analysis on twitter data to be refined later, in this phase data were collected from Twitter and few tag words for search query were selected (i.e. digital coin denomination) without performing any previous semantic pre-analysis. We will use these first raw data in order to refine crawling strategy getting more value from knowledge extraction, i.e. selecting more tagwords or finding influencer twitter account.

One of the feature distinguishing financial tweets from common user tweets is the presence of cashtags: the main Twitter sharing mechanism to track financial information. Cashtag is a blend of cash and hashtag and its use is growing as a mechanism to denote a financial theme in a tweet [4].

Hashtag allows people to follow topics they are interested in. Twitter's own research into hashtags confirms that there is significant advantage to using them. Individuals can see a 100 percent increase in engagement by using hashtags and Brands can see a 50 percent increase. Furthermore, tweets with one or more hashtag are 55 percent more likely to be retweeted.

Cashtag is a clickable term consisting in a short sequence of letters (and sometimes numbers), preceded by \$ symbol. A cashtag can be a company ticker identifying uniquely a company in a specific stock market as well as a cryptocurrency (many of them use the dollar symbol \$ followed by the acronym of the cryptocurrency to indicate that the tweet refers to it). Since 2012, Twitter incorporates cashtag as a mechanism to find and track tweets that address companies by their ticker in a specific stock market. However its usefulness has been deteriorated due to the interference and negative effect of homonymy cashtags of cryptocurrencies, becoming remarkably popular at late 2017. Also similarly to hashtags, Twitter supports tracking tweets that contain a specific cashtag. All of this turns cashtags into one of the most useful mechanisms to easily harvest financial information on Twitter. For these reasons, financial tweets needs to be analyzed with a specific attention to cashtag use and distribution over time.

## 6 Conclusions

This study leads to findings, able to refine and tuning crawling and data storage tasks in ENEAGRID, in order to extract Knowledge from Social Media Networks.

The task of extracting knowledge has different aspects because there are different kind of knowledge: i.e. concrete vs abstract knowledge, general vs domain specific knowledge, etc. ([3] De Jong, 1996). Knowledge can be considered as formed and represented either by abstracts data, or by data & their

---

<sup>2</sup> It is acknowledged a positive correlation between stock capitalization (financial importance) and mentions (social importance), meaning that high-capstocks are mentioned more frequently than low-capstocks. This positive relation has already leveraged for predicting stockprices.

relationships among data, or by correlated data in the real-world (that is, data to act and/or to explain a behavior). All this knowledge is mirrored in Social Media Networks. Furthermore, there is not a unique Social Media Landscape but there are many, differing by country, by culture, by sphere of influence (political/commercial), by language, by application domain, by service integration in social platform<sup>3</sup>, by mobile diffusion and uptake<sup>4</sup>, by user age and engagement. For each Social Media Landscape there is a different Social User behaviour, different usage of social features and a different User Knowledge to be extracted. Also the different nature of web sources (online news, websites, social media) affects web crawling, data storage and text mining tasks. Studying Knowledge Extraction from Web Sources, and in particular from Social Media Networks, we find that: (i) User cultural identity influences user knowledge extraction from social media content; (ii) for each Social Media Landscape, User shows a different Social Media behaviour and usage, leading to different User Knowledge to be extracted. For these reasons, Crawling tasks and raw Data Storage of Big Data from Social Networks need to diversify strategies, on the base of many elements (application domain, research purpose, data sources, user target/age/language/education/social status) as well as also the frequency (time: daily, weekly, monthly, quarterly, ecc.) and size (limitations imposed: by web sources policy, i.e. social media crawling policy of Twitter is 3% per day/per crawling; or by machine limitations (computational power required for crawling, data storage and analysis). Furthermore, evidences show that: cross-cultural differences mirrors on social media platforms design and use of social media features; different domains (common vs. financial or fashion domain) shows different tweets structure and feature preferences, etc.; the role of users changes and is becoming more oriented to information sharing and to purchase. The consequent shift from User to Consumer/Customer Knowledge affects the extraction of knowledge from Social Media. Taking into considerations all this finding will enable a refinement of crawling and data storage tasks, optimizing the allocation of computational resources.

## References

- [1] D. Alderuccio et al. TaLTaC in ENEAGRID Infrastructure in Proceedings JADT 2018 -14th Int. Conference on Statistical Analysis of Textual Data)- pp. 501-508 - Roma 15.6.2018 c/o CNR, (2018)
- [2] D. Alderuccio, et al.: Parallel Text Mining of Large Corpora in GRID Environment – TaLTaC in ENEAGRID Infrastructure. In: “High Performance Computing on CRESCO Infrastructure: research activity and results 2018”, pp. 227-232 - ISBN: 978-88-8286-390-6 - Publisher: ENEA (2019)
- [3] T., De Jong, Ferguson-Hessler, Types and Qualities of Knowledge, in Educational Psychologist, 31 (2), 105-113, Laurence Erlbaum Associates, (1996) - <https://ris.utwente.nl/ws/portalfiles/portal/6401593/types.pdf>
- [4] A. Fernandez Vilas, R. P. Diaz Redondo, A. L Garcia, The Irruption of Cryptocurrencies Into Twitter Cashtags: A Classifying Solution - vers. Feb 25, 2020 doi: 0.1109/ACCESS.2020.2973735 <https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=8998217> - (2020)
- [5] F. Iannone et al., "CRESCO ENEA HPC clusters: a working example of a multifabric GPFS Spectrum Scale layout," 2019 International Conference on High Performance Computing & Simulation (HPCS), Dublin, Ireland, pp. 1051-1052, doi: 10.1109/HPCS48598.2019.9188135 (2019)
- [6] G. Santomauro, D. Alderuccio, F. Ambrosino, S. Migliori. A Brand Scoring System for Cryptocurrencies Based on Social Media Data - In book: Mining Data for Financial Applications. MIDAS 2019 - Publisher: Springer Nature Switzerland (2020) - doi: 10.1007/978-3-030-37720-5\_11
- [7] Yuexin Mao, Wei Wei, Bing Wang, and Benyuan Liu. Correlating S&P 500 stocks with Twitter data. In Proceedings of the 1st International Workshop on Hot Topics on Interdisciplinary Social Networks Research (SIGKDD'12 Workshops). ACM,69–72 (2012)

---

<sup>3</sup> Service integration in Social Networks refers to e-commerce features, including e-payment systems, online stores and cashless purchase.

<sup>4</sup> High mobile penetration also helps driving the performance of massive online events.

# NEUTRONICS ANALYSIS OF DEMO DIVERTOR

Rosaria Villari<sup>\*1</sup>, Davide Flammini<sup>1</sup>, Giovanni Mariano<sup>1</sup>, Giuseppe Mazzone<sup>1</sup>, Fabio Moro<sup>1</sup>,  
Simone Noce<sup>2</sup>

<sup>1</sup>*ENEA, Department of Fusion and Nuclear Safety Technology, I-00044 Frascati (Rome), Italy*

<sup>2</sup>*University of Rome Tor Vergata, Industrial Engineering Department, I-00133, Rome, Italy*

**ABSTRACT.** Detailed neutronics analyses have been performed with MCNP5 Monte Carlo code on HPC CRESCO cluster aimed at evaluating the divertor nuclear responses to provide inputs for thermo-hydraulic and thermo-mechanical analyses and to assess its shielding performances.

A heterogeneous neutronics model of the last divertor design has been developed and integrated in the DEMO MCNP model with a Water Cooled Lithium Lead (WCLL) blanket. Neutron flux, nuclear heating, cumulated damage and helium production have been calculated in all divertor components as well as on the fixation systems, rails, in-vessel coil and vacuum vessel.

## 1 Introduction

In the framework of the work package ‘Divertor’ (WPDIV) within the EUROfusion Power Plant Physics and Technology (PPPT) program, detailed neutronics analyses have been performed aimed at evaluating the divertor nuclear responses to provide inputs for thermo-hydraulic and thermo-mechanical analyses and to assess its shielding performances. An heterogeneous neutronics model of DEMO divertor has been developed and integrated in the DEMO MCNP model with a Water Cooled Lithium Lead (WCLL) blanket.

Neutron flux, nuclear heating, cumulated damage and helium production have been calculated in all divertor components as well as on the fixation systems, rails, in-vessel coil and vacuum vessel with MCNP5v1.60 [1] code and JEFF 3.3 [2] nuclear data library on HPC CRESCO6 cluster [3].

## 2. DEMO MCNP model with full heterogeneous divertor

The heterogeneous neutronics model divertor was developed on the basis of the CAD file of the last divertor design. The original CAD model was simplified and then converted into MCNP representation using SuperMC code [4]. The stand-alone model of the water-cooled Eurofer divertor cassette has been integrated in the reference 11.25° DEMO 2017 generic model with simplified representation of WCLL breeding blanket. Reflective boundary condition has been applied to simulate the full reactor. The model is shown in Figure 1.



**Fig.1:** MCNP model of DEMO 2017 with integrated heterogeneous divertor model based on 2019 design

### 3 Nuclear analyses results

The three-dimensional neutron and photon transport calculations were performed in parallel through MPI protocol on HPC CRESCO6 cluster. The results refer to simulations of  $5 \times 10^9$  source particles and have been normalized to 1998 MW of fusion power corresponding to  $7.09 \times 10^{20}$  n/s.

The nuclear power loads in divertor components are reported in Table 1

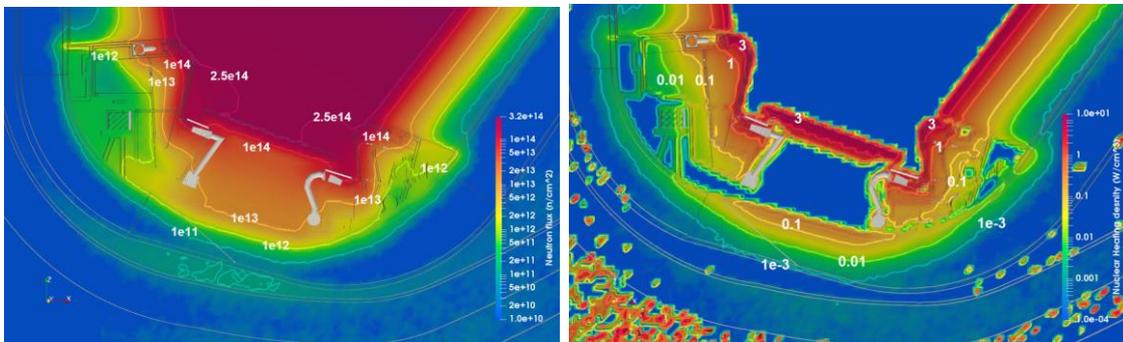
**Table 1** – Nuclear power loads in divertor components for 1 and 48 cassettes

Component	Sub-Component	kW – 1 cassette	MW-48 cassettes
Total	-	2779.1	133.40
CB	CB Eurofer	672.8	32.29
	CB Water	342.7	16.45
	IB_fixation	0.2	0.01
	OB_wishbone	0.7	0.03
PFC	Ist layer- W	99.6	4.78
	II layer-CuCrZr-water-W	184.8	8.87
	III layer W	55.4	2.66
	support Eurofer	26.6	1.28
Liner	I layer- W	65.4	3.14
	II layer -Eurofer	42.9	2.06
	Shield block Eurofer	707.2	33.95
	coolant –Water	344.6	16.54
	support Eurofer	3.2	0.16

<b>Reflector Plates</b>	Ist layer- W	18.9	0.91
	II layer-Eurofer	18.5	0.89
	Shield block Eurofer	56.5	2.71
	coolant –Water	29.5	1.42
	support Eurofer	12.5	0.60
Pipes	Eurofer	47.2	2.27
	water	50.1	2.40

About 133 MW of nuclear power is deposited on the divertor: 41.9% on the liner (56 MW), 36.6% on the cassette body (32 MW on Eurofer and ~16 MW on water), 13,2% on the plasma facing components (17.6 MW) and 4.9% on reflector plates (6.5 MW), the rest on the cooling pipes (5.5 MW).

The neutron flux and nuclear heating density maps in the middle of the lateral cassette are shown in figure 2.



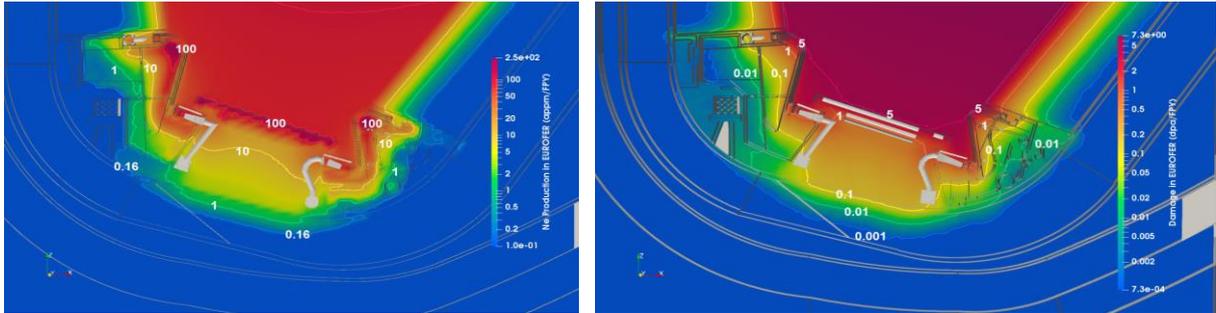
**Fig.2:** Neutron flux ( $\text{n}/\text{cm}^2/\text{s}$ ) (left) and Nuclear heating density ( $\text{W}/\text{cm}^3$ ) (right) maps in the middle of lateral cassette

The neutron flux in divertor components ranges from  $\sim 2.5 \times 10^{14}$   $\text{n}/\text{cm}^2/\text{s}$  in the plasma facing components (PFC) to  $3 \times 10^{11}$   $\text{n}/\text{cm}^2/\text{s}$  in inboard cassette body close to the attachment. A significant streaming from blanket –divertor gaps and pumping duct is observed.

The maximum nuclear heating densities are in the liner:  $22.4 \text{ W}/\text{cm}^3$  on the first W layer and  $7.7 \text{ W}/\text{cm}^3$  on Eurofer; the maximum values in the inboard and outboard supports which are not actively cooled are  $0.035$  and  $0.1 \text{ W}/\text{cm}^3$ , respectively. In particular, regarding the outboard support, the peak refers to the Inconel pin, whereas the maximum in Ti-6Al-4V wishbone is  $0.08 \text{ W}/\text{cm}^3$  and the average value is  $0.028 \text{ W}/\text{cm}^3$ . The maximum nuclear heating density in the VV is  $0.16 \text{ mW}/\text{cm}^3$ , below the target value of  $0.3\text{-}0.5 \text{ W}/\text{cm}^3$ . The following thermo-hydraulic analyses have shown that the nuclear heating level are acceptable to keep them maximum temperature in Eurofer components below the limit ( $550^\circ\text{C}$ ).

The He-production and damage in Eurofer per Full-Power Year (FPY) in the middle of lateral cassette are shown in figure 3. The maximum He-production is  $93.6$  appm/FPY in the second layer of the liner. The cut and re-welding of the pipes is feasible only outside the  $0.16$  appm contour (zones protected by blanket modules) corresponding to the re-weldability limit of  $1$  appm over  $6$  FPY (i.e. DEMO lifetime).

The damage in dpa varies in the range 0.01 - 5 dpa/FPY. The cumulated damage exceeds the limit of 6 dpa over 1.5 FPY, i.e. target lifetime of the divertor cassette, on the top of the liner and in the PFC support in baffle region (maximum cumulated value 7.4 dpa in liner). The level of cumulated damage on the vacuum vessel over 6 FPY is mostly lower than 0.33 dpa except for a small hotspot of 1 dpa after 6 FPY below the pumping duct, thus the present design fulfils the 2.75 dpa limit on the vacuum vessel stainless steel.



**Fig.3:** Maps of He-production (appm/FPY) (let) and damage (dpa/FPY) (right) in Eurofer in the middle of lateral cassette

The results of the neutronics analyses, reported in detail in [5], showed the validity of the present shielding functions of the present liner design to protect the vacuum vessel. However, the damage in Eurofer exceeds the design limit and further design optimization and studies are needed, otherwise the cassette lifetime should be limited to 1.2 FPY.

## References

- [1] X-5 Monte Carlo Team: MCNP - A General Monte Carlo N-Particle Transport Code, Version 5, Los Alamos National Laboratory, Los Alamos, New Mexico, USA, April 2003.
- [2] JEFF-3.3 evaluated data library, <https://www.oecd-nea.org/dbdata/jeff/jeff33/index.html>
- [3] F. Iannone et al., "CRESCO ENEA HPC clusters: a working example of a multifabric GPFS Spectrum Scale layout," 2019 International Conference on High Performance Computing & Simulation (HPCS), Dublin, Ireland, 2019, pp. 1051-1052, doi: 10.1109/HPCS48598.2019.9188135
- [4] Yican Wu, et al. Multifunctional Neutronics Calculation Methodology and Program for Nuclear Design and Radiation Safety Evaluation, Fusion Science and Technology, 74:4, 321-329 (2018)
- [5] R. Villari et al., Neutronic Analysis of DEMO Divertor 2019, Eurofusion Report DIV-1-T006-D005, <https://idm.euro-fusion.org/?uid=2NL3S> (2020)

# MONTE CARLO SIMULATION OF TEMPORARY STORAGE FACILITIES FOR RADIOACTIVE WASTE: EXPOSURE EVALUATIONS

Luigi Lepore<sup>1\*</sup>, Giada Gandolfo<sup>2\*</sup>, Nadia Cherubini<sup>1</sup>, Giuseppe A. Marzo<sup>1</sup>

<sup>1</sup>ENEA, Italian National Agency for New Technologies, Energy and Sustainable Economic Development, C.R. Casaccia, Rome, Italy

<sup>2</sup>ENEA, Italian National Agency for New Technologies, Energy and Sustainable Economic Development, C.R. Frascati, Rome, Italy

**ABSTRACT.** Licensing of Radioactive Waste Storage Facilities requires accurate evaluations about the radiological exposure conditions for workers and public. Both ‘Normal’ and ‘Accidental’ scenarios have to be evaluated and Radiation Protection Experts can adopt different calculation tools to accomplish their tasks. In this framework, Monte Carlo simulation codes for radiation transport play a significant role but they require High Parallel Computing when Storage Facilities with thousands of ‘Radioactive Waste Packages’ are taken into account. The paper shows a case study with more than 20000 drums stored at a Temporary Storage Facility with dimensions in tens of meters: the comprehensive evaluation for photon external irradiation exposure for workers and public has required about 45000 hours for calculations, reduced to 90 hours –in effective time– thanks to the multi-processing on CRESCO High Parallel Computing Resources.

## 1 Introduction

Monte Carlo techniques are widely used in the field of radiation transport simulation in media. They can be used: i) as ‘direct techniques’ to study photons, neutrons, electrons, etc., paths in media within a specified geometry; ii) as ‘indirect techniques’ to run sensitivity and uncertainty calculations about critical parameters in complex systems modelling.

Some examples are given in [1][2][3] where Monte Carlo radiation transport codes are used to address shielding calculations, as to estimate dose-rates to exposed workers and public, e.g. for transportation to, placement, rearrangement or removal of radioactive waste packages at the storage site. The MCNP (Monte Carlo N-Particle) code [4] is one of the most widespread and used in the field.

On the other hand, Monte Carlo techniques are used in radioactive contamination transport through media [5][6][7]: advection and diffusion equations (e.g. through engineered barriers, aquifers, ground soils, air, etc.) take advantages of Monte Carlo sampling to treat uncertainty in the ‘most sensitive’ parameters of mathematical models used, each parameter given as a distribution of values instead of a single and unique value. Currently, AMBER [8] and GOLDSIM [9] codes belong to the state-of-the-art of the field.

Whatever the usages of Monte Carlo techniques, the examples given before are showing that those applications require large use of High Parallel Computing (HPC) resources as to reduce the time needed to reach statistically significant results in searched quantities, e.g. the effective dose-rate at a certain position, or the concentration of a radionuclide dispersed within the volume of air or water of interest.

Calculation time can significantly increase in the following cases: i) the smaller is the resolution to be achieved (e.g. at a certain position, a photon flux is searched as a discretization in several energy channels, instead of the same quantity integrated in energy); ii) the larger are the physical and geometrical dimensions of the calculation to be solved (e.g. a photon flux to be determined 10 m

away from the radioactive source, instead of a photon flux to be determined 1 m away from the radioactive source).

The paper herein shows a case study in which a Temporary Radioactive Waste Storage Facility with more than 20000 drums is modelled within the MCNP code. The main goal of the calculation was the determination of radiological exposure conditions at the ‘Normal Operation’ scenario, when the photon external irradiation of people is acknowledged as the unique exposure pathway. The comprehensive evaluation has required about 45000 hours for calculations, reduced to 90 hours –in effective time– thanks to the multi-processing on CRESCO High Parallel Computing Resources. The computing resources and the related technical support used for this work have been provided by CRESCO/ENEAGRID High Performance Computing infrastructure and its staff [10]. CRESCO/ENEAGRID High Performance Computing infrastructure is funded by ENEA, the Italian National Agency for New Technologies, Energy and Sustainable Economic Development and by Italian and European research programmes, see <http://www.cresco.enea.it/english> for information.

## 2 Methodology

Currently, the most efficient method to simulate the radiation transport in media through complex (many volumes or “cells”, several radioactive sources, and different materials) and large (few centimetres to hundreds of meters) geometries is the Monte Carlo simulation.

By this approach, a lot of ‘histories’ of the particle of interest are run, in order to reproduce, into a virtual world, all the possible paths travelled by the radiation, by means of random samplings from all the probability density functions ruling physical phenomena. Very complex geometries can be built, and very accurate results can be produced if all the elements important to the transport of the analysed radiation are described into the model; moreover, an adequate computational power should be available for running a statistical significant number of histories. For this work, the MCNPX (Monte Carlo N-Particle eXtended) code has been used, in its version MCNPX 2.7.0.

The horizontal cross-section of the Temporary Storage Facility here studied is measuring 80 m times 30 m, and the building is 10 m tall; the facility is hosting about 20000 Waste Packages, as sketched in Figure 1.

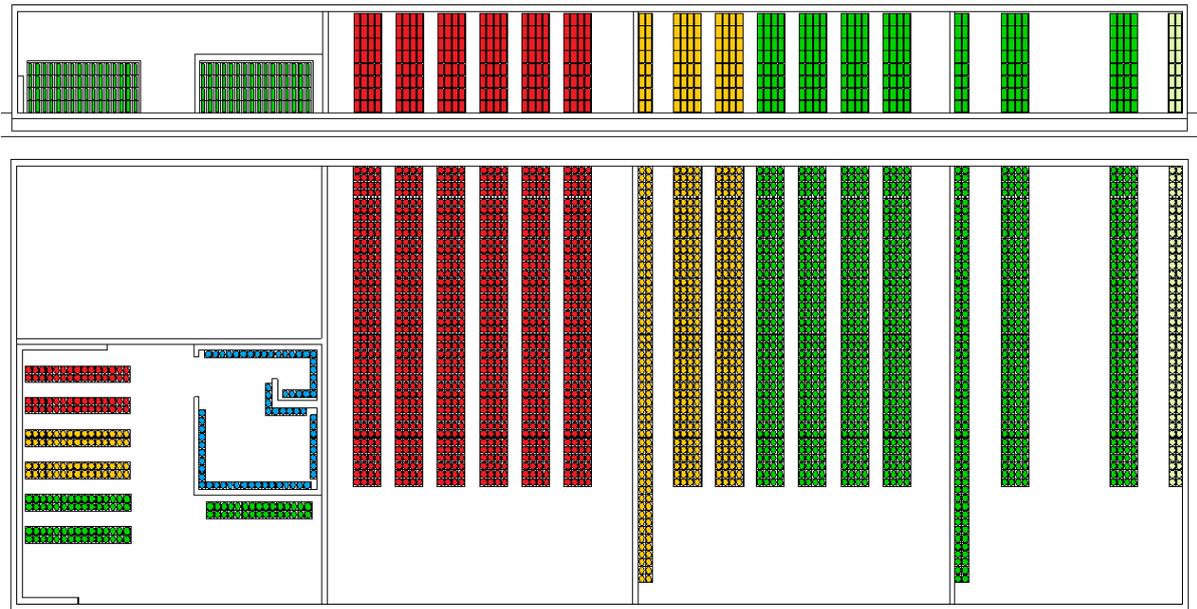
The radioactive inventory of radionuclides contained within the Waste Packages is distributed keeping the most ‘harmful’ radionuclides (higher activity concentration, gamma emitters with photon emission in several hundreds of keV; radionuclides with half-life in the order of decades of years, etc.) at the centre of the facility, using radiologically ‘safer’ Waste Packages at the edges of the facility, as a shielding medium.

To simplify the modelling, all Waste Packages are assumed to be in 220-liters-drum geometry with a waste matrix weight composition in H~22%, C~14%, N~1%, O~18%, Si~7%, Cr~7%, Fe~28%, Ni~3%, constant for all Waste Packages. The apparent density chosen for Waste Matrix is  $0.3 \text{ g cm}^{-3}$ . The radioactive content is supposed to be homogeneously dispersed within the Waste Matrix.

Building walls are modelled as a 30-cm-thick slabs made of concrete with an apparent density corresponding to  $0.8 \text{ g cm}^{-3}$ .

Several tens of dose-rate evaluation points are set all around the building and outside the fence of the site. Dose rate evaluation points are exploiting the MCNP ‘point detector’ tally [4] modified with a dose conversion function taken from the ICRP-116 [11], Table A.1, using the Antero-Posterior (AP) irradiation geometry for people. Evaluation points at the wall are taken 30 cm away from the wall, the height from the floor is 150 cm; points outside the fence of the site are taken about 20 m away from the building walls.

External irradiation dose rates from photons are calculated at each point. Occupancy of exposed workers and public are evaluated. Radiological exposure ‘target’ values are set to be compliant with constraints from the Regulation in force at the moment of this work, the Italian D. Lgs. 230/1995, now superseded by the Italian D. Lgs. 101/2020 in force from August 27<sup>th</sup> on.



**Figure 1** – The Temporary Storage Facility for Radioactive Waste Packages here studied. The horizontal cross-section of the facility here shown is measuring 80 m times 30 m, and the building is 10 m tall. Colour maps is referring to the radiological ‘hazard’ of Waste Packages: from red to green, being red the most active packages.

### 3 Results and Discussion

Having set radiological constraints for exposed workers and public, the MCNP model of the Temporary Storage Facility here studied has been used to optimize the total radiological inventory of the facility by:

1. calculating the maximum radiological inventory allowed;
2. choosing the best placement and facility filling strategy;
3. verifying the compliance with the radiological constraints due to Regulation;
4. optimizing the radiological exposure of workers involved and public.

Activity of the radionuclides in the total inventory contained in Waste Packages are ranging from tens of TBq for ‘safer’ items (green ones in Figure 1) to hundreds of MBq for the most ‘harmful’ items (red ones in Figure 1). Yellow items are in between; blue ones are sealed high activity sources. The comprehensive evaluation has required about 45000 hours for calculations, reduced to 90 hours –in effective time– thanks to the multi-processing on CRESCO High Parallel Computing Resources. The whole study has requested about two months due to: i) preparation of several MCNP input model to study each inventory (‘green’, ‘yellow’, ‘red’, ‘blue’ ones in Figure 1) vs. dose-rate points and determine separated contributions; ii) waiting time due to queues on the CRESCO4 System due to the high number of nodes requested; iii) analyse MCNP output files as to retrieve the information searched for; iv) repeat the whole process as to obtain the radiological inventory vs. exposure optimization searched for.

At the final stage of the optimization process, effective dose-rates at the evaluation points are showing values ranging from 1 to 3  $\mu\text{Sv a}^{-1}$ ; only in two points (closer to the most active Radioactive Waste Packages, the ‘red’ ones) are showing larger values, 5  $\mu\text{Sv a}^{-1}$  and 7  $\mu\text{Sv a}^{-1}$ . Each point is inherently without radiological significance. Thanks to the Computing Resources from CRESCO and the large number of ‘histories’ run, Monte Carlo uncertainties in obtained values are < 0.1%.

The evaluation carried out is taking care of external irradiation from photons only, such pathway acknowledged as the only one happening in ‘Normal Operation’ conditions of the Temporary Storage Facility here studied. A comprehensive evaluation of the exposure need to take care of other pathways also (i.e. dispersion of radionuclide and contamination of workers and public, etc.) e.g. due to ‘Accidental Events’ such as rupture of a Waste Package, fire, flooding, etc.

## References

- [1] T. A. Bayoumi, S. M. Reda, and H. M. Saleh, “Assessment study for multi-barrier system used in radioactive borate waste isolation based on Monte Carlo simulations,” *Appl. Radiat. Isot.*, vol. 70, no. 1, pp. 99–102, 2012.
- [2] T. Özdemir, “Monte Carlo simulations of radioactive waste embedded into EPDM and effect of lead filler,” *Radiat. Phys. Chem.*, vol. 98, pp. 150–154, 2014.
- [3] C. J. Park, J. J. Park, and K. C. Song, “Shielding analysis for an Advanced Voloxidation Process with Monte Carlo calculations,” *Ann. Nucl. Energy*, vol. 36, no. 5, pp. 694–699, 2009.
- [4] J. F. Briesmeister, Ed., “MCNP: A General Monte Carlo N-Particle Transport Code,” 2000.
- [5] P. Poškas, R. Kilda, A. Šimonis, H. Jouhara, and R. Poškas, “Disposal of very low-level radioactive waste: Lithuanian case on the approach and long-term safety aspects,” *Sci. Total Environ.*, vol. 667, pp. 464–474, 2019.
- [6] F. Cadini, J. De Sanctis, T. Girotti, E. Zio, A. Luce, and A. Taglioni, “Monte Carlo-based assessment of the safety performance of a radioactive waste repository,” *Reliab. Eng. Syst. Saf.*, vol. 95, no. 8, pp. 859–865, 2010.
- [7] F. Cadini, J. De Sanctis, I. Bertoli, and E. Zio, “Monte Carlo simulation of radionuclide migration in fractured rock for the performance assessment of radioactive waste repositories,” *Reliab. Eng. Syst. Saf.*, vol. 111, pp. 241–247, 2013.
- [8] QUINTESSA, “AMBER 6.4 Reference Guide,” 2019.
- [9] G. T. Group, “GoldSim User’s Guide, Probabilistic simulation environment, version 12.1.5,” 2020.
- [10] F. Iannone *et al.*, “CRESCO ENEA HPC clusters: a working example of a multifabric GPFS Spectrum Scale layout,” in *2019 International Conference on High Performance Computing Simulation (HPCS)*, 2019, pp. 1051–1052.
- [11] N. Petoussi-Hens *et al.*, “Conversion Coefficients for Radiological Protection Quantities for External Radiation Exposures,” *Ann. ICRP*, vol. 40, no. 2–5, pp. 1–257, Apr. 2010.

# STUDY OF ADVANCED CONFIGURATIONS ON TOKAMAK EAST WITH THE EMC3 EIRENE

Meineri Carlo<sup>1</sup>

<sup>1</sup>*Politecnico di Torino, Dipartimento Energia, Turin, Italy*

**ABSTRACT.** The aim of this work is to study the impact of a particular magnetic geometry, "Two Null Nearby Divertor" (TNND) on the transport of particles and energy in the SOL to the Tokamak EAST. Experimental machine for the study of fusion plasma, located in Hefei, China. The standard single-null magnetic configuration is perturbed, through another secondary null placed outside the divertor walls. The main effect is the reduction of the poloidal magnetic field  $B_p$  near the wall, which makes the field lines diverge. This maximizes the dissipative processes where  $B_p$  is lower, obtaining the detachment condition for lower puffing and equator density. This particular phase is characterized by a strong reduction in the flow of heat and ions to the wall. This analysis was carried out with a tool for modeling the physics of the SOL, EMC3-EIRENE. It is necessary not only because the diagnostics are not sufficient on their own to interpret the experiments, but also to carry out a predictive study to optimize them.

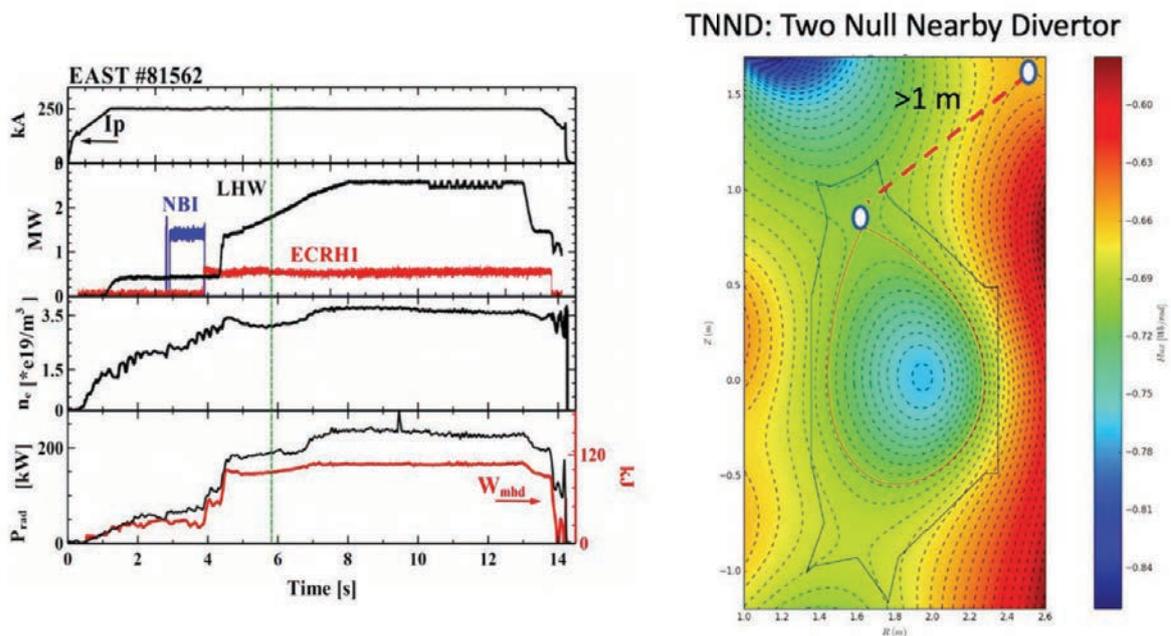
## 1 Introduction

Nuclear fusion could be the solution for the current dependency on fossil fuels and the most scientifically and technologically advanced concept with which to achieve fusion in a controlled environment is the divertor tokamak. It's a magnetic confinement device and will be the configuration for the next step experimental international fusion machine, ITER [1]. In the divertor tokamak a region of closed field lines, the confinement region, and one of open field lines, the Scrape Off Layer (SOL) exist. In the confinement region field lines lie on closed surfaces, whereas in the SOL the field lines intersect with target plates located in the divertor. Energy and particles which diffuse from the confinement region into the SOL are guided along the field towards these target plates. One of the key problems of fusion research is to limit the peak heat loads on the target plates. If no energy is dissipated on the way to the targets the power flux density would exceed  $100 \text{ MW}/\text{m}^2$  in future divertor tokamaks and it is currently estimated that the steady state heat flux has to be limited to  $5 - 10 \text{ MW}/\text{m}^2$  in order to guarantee a sufficiently long material lifetime. The magnetic configuration XD [2] or 'Two Null nearby divertor' (TNND) [3], is created by inducing a secondary X-point outside the divertor region, downstream of the main X-point, with the beneficial effect of increase the field line lengths, connection length, and the distance between the flux surfaces, the poloidal flux expansion, from the core X-point to the wall. This positive gradient in flux expansion toward the target is called poloidal flaring, for the divergent character of the field lines. These reduce the peak power flux and the temperature at the strike points. This work studies the impact of the distance of the second null on the power deposited on the divertor plate at EAST [4] tokamak and the detachment threshold, a regime characterized by low target heat and ionized particle fluxes, and low plasma temperatures at the walls. The study is conducted with EMC3-EIRENE [5], a 3D edge plasma transport modelling tool developed to study helical devices, but recently it has been adapted to tokamak also. The computing resources and the related technical support used for this work have been provided by CRESCO/ENEAGRID High Performance Computing infrastructure and its staff [6]. CRESCO/ENEAGRID High Performance Computing

infrastructure is funded by ENEA, the Italian National Agency for New Technologies, Energy and Sustainable Economic Development and by Italian and European research programmes, see <http://www.cresco.enea.it/english> for information.

## 2 Experimental magnetic configurations

During the last experimental campaigns, high confinement (H-mode) plasma discharges in steady state have been done with a TNND configuration. During these discharges, the edge Localized Modes (ELM) were absent [7], indicating a possible non-linear interaction between the downstream magnetic topology. The same type of alternative heating system has been applied to standard SN, showing the presence of ELMs activity. The reasons for the differences are under investigation. The experimental discharge getting as reference discharge is #81562. Fig. 1.b shows its poloidal section, where the red dashed line represents the distance between the main and the secondary null. The plasma current  $I_p = 0.25\text{MA}$  (Fig. 1.a.a), the toroidal magnetic field  $B_t = 2.5\text{T}$ , and the edge safety factor  $q_{95} = 10$ . The heating scheme (Fig 1.a.b) consists of the ECRH ( $\sim 0.5\text{MW}$ ) and LHW at  $2.45\text{ GHz}$  ( $\sim 0.45\text{MW}$ ), from the start of the flat-top and a ramp-up of LHW starting at  $4.6\text{ GHz}$  at  $4.5\text{ s}$ , from  $1\text{ MW}$  to  $2\text{ MW}$ . The central line-averaged electron density (Fig. 1.a.c) is  $\bar{n}_e = 3.5 \times 10^{19}\text{m}^{-3}$ , from the Polarimeter-Interferometer (POINT) system, with corresponding Greenwald fraction  $0.51 - 0.58$ .



**Fig. 1:** (a) Time evolution of main plasma quantities for TNND #81562: plasma current  $I_p$  (a), additional heating PHEAT (b) (LH and ECRH power for discharges), interferometer line average electron density  $n_e$  (c), total radiated power PRAD from Bolometer diagnostic (d); (b) poloidal section of magnetic configuration #81562, where the red dashed line represents the distance between the main and the secondary null.

The initial position of the reference second null (1.3 m from the main null) is modified, to obtain two other configurations, respectively with the second null at 1.0 m and 0.8m. they are realized with MAXFEA 2D MHD, a non-linear Equilibrium solver. The connection length, poloidal flux expansion and the angle between the magnetic field lines and the wall, grazing angle are compared in Table 1.

**Table 1:** MAIN MAGNETIC GEOMETRY FOR THE TNND DISCHARGES

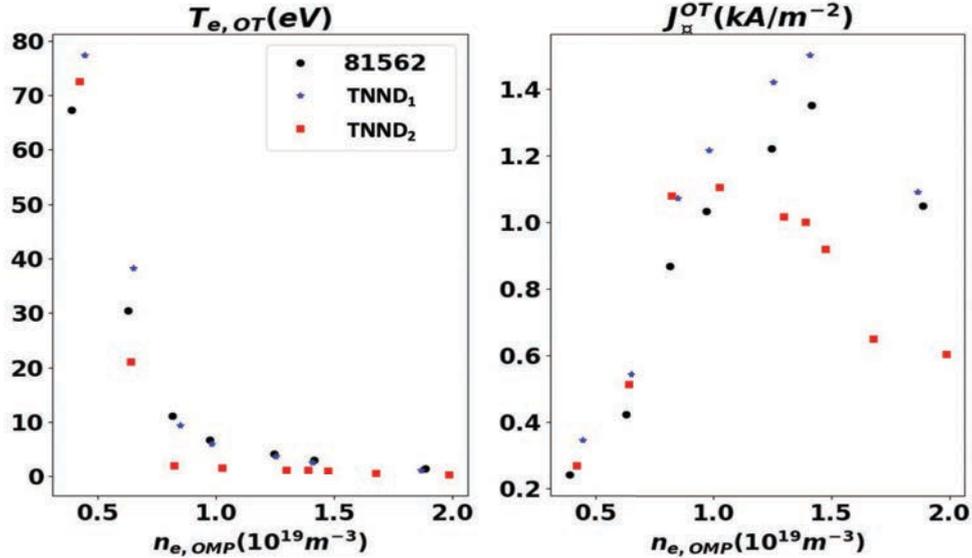
	#81562	#TNND1	#TNND2
Connection Length [m]	32	33	40
Poloidal flux expansion at outer SP	4.6	4.8	6
Grazing angle Outer Target [deg]	5	4.86	2

The first two cases, #81562 and the #TNND1, have similar properties in terms of connection length and flux expansion. In the third case, #TNND2 has a high reduction of Bp toward the target, with a consequent reduction of the flux expansion by 30% greater than the reference case.

### 3 Density Scan

A scan in separatrix electron density  $n_{e,u}$  has been performed to study the detachment threshold for the three cases. Fig. 2 shows the values of temperature  $T_{e,low\ field\ side\ target\ plate}$  (a) and particle flux density  $J_{low\ field\ side\ target\ plate}$  (b) deposited on the upper outer target at the strike point.

The density scan covers the range  $n_e = 0.3 - 2.0 \times 10^{19} m^{-3}$ , where the black dots represent the reference case #81562, the blue star the TNND1, and the red square the closed case TNND2. According to the figure, the electron target temperature decreases rapidly as  $T_{e,low\ field\ side\ target\ plate} \propto 1/n_{e,u}$  and the particle flux deposited on the target increases as  $J_{low\ field\ side\ target\ plate} \propto n_{e,u}^2$  until  $n_{e,u} \approx 1 \times 10^{19} m^{-3}$  for the TNND2 case and  $n_{e,u} \approx 1.5 \times 10^{19} m^{-3}$  for the other two cases. In this range, the plasma can be well described by a high recycling regime and in agreement with the basic two-point model for the divertor SOL (presented in chapter 5.2 in [8]). When  $P_{SOL}$  is not sufficient to sustain ionization in front of the targets, the ionization front gradually shifts towards the separatrix and  $n_{e,low\ field\ side\ target\ plate}$  decreases, and the plasma enters into a state of detachment.



**Fig. 2:** The Figure showing the density scaling of Electron temperature (on the left) and particle flux (on the right) on upper outer target. The black circles represent the #81562 reference case, the blue start the TNND1 and the red square the TNND2.

The plot depicts that there is a clear correlation between the starting of detachment and distance between two nulls. This is caused probably by both the energy lost with neutral-plasma interaction, as shown in table 2 and both geometrical effects. The first two cases are very similar instead the third cases arrive at  $\sim 35\%$  of power lost with neutral interaction.

**Table 2:** Integrated power arriving to both targets, normalized to the total input power  $P_{in}$ , fraction of power transferred to the neutral particles (charge exchange and elastic collisions) and power fraction leaving the computational domain radially according to EMC3-Eirene simulations for the two discharges.

<i>Case <math>n_{e,OMP} = 0.4e^{19}m^{-3}</math></i>	#81562	#TNND1	#TNND2
Ptargets (%)	89.02	86.86	83.77
Neutrals (%)	7.73	8.16	14.83
Boundary (%)	3.26	4.97	1.41
<i>Case <math>n_{e,OMP} = 0.8e^{19}m^{-3}</math></i>	#81562	#TNND1	#TNND2
Ptargets (%)	83.98	83.84	74.31
Neutrals (%)	14.97	14.47	16.66
Boundary (%)	1.05	1.69	8.82
<i>Case <math>n_{e,OMP} = 1.5e^{19}m^{-3}</math></i>	#81562	#TNND1	#TNND2
Ptargets (%)	72.95	65.79	53.56
Neutrals (%)	18.38	22.29	35.30
Boundary (%)	8.63	11.84	11.10

## References

- [1] ITER tokamak, “<https://www.iter.org>”
- [2] Kotschenreuther M., “X-divertor,” 2004.
- [3] G. Calabrò, “EAST Alternative Magnetic Configurations: Modelling and First Experiments,” in IAEA, 2014.
- [4] EAST tokamak, “<http://english.ipp.cas.cn/rh/east/>”
- [5] Y. Feng, “Recent improvements in the emc3-eirene code,” *Contrib. Plasma Phys.*, vol. 54, no. 4-6, pp. 426–431, 2014.
- [6] F. Iannone et al., “CRESCO ENEA HPC clusters: a working example of a multifabric GPFS Spectrum Scale layout,” 2019 International Conference on High Performance Computing & Simulation (HPCS), Dublin, Ireland, 2019, pp. 1051-1052, doi: 10.1109/HPCS48598.2019.9188135.
- [7] G. Calabrò, “Developing steady state elm-absent h-mode scenarios with advanced divertor configuration in east tokamak,” in IAEA, 2018.
- [8] P. Stangeby, *The Plasma Boundary of Magnetic Fusion devices*, 2000.

# SIMULATION OF ELECTRICAL AND THERMAL PROPERTIES OF CARBON NANOTUBES AT NANOSCALE

Y. Zhao<sup>1\*</sup>, G. Donati<sup>2</sup>, A. De Nicola<sup>2</sup>, A. Pizzirusso<sup>2</sup>, G. Munao<sup>2</sup>, S. Caputo<sup>2</sup>, G. Milano<sup>2,3</sup>

<sup>1</sup> *Dalian Minzu University, Institute of Nano-Photonics, School of Physics and Materials Engineering, 116600, Dalian, China.*

<sup>2</sup> *Salerno University, Department of Chemistry and Biology "Adolfo Zambelli", 84084, Via Giovanni Paolo II, 132, Fisciano (SA), Italy.*

<sup>3</sup> *Department of Organic Materials Science, Yamagata University 4-3-16 Jonan Yonezawa, Yamagata-ken 992 8510, Japan*

**ABSTRACT.** A huge scientific and technological interest is continuously growing for the so-called *smart materials*. This classification includes composite materials made of a polymer matrix (such as a resin) and a filler whose electrical and thermal properties are of key importance, and carbon nanotubes (CNTs) are among the possible fillers (beside graphite, carbon fibers etc.). When such filler is added to polymer matrices, it makes them conductive once they overcome a specific (so-called percolative) concentration. Once they are conductive, Joule effect (heat generated because of the current flow) and piezoresistivity (change of resistance after application of external stress or strain), are important features that characterize these materials and that can be employed for self-healing or self-curing (Joule effect), i.e. the material is able to detect and react to its own condition, or for self-sensing (piezoresistivity), i.e. the material detects if an external stress is applied by changing its resistance. For these reasons they are called *smart*.

We simulate the self-assembly processes of carbon nanotubes dispersed in linear polymer or monomer phase using a hybrid particle-field molecular dynamics technique (MD-SCF). This efficient computational approach allows to simulate large-scale systems (up to  $\sim 1\,500\,000$  particles) of flexible rod-like particles in different matrices on the millisecond time scale. It is indeed very important to catch the impact of different morphologies on the electrical properties and how these affect the performances of the system. The equilibrium morphologies are in good agreement with those proposed by several experimental studies that hypothesized a two level "multiscale" organization of CNT assemblies. We have successfully modelled Joule effect on the nanoscale, simulating the heat generation due to the conductive filler and the diffusion of the produced heat. Our studies are focused on the understanding of the effect of the applied voltage, the morphology and the filler concentration, with the final aim to help understanding which are the limits of these materials and how to improve their performances. Interesting results are found also on piezoresistivity.

## 1 Introduction

Theoretical-computational modelling represents nowadays an established crucial approach to help in understanding experimental results, giving the molecular picture necessary to interpret the phenomena under study in chemical field. In-silico experiments are very important to make previsions on processes under study thanks to the capability to model a system and/or a process

before actually perform it in practice, however there are often limits in the size and timescales that can be simulated.

Composite materials are showing a huge interest in the scientific and technological community because of their unique electrical and thermal properties, and can be employed in many different technological and industrial fields from automotive to aeronautics to civil engineering. In particular, carbon nanotube/polymer nanocomposites have received attention from the scientific and technological point of view since CNTs were first observed by Iijima[1] in 1991. Once the CNT concentration overcomes a specific threshold (*percolation threshold*), a continuous conductive network of CNTs is formed and the material becomes conductive. This critical concentration depends on several aspects, among which the nature of the filler itself, the CNTs length, the aspect ratio, the polymer nature, the alignment of the CNTs, the composite processing methods. These conductive materials give rise to Joule effect due to the current flow and also show a piezoresistive behavior when an external stress is applied. These are the features allowing these materials to be called smart since these can be employed as self-responsive materials.

In this project an efficient Molecular Dynamics (MD) simulation scheme, based on hybrid particle-field representation called Molecular Dynamics-Self Consistent Field (MD-SCF)[2] is employed. We investigate the percolation threshold and electrical conductivity of nanocomposites in homopolymer matrices using parallel hybrid particle-field MD-SCF simulations[3]. Self-assembly processes of flexible rod-like particles (mimicking CNTs) dispersed in polymer phases have been examined at different concentrations near percolation. The obtained CNT morphologies have then been used to calculate the conductivity behavior using a resistor network approach.[4]

We deeply investigate such morphologies to catch the dependence from them of Joule heating efficiency. We also investigate concentration and applied voltage impact on it. The optimization of the Joule heating power performances is very important since it has to be a compromise between the optimal working conditions and the limits imposed by the practical applications such as maximum applicable voltage or maximum amount of CNTs in the polymer matrix without altering the mechanical properties. We performed several MD simulations to investigate the importance of assembled vs. random CNT conformations, showing that an optimized morphology can increase the Joule heat produced. We also evaluate piezoresistive behavior on different morphologies.

## 2 Methods and Results

The computing resources and the related technical support used for this work have been provided by CRESCO/ENEAGRID High Performance Computing infrastructure and its staff[5]. CRESCO/ENEAGRID High Performance Computing infrastructure is funded by ENEA, the Italian National Agency for New Technologies, Energy and Sustainable Economic Development and by Italian and European research programmes, see <http://www.cresco.enea.it/english> for information.

The simulation of CNTs enforced polymer matrices at the nanoscale is computationally expensive for both the sizescale ( $> 100$  nm) and the timescale necessary to simulate the assembly process of the CNTs. For this reason we employed the hybrid MD-SCF approach, based on the evaluation of the non-bonded force and its potential between atoms of different molecules by the evaluation of an external potential dependent on the local density at position  $\mathbf{r}$ . According to the SCF theory, a many-body problem such as molecular motion can be reduced to a problem of deriving the partition function of a single molecule with an external potential  $V(\mathbf{r})$ . Then, the nonbonded force between atoms of different molecules can be obtained from a suitable expression of  $V(\mathbf{r})$  and its derivatives. In the framework of SCF theory, a molecule is regarded to be interacting with the surrounding molecules through a mean field, rather than direct interactions among the molecules. Assuming that the density dependent interaction potential  $\mathbf{W}$ , where each species is specified by the index  $K$ , takes the following form:

$$W[\{\phi_K(\mathbf{r})\}] = \int d\mathbf{r} \left( \frac{k_B T}{2} \sum_{KK'} \chi_{KK'} \phi_K(\mathbf{r}) \phi_{K'}(\mathbf{r}) + \frac{1}{2\kappa} \left( \sum_K \phi_K(\mathbf{r}) - \phi_0 \right) \right)$$

where  $\phi_K(\mathbf{r})$  is the coarse-grained density of species K at position  $\mathbf{r}$  and  $\chi_{KK'}$  are the mean field parameters for the interaction of a particle of type K with the density fields due to particles of type K', it can be shown using the so-called saddle point approximation that the external potential is given by:

$$V_K(\mathbf{r}) = \frac{\delta W[\{\phi_K(\mathbf{r})\}]}{\delta \phi_K(\mathbf{r})} = k_B T \sum_{K'} \chi_{KK'} \phi_{K'}(\mathbf{r}) + \frac{1}{\kappa} \left( \sum_K \phi_K(\mathbf{r}) - \phi_0 \right)$$

All simulations (MD-SCF) reported in the present work have been performed using the parallelized version of the OCCAM MD code [6]. The system is composed by an insulator part (polymer matrix) and conductive part (CNTs).

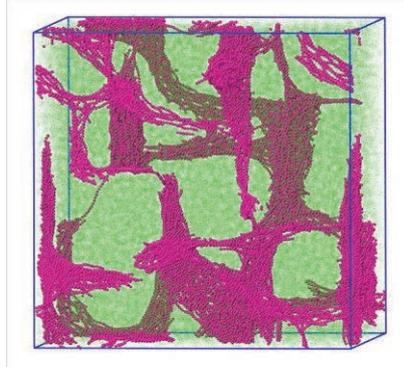


Figure 1 – schematic picture of a configuration of assembled CNTs (purple) in a polymer matrix (green).

The calculation of the CNT resistor network conductivity has been performed by explicitly calculating the resistance contributions coming from both the direct contacts among CNTs (intrinsic resistance) and the one coming from the electrons tunneling (contact resistance) among CNTs closer than a threshold distance under which the electron hopping is possible. In this way we know when the material becomes conductive and can start modelling the Joule heating power:

$$P = I^2 R$$

where R is the local resistance and I is the current intensity.

We propose a modelling strategy to simulate the Joule effect and its heat diffusion at the nanoscale for the first time for these systems. [7] We locally employ the percolation law to evaluate local resistance (depending on local CNT concentration) and by employing the Joule heating power (shown before), we evaluate local temperature.

These temperatures are added to the resolution of the heat diffusion equation solved numerically by using the finite difference method approximation:

$$T_{new(i)} = T_{old(i)} + \Delta t D \left( \sum_{\alpha} \frac{T_{i+1}^{\alpha} - 2T_i^{\alpha} + T_{i-1}^{\alpha}}{\Delta r_{\alpha}^2} \right) + T_{Joule}(i)$$

this equation is solved on a grid (in which the explicit system is translated), where the information on the CNT local concentration is kept. In the equation,  $D$  is the thermal diffusivity, the subscript is referred to the grid elements,  $\Delta_{ra}$  is the grid spacing in  $a$  direction.

We can follow the temporal evolution of the temperature of the system during the time and from it we calculate the heating rate (Celsius/s) to evaluate the performance of the system, the dependence on the applied voltage, concentration and morphology.

In Figure 2 the heating rate is shown as function of the concentration at different voltages: the discrepancy due to the CNT content increases with the applied voltage. Moreover, more heat is generated with a larger CNT content.

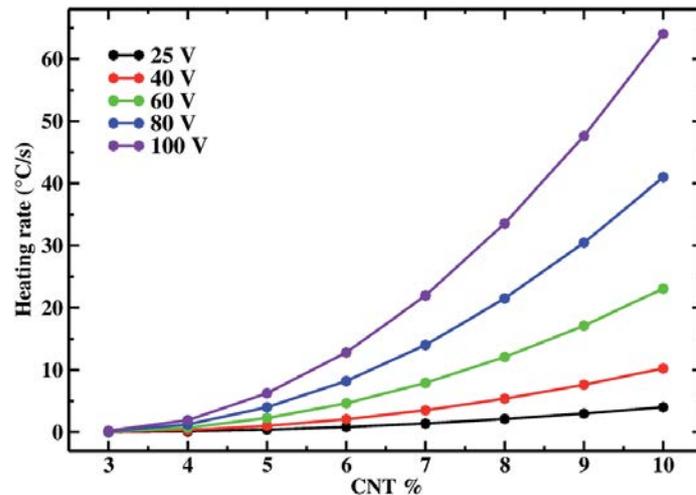


Figure 2 – heating rate as function of CNT concentration for different applied voltages.

We also observed that assembled morphologies are more performing than random ones since CNT networks are optimized and give rise to larger heating rates.[7]

We also successfully compare our simulated heating rates with experimental ones obtaining a very good agreement.

In Table 1 we show the comparison with experimental results: a good agreement is found, moreover the effect of the morphology is also shown: higher heating rates are found with assembled morphologies compared to randomly dispersed CNTs.

Voltage (V)	Random	Monomer	Polymer	Expt. [8]
12	0.033	0.8	0.67	0.16
24	0.13	3.22	2.67	0.6

Voltage (V)	Random	Monomer	Polymer	Expt. [9]
15	0.052	1.26	1.04	1.5
17.5	0.071	1.71	1.42	2.05
20	0.092	2.24	1.85	3.28

Voltage (V)	Random	Monomer	Polymer	Expt. [10]
30	0.17	4.28	3.55	0.78
40	0.31	7.6	6.32	2.33
50	0.48	11.88	9.87	3.88
60	0.69	17.11	14.12	5.43
70	0.94	23.28	19.35	7.75
80	1.23	30.42	23.28	10.08
90	1.56	38.49	31.99	13.18
100	1.92	47.52	39.49	16.28

Table 1 – calculated heating rates compared to experimental ones.

In Figure 3 we show the change of electrical resistance due to applied strain (compression or elongation). In agreement with experimental trends, larger discrepancies are observed for CNT concentrations close to the percolation threshold.[11]

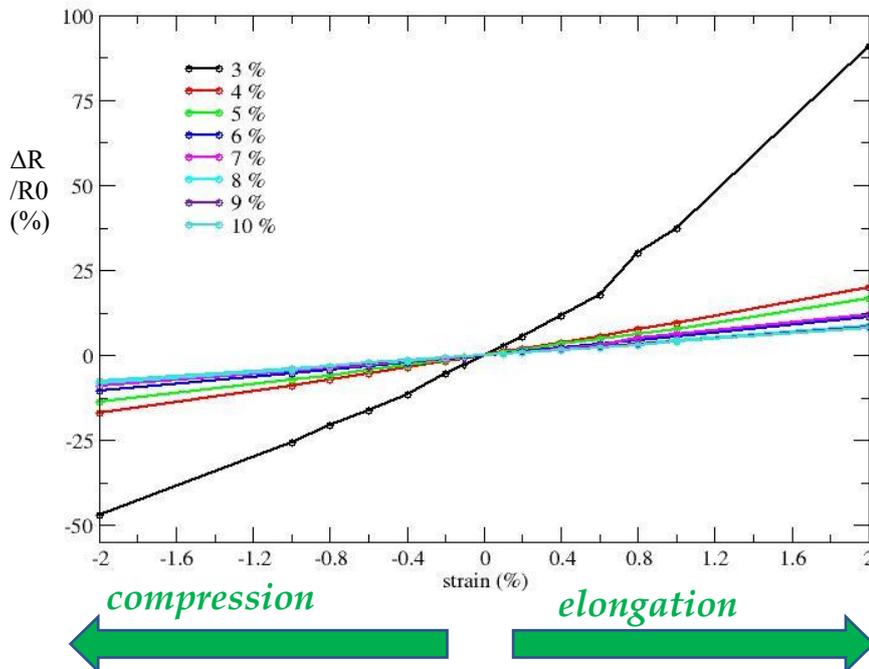


Figure 3 – piezoresistive behavior as function of applied elongation or compression for different CNT concentrations.

## References

- [1] S. Iijima, *Nature*, **354**, 56-58 (1991).
- [2] G. Milano and T. Kawakatsu, *J. Chem. Phys.*, **130**, 214106 (2009).
- [3] Y. Zhao, M. Byshkin, Y. Cong, T. Kawakatsu, L. Guadagno, A. De Nicola, N. Yu, G. Milano, B. Dong, *Nanoscale*, **8**, 15538-15552 (2016).
- [4] M. S. Byshkin and A. A. Turkin, *J. Phys. A: Math. Gen.*, **38**, 5057 (2005).
- [5] F. Iannone et al., "CRESCO ENEA HPC clusters: a working example of a multifabric GPFS Spectrum Scale layout," 2019 International Conference on High Performance Computing & Simulation (HPCS), Dublin, Ireland, 2019, pp. 1051-1052, doi: 10.1109/HPCS48598.2019.9188135
- [6] Y. Zhao, A. DeNicola, T. Kawakatsu and G. Milano, *J. Comput. Chem.*, **33**, 868–880 (2012).
- [7] G. Donati et al. *Nanoscale Adv.*, **2**, 3164 (2020).
- [8] A. Dorigato, M. Brugnara and A. Pegoretti, *Adv. Polym. Technol.*, **37**, 1744 (2018).
- [9] L. Sangroniz, A. Sangroniz, M. Fernandez, A. Exteberria, A. J. Muller and A. Santamaria, *Polymers*, **11**, 1180 (2019).
- [10] Y. G. Jeong and G. W. Jeon, *ACS Appl. Mater. Interfaces*, **5**, 6527 (2005).
- [11] G. Donati et al. *Manuscript in preparation*.

# AB-INITIO CALCULATIONS OF CATHODE MATERIALS FOR RECHARGEABLE BATTERIES

Adriano Mosca Conte<sup>1</sup>, Francesco Buonocore<sup>2</sup>, Paola Gori<sup>3</sup>, Massimo Celino<sup>2</sup>,  
Pier Paolo Prosini<sup>5</sup> and Olivia Pulci<sup>5</sup>

<sup>1</sup>*CNR-ISC, Unit Sapienza, Piazzale Aldo Moro 5, I-00185 Rome, Italy*

<sup>2</sup>*ENEA-TERIN-ICT Division, C.R. Casaccia, Via Anguillarese 301, I-00123 Rome, Italy*

<sup>3</sup>*Dipartimento di Ingegneria, Universita' Roma Tre, Via della Vasca Navale 79, I-00146 Roma, Italy*

<sup>4</sup>*ENEA-TERIN-PSU Division, C.R. Casaccia, Via Anguillarese 301, I-00123 Rome, Italy*

<sup>5</sup>*University of Rome Tor Vergata, Via della Ricerca Scientifica 1, I-00133 Rome, Italy*

**ABSTRACT.** We present a study based on first principles calculations of the electronic properties of the  $\text{MeMnO}_2$  base crystal (with  $\text{Me} = \text{Li}$  or  $\text{Na}$ ). The formation energies were evaluated, and the energy of the valence and conduction electronic bands of the base crystal were determined, in order to identify the phenomena underlying the electron transport within the material. The final goal is to use these crystals as cathode material for rechargeable sodium or lithium batteries.

## 1 Introduction

In the last years, the research to replace the widespread Lithium batteries (LIBs) by sodium-ion batteries (NIBs) has received more and more attention, mainly because the limited Li availability makes it difficult to meet the forecast growth in market demand. In contrast, Na is widely distributed around the world and is inexpensive, making NIBs a promising alternative for LIBs. Sodium-ion batteries are rechargeable batteries analogue to the more known and nowadays worldwide used Lithium-ion batteries. In both devices, the cathode is a solid state material containing the alkaline ions, the anode is a solid state material not necessarily containing the alkaline ions, and the electrolyte is a solution containing the alkaline ions. Similar to all lithium-ion cathodes, sodium-ion cathodes also store sodium via intercalation reaction mechanism. Owing to their high operating potentials and high capacities, cathodes based on sodium transition metal oxides have received increasing interest.

In this report we analyze the structure and electronic properties of the basic P2- $\text{NaMnO}_2$  crystals (hereinafter also referred to as P2-Na) and T2- $\text{Li}_{0.67}\text{Ni}_{0.33}\text{Mn}_{0.67}\text{O}_2$  (T2-Li) with calculations based on the density functional theory (DFT). The article by Prosini and coauthors [1] showed that the cathode of the sodium battery with formula  $\text{NaLi}_{0.2}\text{Ni}_{0.25}\text{Mn}_{0.75}\text{O}_\delta$  ( $\delta \approx 2$ ) which crystallizes according to the hexagonal structure P2, is able to reversibly cycle sodium ion. The aim of this study is to

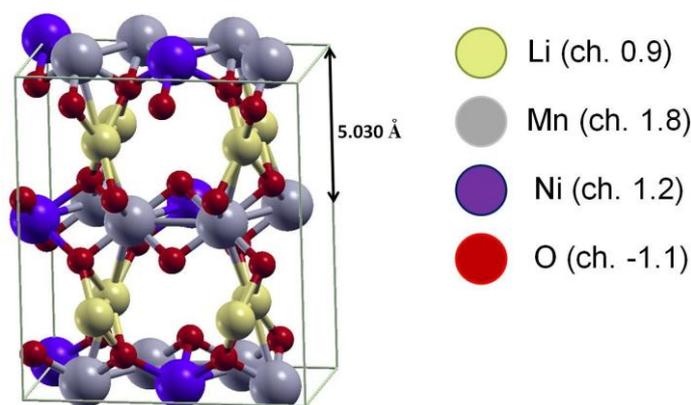
characterize the base crystal of the cathode group of the sodium batteries. Comparison with a similar material containing Li ion allows to determine the differences in performance between sodium and lithium batteries. The calculations were performed on the CRESCO6 computing cluster of ENEA's ENEAGRID computing infrastructure. The energies of formation of the crystals are calculated to evaluate their stability, and the energy of the electronic valence and conduction bands of the base crystal has been determined in order to identify the phenomena underlying the electron transport within the material. In fact, the presence of conduction bands near the Fermi level facilitates the electrons transfer within the material by lowering the electrical resistance, and therefore improves the electrochemical performance of the material. In our calculations we take into account the magnetism of the material to identify the electronic configurations with the lowest energy, and therefore the most stable ones. This information will be used in the near future research to evaluate how small amounts of dopants etheroatoms (such as Ti, Zn and Cu) can influence the electronic structure of the base crystal.

## 2 Results and Discussion

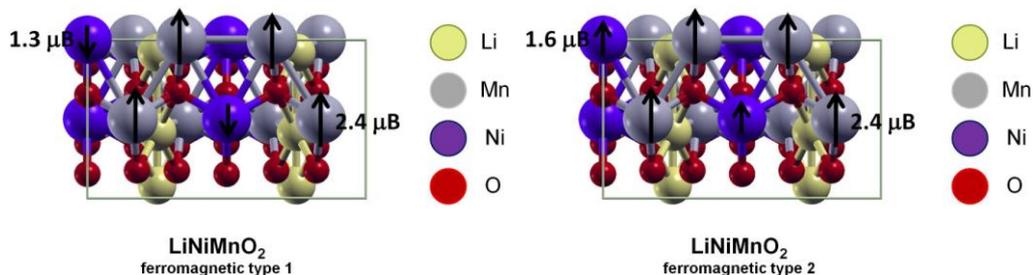
Calculations were performed within Density Functional Theory (DFT) [2,3] using the Quantum Espresso software [4]. The exchange-correlation was approximated with the functional GGA PBE [5]. Van der Waals interactions are included through the empirical correction of GRIMME [6].

### 2.1 T2 $\text{Li}_{0.67}\text{Ni}_{0.33}\text{Mn}_{0.67}\text{O}_2$ as cathode material for lithium batteries

The geometry of the primitive cell of  $\text{Li}_{0.67}\text{Ni}_{0.33}\text{Mn}_{0.67}\text{O}_2$  was obtained following the indications of Paulsen et al. [7] based on the results of X-ray diffraction measurements. The cell has T2 symmetry, contains 44 atoms (24 O, 8 Mn, 4 Ni, 8 Li) and is shown in Figure 1. Each Lithium atom is bonded to 4 oxygens, with a bond distance of approximately  $d_{\text{Li}-\text{O}} = 2.4 \text{ \AA}$ . The Lithium ions are interposed between the metal planes (Ni or Mn) which are about  $5 \text{ \AA}$  apart. Two ferromagnetic configurations were obtained, almost with the same geometry but with the spins ordered differently, as can be seen in Figure 2.



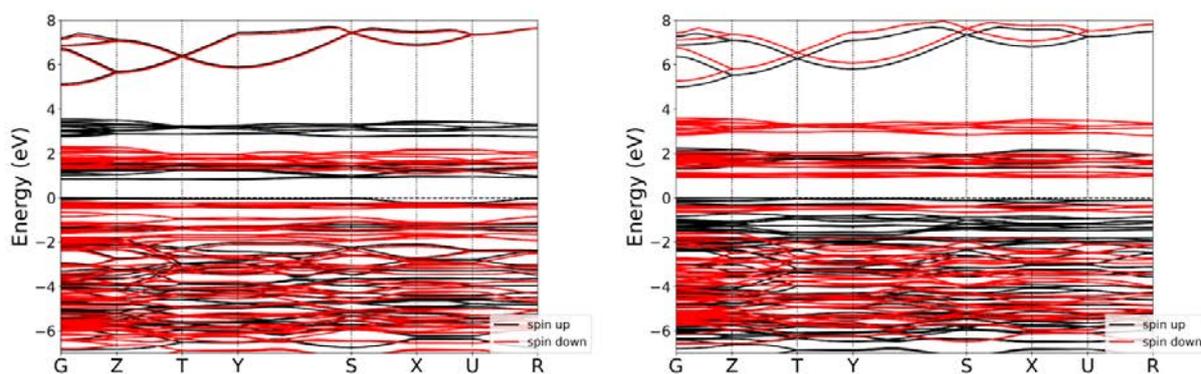
**Fig. 1:** Sticks and balls representation of the unit cell of  $\text{Li}_{0.67}\text{Ni}_{0.33}\text{Mn}_{0.67}\text{O}_2$  (T2-Li).



**Fig.2:** Top view of the atomic configuration and the spin magnetic moments of the unit cell of  $\text{Li}_{0.67}\text{Ni}_{0.33}\text{Mn}_{0.67}\text{O}_2$  (T2-Li) in the two found ferromagnetic configurations. For sake of clarity, the spin magnetic moments lower than  $0.1 \mu\text{B}$  are not shown such as those of O ( $0.03 \mu\text{B}$ ) and Li ( $0.002 \mu\text{B}$ ) atoms.

The formation energy was calculated from the difference between the total energy of the cell and that of its constituents, taken in the stable form at room temperature. The formation energy per atom of the T2-Li cell is equal to  $-1.9 \text{ eV/atom}$ . Being negative, the crystal is stable.

The band structure of the T2-Li crystal is shown in Figure 3. Both phases have an energy gap between full and empty bands. The predicted energy gap with the PBE exchange-correlation is  $0.85 \text{ eV}$  in the FM type 1 phase and  $0.9 \text{ eV}$  in the FM type 2 phase. The presence of partially occupied conduction bands is generally important for electrochemical performance. Indeed the presence of conduction bands close to the Fermi level facilitates electronic transfer within the material by lowering the electrical resistance. The presence of the gap indicates that the material is semiconductor. The low dispersion of the conduction bands close to the Fermi level, which therefore appear to be rather flat, indicates however a not small electrical resistance of the crystal. In fact, wide dispersion bands are connected to good conductive properties, and vice versa low dispersion bands imply low electrical conduction.



**Fig.3:** Electronic structures of  $\text{Li}_{0.67}\text{Ni}_{0.33}\text{Mn}_{0.67}\text{O}_2$  (T2-Li) in FM phases of type 1 (left panel) and 2 (right panel). Black bands correspond to spin up, red bands correspond to spin down.

## 2.1 P2-NaMnO<sub>2</sub> as cathode material for sodium batteries

The P2 phase of the MnNaO<sub>2</sub> (or P2-Na) structure as determined by X-ray diffraction by Paulsen and co-authors was taken as a preliminary model [7]. The structure consists of 8 overlapping hexagonal layers, as can be seen in Figure 4 and as described in Ref. [8]. Both the ferromagnetic (FM) and anti-ferromagnetic (AF) configurations obtained by placing the spin magnetic moments of the non-coplanar manganese atoms in the opposite direction were considered. The distance between the metal layers (Mn) is 5.414 Å in the ferromagnetic phase and 5.415 Å in the anti-ferromagnetic one. The fact that the non-coplanar Mn atoms are quite distant from each other and are interposed with 3 atomic layers of different species means that the difference in total energy between the ferromagnetic and antiferromagnetic phase is relatively low: the total energy of the FM configuration is lower than 1 meV / atom. It will be subject of research to find out whether the structure remains ferromagnetic even at lower concentrations of Na. The formation energy per atom of the P2-Na cell is equal to -1.7 eV / atom, so the crystal is stable.

We show the band structure (both FM and AF) in Figure 5. The analysis of the states of the lower energy phase (FM), not shown here, highlights that ferromagnetism is mainly induced by Mn atoms, and to a lesser extent by O atoms. This is in agreement with the atomic spin magnetic moments. The system is metallic in that it has partially occupied conduction bands. From PDOS it can be inferred that these bands are associated with Mn-3d and O-2p states in strong hybridization between them. The metallic character of the P2-Na crystal is important for electrochemical performance. In fact, the presence of conduction bands close to the Fermi level facilitates electronic transfer within the material by lowering the electrical resistance. For P2-Na a ferromagnetic state with a total magnetization of 8.0 μB / cell was found.

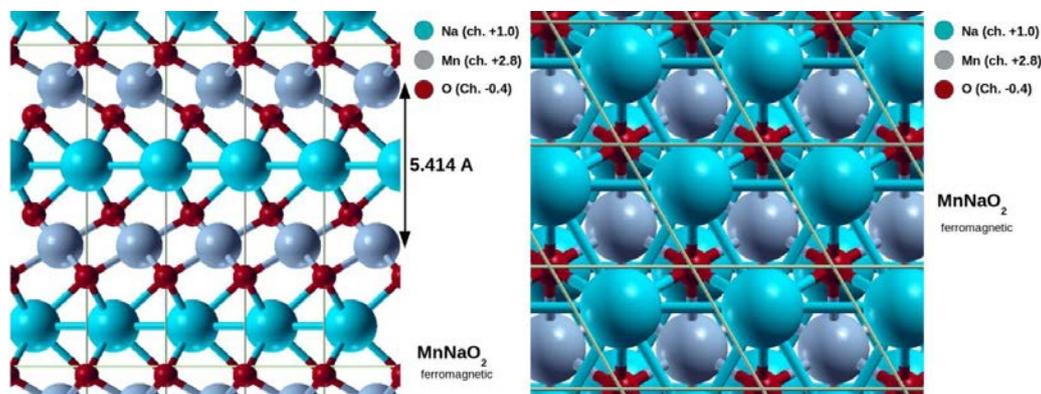
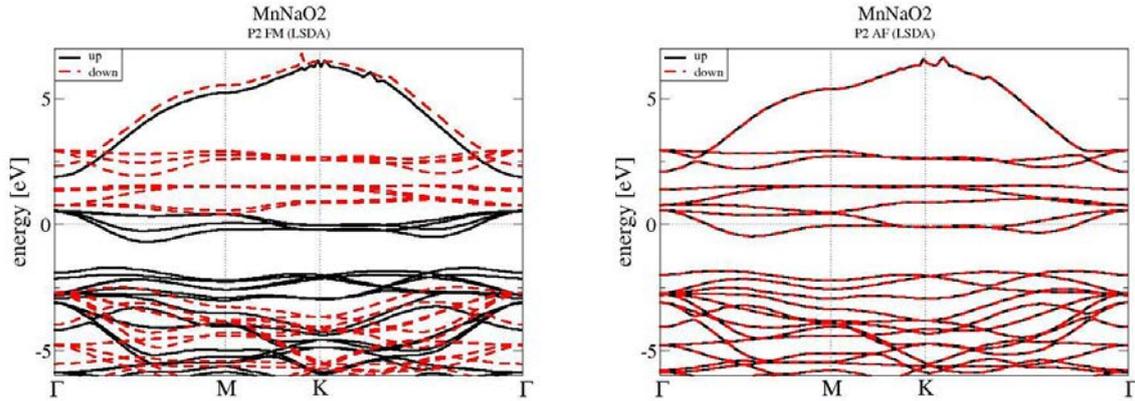


Fig.4: Side and top view of the MnNaO<sub>2</sub> (P2-Na) cell in the ferromagnetic phase.

## 3 CONCLUSIONS

In this study we addressed the determination of the electronic properties of the MnNaO<sub>2</sub> (P2-Na) and Li<sub>0.67</sub>Ni<sub>0.33</sub>Mn<sub>0.67</sub>O<sub>2</sub> (T2-Li) systems: the equilibrium geometry, the formation energies and the electronic band structure. At least two magnetic phases have been found for each system, which seem

to coexist at room temperature as their difference in energy, calculated per atom, is less than  $k_B T$  (about 26 meV). Two important findings are that the T2-Li structure is semiconductor, while the P2-Na is metallic. The latter result will be validated in next works through the use of more appropriate exchange and correlation potentials, such as DFT + U or hybrid potentials. Both structures have negative formation energy and therefore are stable. These results represent the basement to the near future research, in particular to understand the effects of doping on the electronic properties and stability of these materials.



**Fig.5:** Electronic structure of  $\text{MnNaO}_2$  (P2-Na) in the ferromagnetic phase (left panel) and antiferromagnetic (right panel). The black bands are occupied by spin up electrons, the red ones with spin down. In both cases the material is metallic.

## Acknowledgement

The computing resources and the related technical support used for this work have been provided by CRESCO/ENEAGRID High Performance Computing infrastructure and its staff [9, 10]. CRESCO/ENEAGRID High Performance Computing infrastructure is funded by ENEA, the Italian National Agency for New Technologies, Energy and Sustainable Economic Development and by Italian and European research programs. This research has been funded by MiSE Piano Triennale di Realizzazione 2019-2021, “Ricerca di Sistema elettrico”.

## References

- [1] P. P. Prosini, M. Carewska, C. Cento, et al., Tin-Decorated Reduced Graphene Oxide and  $\text{NaLi}_{0.2}\text{Ni}_{0.25}\text{Mn}_{0.75}\text{O}$  as Electrode Materials for Sodium-Ion Batteries. *Materials* (Basel), volume 12 (2019), p. 1074
- [2] P. Hohenberg and W. Kohn, Inhomogeneous Electron Gas, *Phys. Rev. B*, volume 136 (1964), p. 864

- [3] W.Kohn, L. J. Sham, Self-Consistent Equations Including Exchange and Correlation Effects. *Physical Review*, volume 140 (1965), pp. A1133–A1138
- [4] P. Giannozzi, S. Baroni, N. Bonini, et al., QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials. *J. Phys. Condens. Matter.*, volume 21 (2009), p. 395502
- [5] J. P. Perdew, K. Burke and M. Ernzerhof, Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.*, volume 77 (1996), p. 3865
- [6] S. Grimme, Semiempirical GGA- type density functional constructed with a long- range dispersion correction. *J. Comp. Chem.*, volume 27 (2006), p. 1787
- [7] J. M. Paulsen, R.A. Donaberger, J.R. Dahn, Layered T2-, O6-, O2-, and P2-Type  $A_{2/3}[M'_{2+1/3}M_{4+2/3}]O_2$  Bronzes, A = Li, Na; M' = Ni, Mg; M = Mn, Ti, *Chem. Chem. Mater.*, volume 12 (2000), p. 2257
- [8] W. Kobayashi, A. Yanagita, T. Akaba, et al., Thermal Expansion in Layered  $Na_xMO_2$ . *Sci. Rep.*, volume 8 (2018), p. 3988
- [9] F. Iannone et al., CRESCO ENEA HPC clusters: a working example of a multifabric GPFS Spectrum Scale layout 2019 International Conference on High Performance Computing & Simulation (HPCS), Dublin, Ireland, 2019, pp. 1051-1052.
- [10] A. Mariano et al., Fast Access to Remote Objects 2.0 a renewed gateway to ENEAGRID distributed computing resources. *FGCS 94* (2019) 920-928.

# EXACT INVARIANT APPROACH TO THE INVESTIGATION OF NONLINEAR DYNAMICS OF CHIRPING ALFVÉN MODES

Sergio Briguglio<sup>1\*</sup>, Xin Wang<sup>2</sup>, Claudio Di Troia<sup>1</sup>, Giuliana Fogaccia<sup>1</sup>, Valeria Fusco<sup>1</sup> and  
Gregorio Vlad<sup>1</sup>

<sup>1</sup>*ENEA, Fusion and Nuclear Safety Department, C. R. Frascati, Via E. Fermi 45, 00044 Frascati (Roma), Italy*

<sup>2</sup>*Max-Planck-Institut für Plasmaphysik, Boltzmannstraße 2, Garching D-85748, Germany*

**ABSTRACT.** An exact invariant approach to the investigation of nonlinear dynamics of chirping Alfvén modes is described, and some results are shown.

## 1 The need of exact invariants

An exact invariant approach to the investigation of nonlinear dynamics of chirping Alfvén modes has been carried out by means of large scale simulations performed by the hybrid magnetohydrodynamic-gyrokinetic code HMGC [1]. Two main issues have to be addressed when studying such dynamics:

- the identification of the relevant resonances destabilising the Alfvén mode, both in the exponential-growth phase and in the nonlinear one
- the analysis of the nonlinear saturation mechanism. If exact invariants of the system can be identified, both problems can be faced in a more effective way, once such invariants are used as phase-space coordinates.

Concerning the former problem, we can distinguish whether the drive at a certain time comes from the same particles that were destabilising the mode at a previous time (if the maximum drive is associated to the same values of invariants) or, rather, from a different set of particles (in the opposite case).

Concerning the analysis of the nonlinear saturation mechanism, the adoption of invariants as phase space coordinates allows for simplifying the Vlasov equation, which describes the evolution of the particle distribution function. Indeed, gradients of this function along the directions corresponding to the invariant coordinates do not play any role in the equation, as they are multiplied by the (null) time derivatives of the invariants. This means that the phase space can be cut into slices orthogonal to those directions, with the mode-particle dynamics of each slice evolving independently of each other. In particular, we can concentrate our analysis on the slices yielding the larger drive in the different phases of the mode evolution.

## 2 Exact invariants of particle motion in presence of Alfvén modes

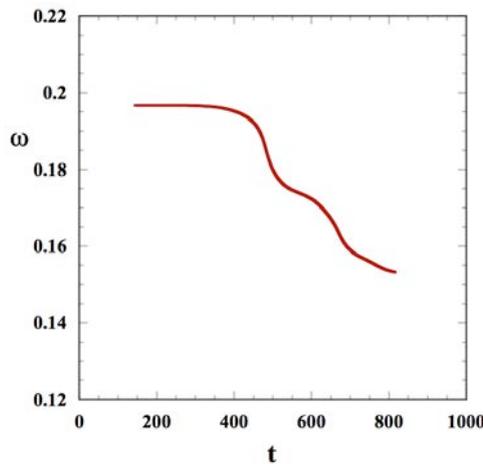
As far as modes characterised by a single toroidal number and a constant frequency are considered, two exact invariants can be identified: the magnetic momentum of the particle and the quantity yielded by the mode frequency times the toroidal angular momentum minus the toroidal mode

number times the particle energy. Unfortunately, the latter quantity is not invariant in the general case; in particular, in case of modes characterised by varying ("chirping") frequency.

A natural set of exact invariants is represented by the initial coordinates of the considered particle. While such set can hardly be used within an analytical approach, because of the complicate backward transformation relating the actual phase space coordinates to the initial ones, it is very easy to use within the numerical simulation approach: the initial coordinates of each simulation particle can be stored in arrays separate from those used for the current coordinates.

### 3 Results

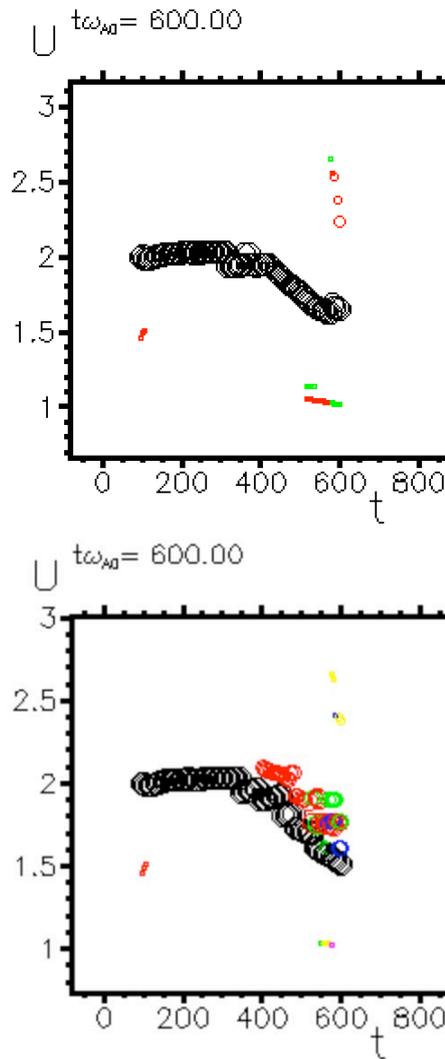
This approach has been applied to a specific case: an Energetic Particle Mode [2], driven unstable, in a Tokamak magnetic equilibrium, by the alpha particles produced by fusion reactions. The analysis of the dynamics of these modes is relevant in the field of controlled nuclear fusion research, as they could deteriorate the alpha particle confinement in next generation, reactor-like, Tokamaks. Figure 1 shows the time evolution of the mode frequency: after the exponential growth of the mode, during which the frequency remains constant, mode saturation is accompanied by a variation in frequency. Then, this is a suitable case for the adoption of the set of exact invariants (initial particle coordinates) described above.



**Fig. 1:** Time evolution of the mode frequency for the Energetic Particle Mode investigated in our simulations.

Figure 2 shows an application of this approach to the identification of the destabilising sets of particles. In the upper frame we see the time evolution of the parallel velocity corresponding to the maximum drive. The black larger circles represent the absolute maximum for each time. In this frame, the current value of the parallel velocity is considered. In principle, the decreasing of the maximum-drive parallel velocity could correspond to a progressive velocity loss of the same set of driving particles, or to a succession of different, slower and slower, driving sets. The lower frame,

which refers to the invariant coordinate (namely, the initial value of the parallel velocity), shows that the latter interpretation is the correct one.

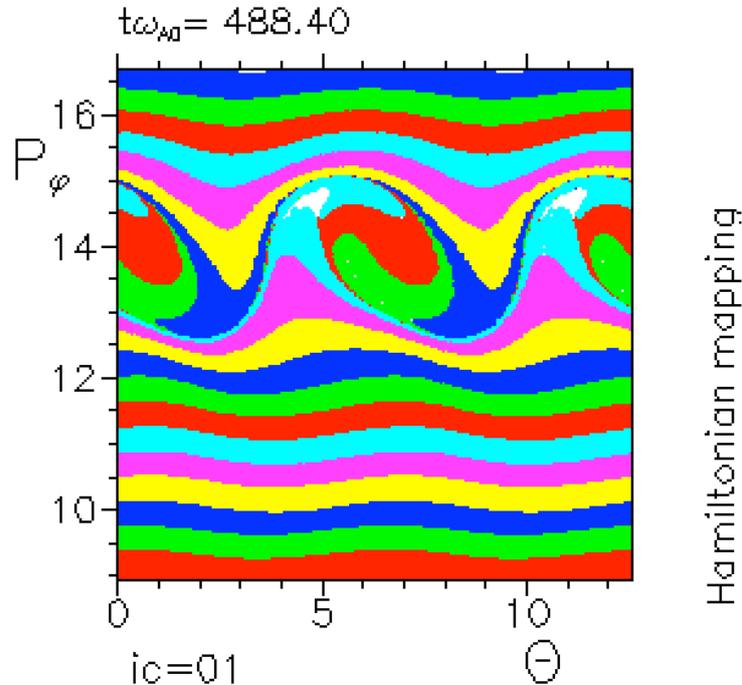


**Fig. 2:** Time evolution of the parallel velocity corresponding to the maximum drive. The black larger circles represent the absolute maximum for each time. In the upper frame, the current value of the parallel velocity is considered; in the lower frame, its initial value.

Figure 3 shows the formation of an island in the "wave-particle phase, toroidal angular momentum" space [3] due to the nonlinear interaction between an Alfvén mode and a population of test particles sampling the phase-space slice corresponding to the maximum drive during the exponential growth of the mode. Such island formation is responsible for the flattening of the density profile in the slice, with consequent depletion of the free-energy source.

The computing resources and the related technical support used for this work have been provided by CRESCO/ENEAGRID High Performance Computing infrastructure and its staff [4]. CRESCO/ENEAGRID High Performance Computing infrastructure is funded by ENEA, the Italian

National Agency for New Technologies, Energy and Sustainable Economic Development and by Italian and European research programmes, see <http://www.cresco.enea.it/english> for information.



**Fig. 3:** Formation of an island in the "wave-particle phase, toroidal angular momentum" space due to the nonlinear interaction between an Alfvén mode and a population of test particles sampling the phase-space slice yielding the maximum drive during the exponential growth of the mode.

## References

- [1] S. Briguglio, G. Vlad, F. Zonca, C. Kar. Hybrid magnetohydrodynamic-gyrokinetic simulation of toroidal Alfvén modes. *Phys. Plasmas* **2**, pp. 3711-3723, (1995)
- [2] L. Chen. Theory of magnetohydrodynamic instabilities excited by energetic particles in tokamaks. *Phys. Plasmas* **1**, p. 1519, (1994)
- [3] S. Briguglio, M. Schneller, X. Wang, C. Di Troia, T. Hayward, V. Fusco, G. Vlad and G. Fogaccia. Saturation of Alfvén modes in Tokamak plasmas investigated by Hamiltonian mapping techniques. *Nucl. Fusion* **57**, p. 072001, (2017)
- [4] F. Iannone et al., "CRESCO ENEA HPC clusters: a working example of a multifabric GPFS Spectrum Scale layout," 2019 International Conference on High Performance Computing & Simulation (HPCS), Dublin, Ireland, 2019, pp. 1051-1052, doi: 10.1109/HPCS48598.2019.9188135

# PEPTIDE BOND RECOGNITION VIA VIBRATION ASSISTED ELECTRON TUNNELING IN GRAPHENE NANODEVICE

Giuseppe Zollo \*

*Dipartimento di Scienze di Base e Applicate per l'Ingegneria- Sapienza Universit di Roma via A. Scarpa 14-16  
00161 Roma- ITALY*

**ABSTRACT.** Inelastic scattering due to vibrations is considered in the transversal tunneling current recognition of amino acids across a nano gap in graphene nanoribbons. Phonons and local vibrational modes assisted current tunneling is calculated by treating the electron-phonon scattering in the context of the lowest order expansion of the self-consistent Born approximation. We study the Gly and Ala homo-peptide cases as an example of very similar, small and neutral amino-acid that would be undistinguishable by means of standard techniques, such as the ionic blockade current, in real peptides. We show that all the inelastic contributions to the tunneling current are in the bias range  $0 V \leq V \leq 0.5 V$  and that they can be classified, from the atomistic point of view, in terms of energy sub-ranges they belong to. Peculiar fingerprints can be found for the typical configurations that have been recently found for peptide bonds recognition by tunneling current.

## 1 Introduction

Nano-devices are being considered for bio-molecule recognition and sequencing [1, 2, 3, 4, 5]. However the protein case is challenging and new methods for fast and reliable amino-acid recognition are mandatory to support the rising importance of proteomics [3, 4, 5, 6, 7, 8, 9, 10, 11, 12]. The amino-acid (AAs) recognition can be attained, at least in principles, by measuring the tunneling current flowing through the gap during the peptide translocation, provided a bias is applied between the electrodes [11, 13]: indeed, the tunneling current measured carries the chemical and physical signature of the piece of molecule that occupies the nano-gap. This scheme, however, requires a controlled translocation dynamics of the peptide or the protein (in its primary structure state) that is still an open technological issue. Several proposals to employ nano-gaps for biomolecule sensing have been found in the recent literature [14, 13]. In this context, nano-gaps in "2D materials" are particularly appealing because of the expected atomistic resolution of the tunneling current and graphene is of course the best candidate for that [15, 16, 17, 18, 19]. Recently a nano-gap device based on an array of semi-infinite graphene nano-ribbons (GNRs) (see Figure 1(a)) has been proposed showing the ability to sense single peptide bonds (PBs) along the peptide back-bone [19]. The above results have been obtained, for simple Glycine (Gly) and Alanine (Ala) template peptides, using the non equilibrium Green functions approach on the density functional theory basis (DFT-NEGF) and the well known Landauer-Büttiker formula [20, 21, 22, 23, 24] (see Section 2). However there are several approximations at this level of theory: among them, one of the most important is that only elastic scattering is considered and one might wonder how molecular vibrations and GNRs phonons affect the tunneling current. Inelastic tunneling current across the gap is expected to carry a larger amount of information related to the peculiar

---

\*Corresponding author. E-mail: giuseppe.zollo@uniroma1.it.

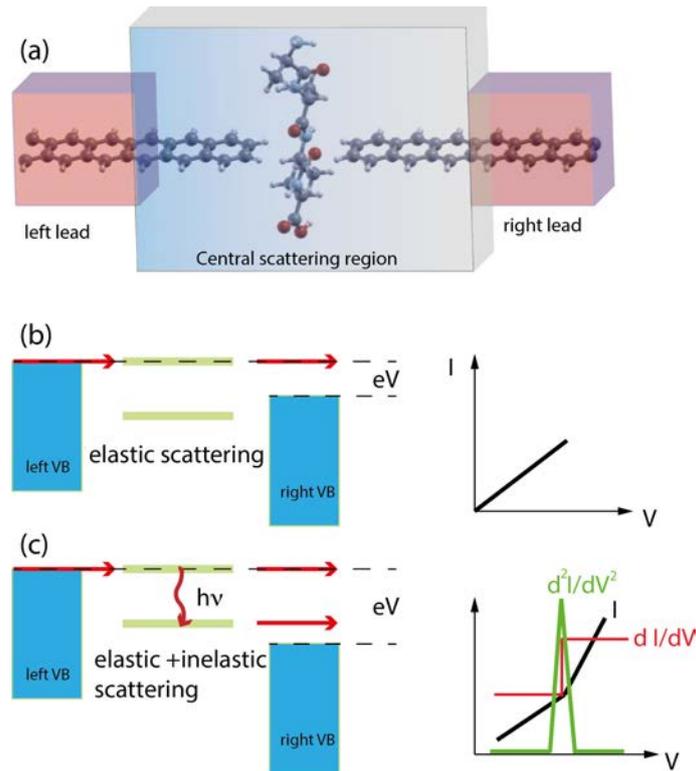


Figure 1: Schematics of nano-gap device between two semi-infinite ZGNR for peptide sequencing: the left(right) leads and the central scattering region are indicated (a). Schematics of the IETS phenomenology: for low bias only elastic scattering occurs (b) while for larger bias the tunneling probability increases due to inelastic scattering with phonon emission (c).

vibrational modes, the chemical formula and the charge states of the amino-acid side chains involved. Indeed, while elastic transport gives direct access to the electronic properties of the system (position and broadening of molecular levels for example), the study of the inelastic processes gives a deeper insight depending on the structure and dynamics of the molecule in the gap. Electrons can exchange energy with the vibrational modes of the molecule thus increasing or reducing the conductance of the junction (Fig 1(b)(c)). In our ideal sensing device the electrodes are made of two semi-infinite hydrogenated zig-zag GNRs (ZGNR) of the even type lying in the  $z$  direction (see Figure 1(a)). Zig- The nano-gap has a hydrogenated armchair pattern at the border and the distance between the two semi-infinite ZGNRs is  $5.0 \text{ \AA}$  (distance between H atoms). Although hydrogenated ZGNR are expected to have a gap in their ground state due to spin interactions [25, 26], it has been reported that the metallic unpolarized configurations is stable when a bias is applied and in presence of ballistic conduction [27, 28]. Moreover it is known that the spin polarized ground state undergoes a transition to semi-metallic conduction properties when the ZGNR is doped with N atoms [29] or when a transverse electric field is applied [30]. Then we have chosen the unpolarized metallic phase of the hydrogenated ZGNR as a paradigmatic case of metallic or half-metallic ideal 2D graphene based electrodes.

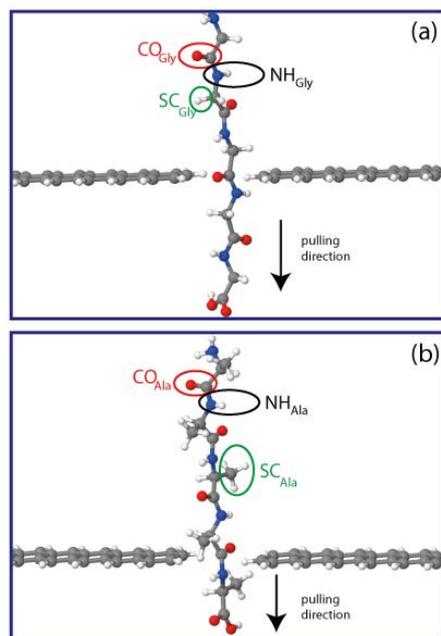


Figure 2: Sketches of the two atomistic systems employed to calculate the inelastic tunneling current signals for Gly (a) and Ala (b) omo-peptides. The translocation configurations have been obtained in water as detailed in the main text. We indicate the relevant groups of the two case treated, namely  $NH$ ,  $CO$  and  $SC$  and the pulling direction employed for the SMD simulations, namely from the N-terminal to the C-terminal of the peptides. Each peptide is made of 5 residues and the signals have been collected in the central part of the peptides, thus far from the peptide terminals.

## 2 Theoretical methods

We have focussed our attention on Ala and Gly omo-peptides because of the similarity of these two AAs that are small, neutral and with simple hydrophobic side chains. The configurations they assume in the ZGNR nano-gap have been obtained as follows: first we have performed a classical non-equilibrium steered molecular dynamics (SMD) simulations [31] at constant velocity and  $T = 300$  K in water: several passages across the nano-gap have been simulated to analyze and collect the most recurrent configurations (see Figure 2).

Then, NVT classical MD run at  $T = 300$  K in water have been performed to relax part of the peptide strain due to the previous SMD stage; the most probable configurations have been further relaxed at  $T = 0$  K in the context of the density functional theory (DFT) and with a large energy threshold in order to approach a local energy minimum. In this third stage the waters have been removed. With the above protocol we retain thermal fluctuations of the configurations occurring during the translocation at  $T = 300$  K in water (first and second stages) and correct them in the third stage for the wrong interactions, arising from classical model potential employed during the classical MD, between the GNR electrodes and the peptide. The final configurations have been employed to calculate the quantum transport and the inelastic scattering at low temperature ( $T = 4.2$  K). The vibration assisted tunneling,

that is schematically drawn in Figure 1(b)(c), is described by the Hamiltonian [32, 33]

$$\mathbf{H} = \mathbf{H}_e + \mathbf{H}_{ph} + \mathbf{H}_{e-ph} \quad (1)$$

$$\mathbf{H}_{ph} = \sum_{\lambda} \hbar\omega_{\lambda} \hat{b}_{\lambda}^{\dagger} \hat{b}_{\lambda} \quad \mathbf{H}_{e-ph} = \sum_{\lambda} \sum_{i,j} \mathbf{M}_{i,j}^{\lambda} \hat{c}_i^{\dagger} \hat{c}_j (\hat{b}_{\lambda}^{\dagger} + \hat{b}_{\lambda}) \quad (2)$$

with  $\mathbf{H}_{e(ph)}$  is the electrons (phonons) Hamiltonian,  $\hat{c}(\hat{b})$ , is the electron (phonon) annihilation operator,  $\lambda$  is a phonon label and  $i, j$  are electron labels. The electron-phonon coupling matrix  $\mathbf{M}^{\lambda}$  is:

$$\mathbf{M}_{i,j}^{\lambda} = \sum_{I\nu} \langle i | \frac{\partial \mathbf{H}_e}{\partial \mathbf{R}_{I\nu}} | j \rangle v_{I\nu}^{\lambda} \sqrt{\frac{\hbar}{2M_I \omega_{\lambda}}} \quad (3)$$

where  $v_{I\nu}^{\lambda}$  and  $\omega_{\lambda}$  are respectively the normalized eigenvector for the phonon mode  $\lambda$  and its frequency that are obtained with the frozen phonons approach. The vibrational density of states and eigen-vectors have been calculated in using a frozen-phonon approach. According to the Self-consistent Born approximation (SCBA), the electron self-energy due to the phonon scattering is:

$$\Sigma_{ph}^{\lessgtr}(\epsilon) = \sum_{\lambda} \mathbf{M}^{\lambda} [(n_{\lambda} + 1) G^{\lessgtr}(\epsilon \pm \hbar\omega_{\lambda}) + n_{\lambda} G^{\lessgtr}(\epsilon \mp \hbar\omega_{\lambda})] \mathbf{M}^{\lambda} \quad (4)$$

where  $G^{\lessgtr}$  are the unperturbed lesser/greater electron Green functions that are related to the spectral functions of the leads through  $G^{\lessgtr}(\epsilon) = \pm i [f(\epsilon \pm \hbar\omega_{\lambda} - \mu_L) A_L(\epsilon) + f(\epsilon \pm \hbar\omega_{\lambda} - \mu_R) A_R(\epsilon)]$ . A suitable approximation of the SCBA scheme can be attained to reduce the computational workload provided that the electron chemical potentials of the electrodes are  $eV = \mu_R - \mu_L = \pm \hbar\omega_{\lambda}$ . In that case it has been shown [33] that the total current  $I(V) = I_{el}(V) + I_{in}(V)$  at the lowest order expansion (LOE), obtained at the first order of  $\Sigma_{ph}^{\lessgtr}(\epsilon)$  (second order of  $\mathbf{M}^{\lambda}$ ), casts into the sum of two terms with the differential conductance  $\frac{\partial I}{\partial V}$  respectively symmetric and asymmetric with respect to the bias. In terms of the second derivative of the current with respect to the bias we have:

$$\frac{\partial^2 I}{\partial V^2} = \sum_{\lambda} \gamma_{\lambda} \frac{\partial^2 I^{sym}}{\partial V^2}(V, \hbar\omega_{\lambda}, T, n_{\lambda}) + \kappa_{\lambda} \frac{\partial^2 I^{asym}}{\partial V^2}(V, \hbar\omega_{\lambda}, T) \quad (5)$$

with  $\gamma_{\lambda} = \gamma_{el,\lambda} + \gamma_{in,\lambda}$  and the coefficients:

$$\begin{aligned} \gamma_{in,\lambda} &= Tr[\mathbf{M}^{\lambda} A_L^{\dagger}(\mu_L) \mathbf{M}^{\lambda} A_R(\mu_R)] - \mathbf{M}^{\lambda} G^a(\mu_R) \Gamma_L(\mu_R) A_R(\mu_R) \mathbf{M}^{\lambda} A_L(\mu_L) \} \\ B_{\lambda} &= \left\{ Tr[\mathbf{M}^{\lambda} A_R(\mu_L) \Gamma_L(\mu_L) G^r(\mu_L) \mathbf{M}^{\lambda} A_R(\mu_R) - \mathbf{M}^{\lambda} G^a(\mu_R) \Gamma_L(\mu_R) A_R(\mu_R) \mathbf{M}^{\lambda} A_L(\mu_L)] \right\} \\ \gamma_{el,\lambda} &= Im\{B_{\lambda}\} \quad \kappa_{\lambda} = 2Re\{B_{\lambda}\} \end{aligned} \quad (6)$$

In the previous formulas we have that  $G^{a(r)}(\epsilon)$  is the advanced (retarded) electron Green function and  $A_{L(R)}(\epsilon) = G^r(\epsilon) \Gamma_{L(R)}(\epsilon) G^a(\epsilon)$ .

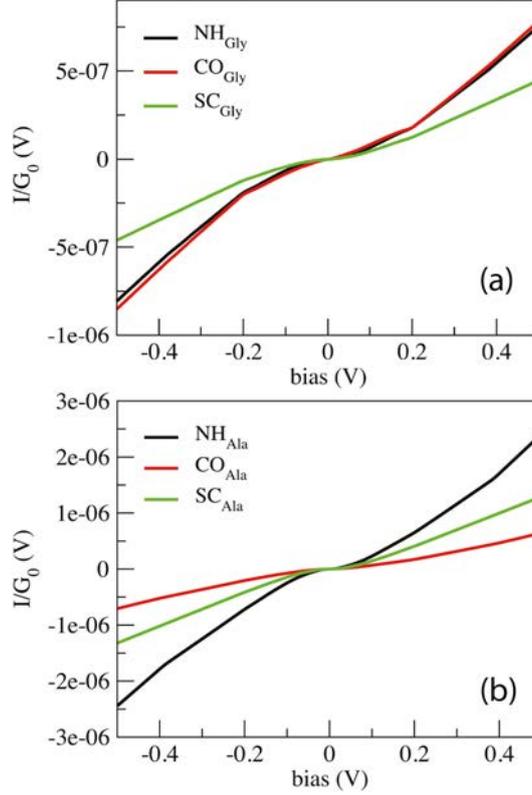


Figure 3: Non linear behavior of the tunneling current across the ZGNR nano-gap for Gly (a) and Ala (b) homo-peptides. The  $NH$ ,  $CO$  and  $SC$  curves indicate the signals collected from the three reference configurations, namely with the  $NH$ ,  $CO$  and the side chain groups in the middle of the gap.

The symmetric and asymmetric current universal functions are:

$$\begin{aligned}
 I^{sym} &= \frac{G_0}{2e} \sum_{\sigma=\pm 1} \sigma(\hbar\omega_\lambda + \sigma eV) \left[ \coth\left(\frac{\hbar\omega_\lambda}{2k_B T}\right) + -\coth\left(\frac{\hbar\omega_\lambda - \sigma eV}{2k_B T}\right) \right] \\
 I^{asym} &= -\frac{G_0}{2e\pi} \sum_{\sigma=\pm 1} \sigma(\sigma\hbar\omega_\lambda + eV) \ln \left| \frac{\sigma\hbar\omega_\lambda + eV}{\sigma\hbar\omega_\lambda} \right|
 \end{aligned} \tag{7}$$

In our system all the atoms of the central scattering region, that contains also part of the left and right ZGNR electrodes, have been allowed to vibrate. Transport and inelastic tunneling analyses have been performed using the Transiesta [24] and Inelastic [32] packages.

### 3 Results

In Figure 3 we report the current in  $G_0 = \frac{2e^2}{h}$  units calculated for the Gly and Ala homo-peptides. We have considered as reference positions the ones in Figure 2 corresponding to the relevant configurations found for the elastic case [19]. The non linear behavior due to inelastic phenomena is marked. For Gly one can observe localized abrupt slope changes while the non linear behavior seems more continuous in the Ala case. The different current values measured for  $NH$  and  $CO$  in the Ala case is due to the amount of the overlap between the side chain and the pseudo- $\pi$  and pseudo- $\pi^*$  orbitals of the nano-

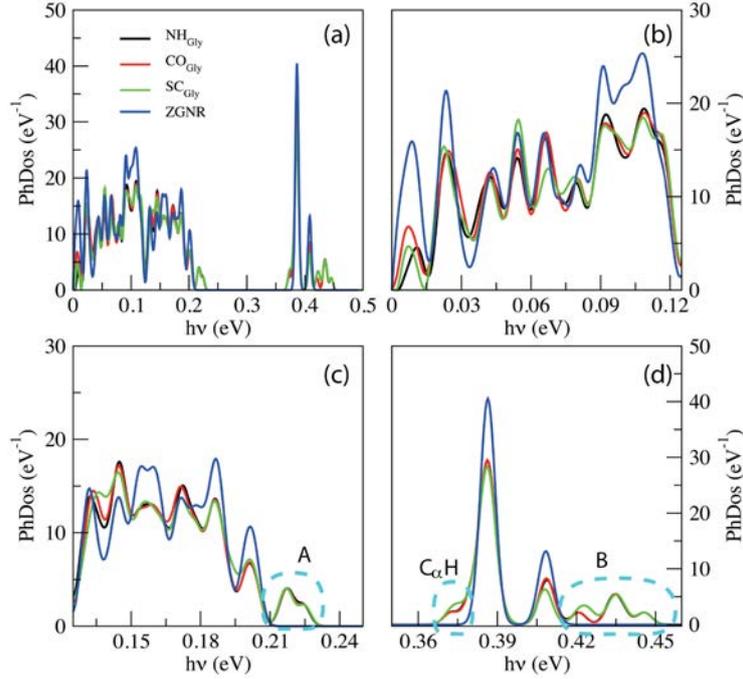


Figure 4: PhDoS of the three configurations analyzed for Gly compared to the PhDoS of the ZGNR without the peptide (a). Three regions of interest of the PhDoS spectra are expanded in (b), (c) and (d).  $NH_{Gly}$ ,  $CO_{Gly}$  and  $SC_{Gly}$  indicate the PhDoS obtained from three notable configurations, namely with the NH, CO and the side chain (SC) groups in the middle of the gap. The "A", "B" and  $C_{\alpha}H$  regions are indicated.

ribbons [19] that depends on the localization of the peptide in the gap along the y direction in the GNR plane, i.e. along the GNR width. Hence the difference of the current level is not meaningful in terms of the inelastic scattering processes.

### 3.1 Inelastic tunneling for Gly peptides

Comparing the phonon density of states (PhDoS) of the three configurations to the one of the ZGNR gap (without the peptide) reported in Figure 4, we find three energy ranges where the PhDoS of the systems does not overlap to the one of the electrodes. Scattering events in that regions, therefore, are due to contributions arising just from the local vibrational modes of the peptide (see Figure 4). The range  $0.42 \text{ eV} \leq hv \leq 0.45 \text{ eV}$  (region "B") contains vibrational modes of the C- and the N-terminals of the homo-peptide and stretching modes of the  $NH$  bonds along the peptide backbone. The region "A"  $0.21 \text{ eV} \leq hv \leq 0.23 \text{ eV}$  contains stretching modes of the  $CO$  bonds along the peptide backbone and contributions from the C and the N-terminals. The last region, labeled as  $C_{\alpha}H$ , is due to vibrational modes from the  $-CH_2$  group closest to, or in, the gap. In all the other regions of the PhDoS spectra, vibrational modes involve also the ZGNR in the device regions.

In Figure 5 we reports the IETS signals ( $IETS = \frac{d^2I/dV^2}{dI/dV}$ ) for the three reference configurations mentioned. Three regions are expanded in the same figure for  $V \geq 0.03 \text{ V}$  because, except for the large peak close to zero occurring in the  $SC$  case, the region close to  $V = 0$  does not show marked differences among the three cases. Comparing the IETS signals with the corresponding PhDoS and the values of

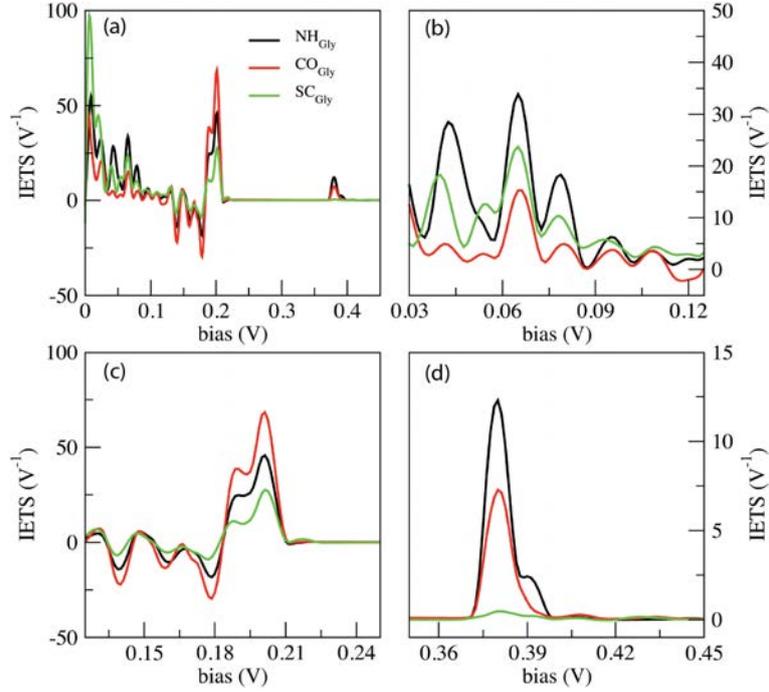


Figure 5: IETS signals calculated for the Gly omo-peptide (a). Three energy regions of the IETS spectra are expanded in (b), (c) and (d). The NH, CO and SC curves indicate the signals collected from three notable configurations, namely with the *NH*, *CO* and the side chain (*SC*) groups in the middle of the gap.

the  $\gamma$  and  $\kappa$  coefficients, four main regions of the IETS spectra can be defined with reference to Figure 5: the high energy "side chain" region in the bias range  $[0.36 \text{ V} \leq V \leq 0.42 \text{ V}]$  (Figure 5(d)), The "ZGNR" region for  $[0.19 \text{ V} \leq V \leq 0.21 \text{ V}]$  (Figure 5(c)), the "dips" region for  $[0.13 \text{ V} \leq V \leq 0.19 \text{ V}]$  (Figure 5(c)) and, finally, the "low energy" region for  $0.03 \text{ V} \leq V \leq 0.13 \text{ V}$  (Figure 5(b)(c)).

The energy values of the most important vibrational modes in the various regions of the spectra are reported in Table 1. The peaks observed in the "side chain region" for *NH* and *CO* are related to local vibrational modes of the closest side chain and  $C_{\alpha}H$  (see Figure 6(d)). The "ZGNR" region, instead is due to the scattering at the pieces of the ZGNR in the device region with modes in the ZGNR plane (Figure 6(c)). The "dips" region shows a complex scenario with contributions from both the phonons from the ZGNR and peptide local vibrations; here reflection scattering events (with  $\gamma_{\lambda} < 0$ ) contribute significantly (Figure 6(b)). In the "low energy" region the inelastic scattering is contributed by both peptide vibrational modes and ZGNR phonons with eigenvectors orthogonal to the ZGNR plane (Figure 6(a)) and only positive differential conductivity coming from symmetric terms ( $\gamma_{\lambda} > 0$ ) is measured.

The three reference configurations *NH*, *CO* and *SC* have different IETS features. The *SC* configuration has no scattering processes in the "side chain" region, shows an enhanced peak at very low energy ( $h\nu \approx 7 \text{ meV}$ ) (see Figure 5(a)) and the lowest IETS signal in the ZGNR region. The first of the three features is related to the lower coupling between the orbitals of the vibrating side chain and the pseudo- $\pi$  and pseudo- $\pi^*$  orbitals of the left and right electrodes. The large peak at very low energy is caused by an enhanced transmission from the central side chain due to long wavelength vibrational modes orthogonal to the ZGNR plane. Lastly, because the modes in ZGNR region are mainly characterized by in-plane vibrations of left/right ZGNR, the increase in the conductance is less effective for *SC* due

Table 1: Energy values in meV of the vibrational modes affecting the most important peaks and dips of the IETS spectra from Gly translocation configurations. Labels in parentheses indicate the symmetric or/and antisymmetric coefficients contributions and their sign:  $\gamma_\lambda^{+(-)}$  and  $\kappa_\lambda^{+(-)}$  stand for positive (negative) symmetric and antisymmetric coefficients respectively.

Region	<i>NH</i>	<i>CO</i>	<i>SC</i>
low energy	65.15 ( $\gamma_\lambda^+$ )	65.52 ( $\gamma_\lambda^+$ )	64.95 ( $\gamma_\lambda^+$ )
	78.2 ( $\gamma_\lambda^+$ )		
	79.2 ( $\gamma_\lambda^+$ )		
dips	139.5 ( $\gamma_\lambda^-$ )	139.7 ( $\gamma_\lambda^-$ )	132.55 ( $\gamma_\lambda^+ \kappa_\lambda^-$ )
	163 ( $\kappa_\lambda^+$ )	162.5 ( $\gamma_\lambda^+ \kappa_\lambda^+$ )	139.68 ( $\gamma_\lambda^- \kappa_\lambda^-$ )
	178.9 ( $\kappa_\lambda^-$ )	179.16 ( $\gamma_\lambda^-$ )	
ZGNR	199.5 ( $\gamma_\lambda^+$ )	200 ( $\gamma_\lambda^+$ )	200 ( $\gamma_\lambda^+$ )
	203.25 ( $\gamma_\lambda^+$ )	203 ( $\gamma_\lambda^+$ )	203.17 ( $\gamma_\lambda^+$ )
	203.3 ( $\gamma_\lambda^+ \kappa_\lambda^-$ )	203.42 ( $\gamma_\lambda^+ \kappa_\lambda^-$ )	203.46 ( $\gamma_\lambda^+ \kappa_\lambda^-$ )
side chain	379.28 ( $\gamma_\lambda^+$ )	379.45 ( $\gamma_\lambda^+$ )	
	380.34 ( $\gamma_\lambda^+$ )	384.6 ( $\gamma_\lambda^+$ )	
	390.65 ( $\gamma_\lambda^+$ )		

to the location of the side chain in the middle of the gap and the lower transmission across the peptide [19].

The main peak satellites in the "side chain" region allows the distinction of *NH* from *CO*. It is worth noticing that in the ZGNR region, the differential conductivity in the *CO* case is the largest due to the enhanced transport related to the off-plane side chain and *NH* groups [19].

### 3.2 Inelastic tunneling for Ala peptides

The Ala IETS signals are reported in Figure 7. The analysis of the PhDoS (not shown) evidences again three regions: the first two (regions "A" and "B") show the same features and energy ranges as the corresponding ones in the Gly case. A difference emerges for the  $C_\alpha H$  region that now is in the range  $0.39 \text{ eV} \leq h\nu \leq 0.4 \text{ eV}$  as a right-hand side satellite of the main PhDoS peak. This shift is related to the larger vibrational energy of the Ala side chain, a methyl group, with respect to the Gly one.

We can appreciate the quite larger scattering peaks in the very low energy range with respect to the Gly case. Moreover the peak heights in the ZGNR region here appear much smaller than for Gly which means that the peptide transmission is much lower than in the Gly case as already evidenced [19].

Similarly to the Gly case, the *SC* configuration differs from the other two mainly because of three aspects: no IETS signals in the "side chain" region, the large peak at very low energy and finally the less pronounced oscillation in the "dips" region. Distinction among the *NH* and *CO* is quite easy because of the additional peak for *CO* at  $h\nu = 438 \text{ meV}$  that does not appear in the *NH* configuration.

Some IETS features can be considered as peculiar of the two cases treated: one of them is the relative

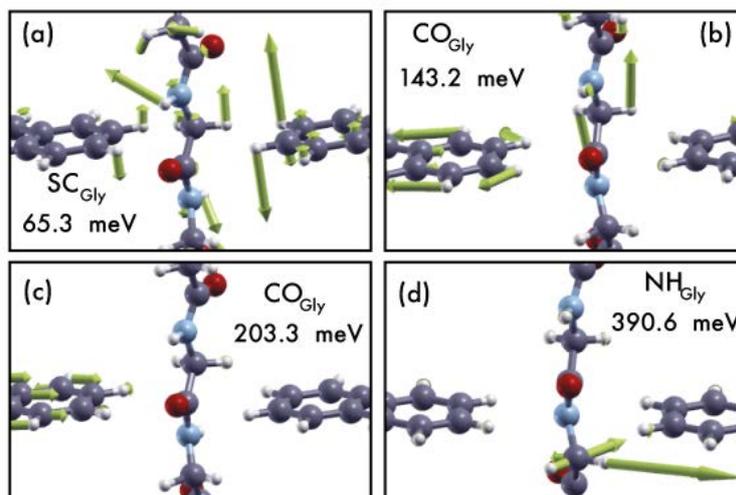


Figure 6: Four representative modes causing the inelastic scattering for the Gly case. The four modes have been chosen each in one of the four energy regions discussed in the text. In the insets are reported the configurations they belong to and their vibrational energy values.

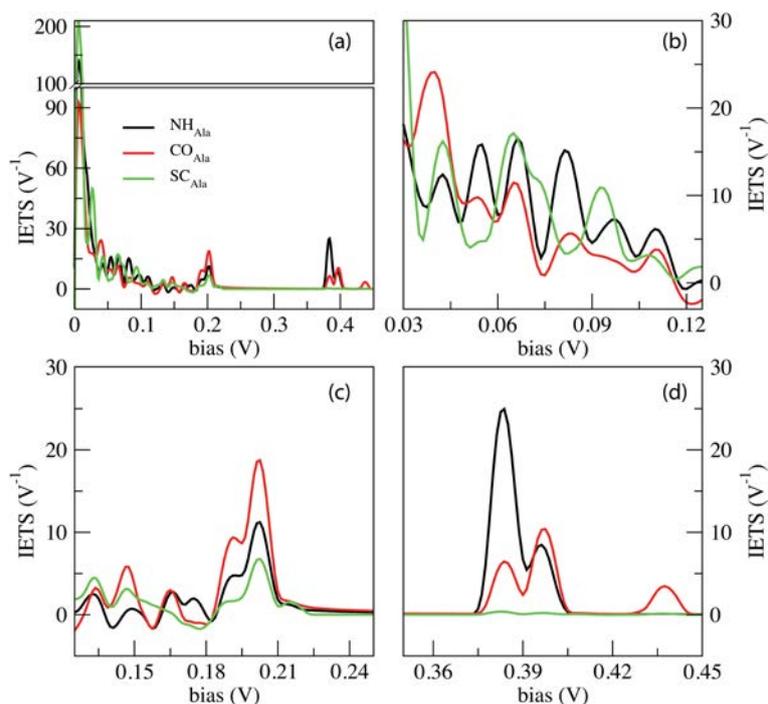


Figure 7: IETS signals from the phonon assisted the tunneling current across the ZGNR nano-gap with three special configurations for the Ala homo-peptide. The NH, CO and SC curves indicate the signals collected from three notable configurations, namely with the NH, CO and the side chain (SC) groups in the middle of the gap.

Table 2: Energy values in meV of the main peaks and dips of the IETS spectra from Ala translocation configurations. In parentheses are indicated if symmetric or/and antisymmetric coefficients contribute and their sign:  $\gamma_\lambda^{+(-)}$  and  $\kappa_\lambda^{+(-)}$  stand for positive (negative) symmetric and antisymmetric coefficients respectively.

Region	<i>NH</i>	<i>CO</i>	<i>SC</i>
low energy	9 ( $\gamma_\lambda^+$ )	11.4 ( $\gamma_\lambda^+$ )	8.7 ( $\gamma_\lambda^+$ )
		14.3 ( $\gamma_\lambda^+$ )	28.23 ( $\gamma_\lambda^+$ )
		39.38 ( $\gamma_\lambda^+$ )	43.7 ( $\gamma_\lambda^+$ )
	55.14 ( $\gamma_\lambda^+$ )	55.78 ( $\gamma_\lambda^+$ )	
	65.63 ( $\gamma_\lambda^+$ )	65.72 ( $\gamma_\lambda^+$ )	64.17 ( $\gamma_\lambda^+$ )
	80.55 ( $\gamma_\lambda^+$ )	93.25 ( $\gamma_\lambda^+$ )	
dips	133.22 ( $\gamma_\lambda^+$ )		136.22 ( $\gamma_\lambda^- \kappa_\lambda^+$ )
		146.03 ( $\gamma_\lambda^+ \kappa_\lambda^+$ )	146.5 ( $\gamma_\lambda^+$ )
	162.64 ( $\kappa_\lambda^+$ )	162.67 ( $\kappa_\lambda^+$ )	
		172 ( $\gamma_\lambda^-$ )	179.4 ( $\gamma_\lambda^-$ )
ZGNR		191.46 ( $\gamma_\lambda^+ \kappa_\lambda^+$ )	
	199.97 ( $\gamma_\lambda^+$ )	199.98 ( $\gamma_\lambda^+$ )	200 ( $\gamma_\lambda^+$ )
	202.9 ( $\gamma_\lambda^+ \kappa_\lambda^+$ )	202.8 ( $\gamma_\lambda^+ \kappa_\lambda^+$ )	202.9 ( $\gamma_\lambda^+$ )
	203.44 ( $\gamma_\lambda^+ \kappa_\lambda^-$ )	203.47 ( $\gamma_\lambda^+ \kappa_\lambda^-$ )	203.45 ( $\gamma_\lambda^+ \kappa_\lambda^-$ )
side chain	384.5 ( $\gamma_\lambda^+$ )	384.3 ( $\gamma_\lambda^+$ )	
	394.8 ( $\gamma_\lambda^+$ )		
	397.7 ( $\gamma_\lambda^+$ )	397.4 ( $\gamma_\lambda^+$ )	
		438 ( $\gamma_\lambda^+$ )	

importance of the vibrational scattering at very low bias with respect to the ZGNR region: in the Gly case the IETS peaks in these two regions are of the same order of magnitude while in the Ala case the IETS peaks in the ZGNR region is about one order of magnitude lower than the one at very low energy. This circumstance is due to the larger peptide transmission for Gly than for Ala in the ZGNR region. Moreover at very low energy, where the ZGNRs oscillate perpendicularly to their plane, Ala vibrational modes are more numerous due to the larger side chain thus enhancing the transmission through the side chains and  $C_\alpha H$  group. A second feature is the relative importance of the ZGNR and the side chain regions: in the Gly case the first one prevails while in the Ala case they are almost comparable. This is related to the different side chain sizes and the larger elastic transport across the "frozen" peptide in the Gly case, typical of the ZGNR region.

Looking at the fine details in the region of interest of Figs. 5, 7, it is worth emphasizing that "side chain" region is shifted to larger energy for Ala where a well separated double peak feature emerges for *CO* and *NH* (see tables 1 and 2), the energy shift being related to different vibrational modes of the side chains. Moreover, it is important to focus the reader's attention to the additional peak at  $h\nu = 438 \text{ meV}$  for  $CO_{Ala}$  that is related to the vibrations of the hydrogen bonded to the closest  $C_\alpha$  atom and of the closest (lower) *NH* group. Similar modes observed in the  $CO_{Gly}$  are not effective in terms of scattering due to the different orientation of the  $C_\alpha H$  group. The absence of a similar peak for  $NH_{Ala}$  tells us that this peak benefits of the tunneling current flowing across both the  $C_\alpha H$  and the *NH* groups that are out

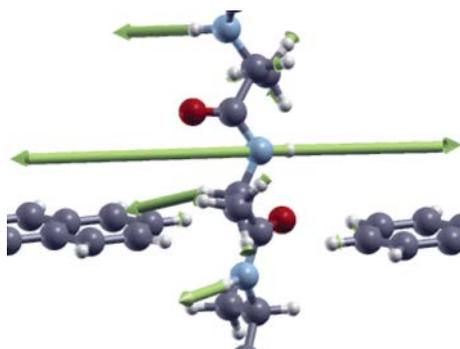


Figure 8: Vibrational mode corresponding to the inelastic scattering peak at  $h\nu = 438\text{meV}$  for the  $CO$  configuration of the Ala homo-peptide.

of the ZGNR plane: this circumstance just occurs for  $CO_{Ala}$ .

Therefore we can state that, although Gly and Ala have similar sizes and equal charge states, and are not distinguishable through ionic blockade current measurements, here can be discriminated.

There are, however, some interesting similarities in the cases treated: in both Gly and Ala cases the region "A" does not contribute to the inelastic scattering while the other two make the difference. Therefore tunneling current measurements with bias  $V \leq 0.35\text{ V}$  are affected only by vibrational modes that involve both the graphene nano-ribbons and the peptide. Interestingly there exists a bias range  $0.203\text{ V} \leq V \leq 0.35\text{ V}$  where the current shows a local linear behavior, with no additional inelastic events, for all the cases studied. By measuring the tunneling current with two different bias values in this range, slope values could be obtained and employed as reference for differential tunneling current measurement with larger bias  $V \geq 0.45\text{ V}$ ; this differential measurements would be affected, as already shown, just by the peptide vibrational modes, more specifically by the side chain and the  $C_\alpha H$  modes and, for the  $CO_{Ala}$  case, also by the closest  $NH$  stretching mode. Then, three measurement points in the  $I - V$  characteristic are supposed to be sufficient to obtain tunneling current signals affected just by the vibrational properties of the side chains and the  $C_\alpha H$  bond: in this way it would be possible to obtain specific fingerprints for side chain recognition also in the case of narrow electrodes. A more detailed discussion on the findings concerning vibration assisted tunneling current for amino acid recognition can be found in the recent literature [34]

## 4 Conclusions

The previous results shed some light on the inelastic scattering events taking place during tunneling current measurements across the gap of a graphene nano-ribbon when a peptide translocates across it. If the bias is in the "dips" or the "low energy" regions, then a very complex scenario arises where modes may cause reflection ( $\gamma_\lambda > 0$ ) or contribute to the inelastic scattering through the asymmetric terms with dips and peaks features ( $\kappa_\lambda \neq 0$ ). This behavior is caused by in-plane vibrational modes of the ZGNR hydrogen atoms in conjunction with local vibrations of relevant groups in the peptide close to the gap. The important findings are mostly related to the existence of two bias ranges for  $V \geq 0.19\text{ meV}$ . In the first one (the ZGNR region) the scattering is induced just by the in-plane ZGNR vibrational modes: here  $C$  and  $H$  atoms vibrations cause a distortion of the pseudo- $\pi$  and pseudo- $\pi^*$  ZGNR orbitals that result in an enhanced transmission through the peptide thus following the conductivity hierarchy already found

in the elastic modeling at low temperature [19]. Then a hypothetical measurement of the differential conductance in this region is expected to follow the same curve as the one found in the elastic regime.

Lastly the "high energy" bias region  $0.35 \text{ V} \leq V \leq 0.45 \text{ V}$  preserves unique informations on the side chain together with additional high energy peaks related to the vibrational modes of the  $NH$  and the  $C_{\alpha}H$  atomic groups and could be employed, in principles, as a source of observables for side chain recognition also with such narrow ZGNR electrodes.

## 5 Acknowledgments

Computing resources and the related technical support used for this work have been provided also by CRESCO/ENEAGRID High Performance Computing infrastructure and its staff [35]. CRESCO/ENEAGRID High Performance Computing infrastructure is funded by ENEA, the Italian National Agency for New Technologies, Energy and Sustainable Economic Development and by Italian and European research programmes, see <http://www.cresco.enea.it/english> for information.

## References

- [1] Miten Jain, Ian T Fiddes, Karen H Miga, Hugh E Olsen, Benedict Paten, and Mark Akeson. Improved data analysis for the minion nanopore sequencer. *Nature methods*, 2015.
- [2] Grégory F Schneider and Cees Dekker. Dna sequencing with nanopores. *Nature biotechnology*, 30(4):326–328, 2012.
- [3] Eamonn Kennedy, Zhuxin Dong, Clare Tennant, and Gregory Timp. Reading the primary structure of a protein with 0.07 nm<sup>3</sup> resolution using a subnanometre-diameter pore. *Nat Nano*, 11(11):968–976, 11 2016.
- [4] James Wilson, Leila Sloman, Zhiren He, and Aleksei Aksimentiev. Graphene nanopores for protein sequencing. *Advanced Functional Materials*, 26(27):4830–4838, 2016.
- [5] Giovanni Di Muccio, Aldo Eugenio Rossini, Daniele Di Marino, Giuseppe Zollo, and Mauro Chinappi. Insights into protein sequencing with an alpha-hemolysin nanopore by atomistic simulations. *Scientific Reports*, 9(1):6440, 2019.
- [6] Abdelghani Oukhaled, Laurent Bacri, Manuela Pastoriza-Gallego, Jean-Michel Betton, and Juan Pelta. Sensing proteins through nanopores: fundamental to applications. *ACS chemical biology*, 7(12):1935–1949, 2012.
- [7] Daniele Di Marino, Emma Letizia Bonome, Anna Tramontano, and Mauro Chinappi. All-atom molecular dynamics simulation of protein translocation through an  $\alpha$ -hemolysin nanopore. *The journal of physical chemistry letters*, 6(15):2963–2968, 2015.
- [8] Loredana Mereuta, Mahua Roy, Alina Asandei, Jong Kook Lee, Yoonkyung Park, Ioan Andricioaei, and Tudor Luchian. Slowing down single-molecule trafficking through a protein nanopore reveals intermediates for peptide translocation. *Scientific reports*, 4:3885, 2014.
- [9] Omid Tavassoly, Sergiy Nokhrin, Oleg Y Dmitriev, and Jeremy S Lee. Cu (ii) and dopamine bind to  $\alpha$ -synuclein and cause large conformational changes. *Febs Journal*, 281(12):2738–2753, 2014.

- [10] David Rodriguez-Larrea and Hagan Bayley. Multistep protein unfolding during nanopore translocation. *Nature nanotechnology*, 8(4):288–295, 2013.
- [11] Takahito Ohshiro, Makusu Tsutsui, Kazumichi Yokota, Masayuki Furuhashi, Masateru Taniguchi, and Tomoji Kawai. Detection of post-translational modifications in single peptides using electron tunnelling currents. *Nature nanotechnology*, 9(10):835–840, 2014.
- [12] Christian B Rosen, David Rodriguez-Larrea, and Hagan Bayley. Single-molecule site-specific detection of protein phosphorylation with a nanopore. *Nature biotechnology*, 2014.
- [13] Stephanie J. Heerema and Cees Dekker. Graphene nanodevices for dna sequencing. *Nature Nanotechnology*, 11:127 EP –, 02 2016.
- [14] Massimiliano Di Ventra and Masateru Taniguchi. Decoding dna, rna and peptides with quantum tunnelling. *Nature nanotechnology*, 11(2):117–126, 2016.
- [15] F Traversi, C Raillon, SM Benameur, K Liu, S Khlybov, M Tosun, D Krasnozhan, A Kis, and A Radenovic. Detecting the translocation of dna through a nanopore using graphene nanoribbons. *Nature nanotechnology*, 8(12):939–945, 2013.
- [16] Eugene Paulechka, Tsjerk A Wassenaar, Kenneth Kroenlein, Andrei Kazakov, and Alex Smolyanitsky. Nucleobase-functionalized graphene nanoribbons for accurate high-speed dna sequencing. *Nanoscale*, 8(4):1861–1867, 2016.
- [17] Rodrigo G. Amorim, Alexandre R. Rocha, and Ralph H. Scheicher. Boosting dna recognition sensitivity of graphene nanogaps through nitrogen edge functionalization. *The Journal-2016 of Physical Chemistry C*, 120(34):19384–19388, 2016.
- [18] Jariyane Prasangkit, Gustavo T. Feliciano, Alexandre R. Rocha, Yuhui He, Tanakorn Osotchan, Rajeev Ahuja, and Ralph H. Scheicher. Theoretical assessment of feasibility to sequence dna through interlayer electronic tunneling transport at aligned nanopores in bilayer graphene. *Scientific Reports*, 5:17560 EP –, 2015.
- [19] Aldo Eugenio Rossini, Fabrizio Gala, Mauro Chinappi, and Giuseppe Zollo. Peptide bond detection via graphene nanogaps: a proof of principle study. *Nanoscale*, 10:5928–5937, 2018.
- [20] K. Stokbro, J. Taylor, M. Brandbyge, J.-L. Mozos, and P. Ordejón. Theoretical study of the non-linear conductance of di-thiol benzene coupled to au(111) surfaces via thiol and thiolate bonds. *Computational Material Science*, 27:151–160, 2002.
- [21] Supriyo Datta. *Electronic Transport in Mesoscopic Systems*. Cambridge Studies in Semiconductor Physics and Microelectronic Engineering. Cambridge University Press, 1995.
- [22] Tchavdar N Todorov. Tight-binding simulation of current-carrying nanostructures. *Journal of Physics: Condensed Matter*, 14(11):3049, 2002.
- [23] K. Stokbro, J. Taylor, M. Brandbyge, and P. Ordejón. Transiesta - a spice for molecular electronics. *Annals of the New York Academy of Sciences*, 1006(1):212–226, 2003.
- [24] Nick Papior, Nicols Lorente, Thomas Frederiksen, Alberto Garca, and Mads Brandbyge. Improvements on non-equilibrium and transport green function techniques: The next-generation transiesta. *Computer Physics Communications*, 212(Supplement C):8 – 24, 2017.
- [25] Young-Woo Son, Marvin L. Cohen, and Steven G. Louie. Energy gaps in graphene nanoribbons. *Phys. Rev. Lett.*, 97:216803, Nov 2006.

- [26] Y. Y. Li, M. X. Chen, M. Weinert, and L. Li. Direct experimental determination of onset of electron–electron interactions in gap opening of zigzag graphene nanoribbons. *Nature Communications*, 5:4311 EP –, 07 2014.
- [27] Daniel Gunlycke, Denis A. Areshkin, Junwen Li, John W. Mintmire, and Carter T. White. Graphene nanostrip digital memory device. *Nano Letters*, 7(12):3608–3611, 12 2007.
- [28] Zuanyi Li, Haiyun Qian, Jian Wu, Bing-Lin Gu, and Wenhui Duan. Role of symmetry in the transport properties of graphene nanoribbons under bias. *Phys. Rev. Lett.*, 100:206802, May 2008.
- [29] Yafei Li, Zhen Zhou, Panwen Shen, and Zhongfang Chen. Spin gapless semiconductor?metal?half-metal properties in nitrogen-doped zigzag graphene nanoribbons. *ACS Nano*, 3(7):1952–1958, 07 2009.
- [30] Young-Woo Son, Marvin L. Cohen, and Steven G. Louie. Half-metallic graphene nanoribbons. *Nature*, 444(7117):347–349, 2006.
- [31] J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kale’, and K. Schulten. Scalable molecular dynamics with namd. *Journal of Computational Chemistry*, 26(16):1781–1802, 2005.
- [32] Thomas Frederiksen, Magnus Paulsson, Mads Brandbyge, and Antti-Pekka Jauho. Inelastic transport theory from first principles: Methodology and application to nanoscale devices. *Phys. Rev. B*, 75:205413, May 2007.
- [33] Jing-Tao Lü, Rasmus B. Christensen, Giuseppe Foti, Thomas Frederiksen, Tue Gunst, and Mads Brandbyge. Efficient calculation of inelastic vibration signals in electron transport: Beyond the wide-band approximation. *Phys. Rev. B*, 89:081405, Feb 2014.
- [34] Giuseppe Zollo and Aldo Eugenio Rossini. Vibration assisted electron tunneling through nanogaps in graphene nano-ribbons for amino-acid and peptide bond recognition. *Nanoscale Adv.*, 1:3547–3554, 2019.
- [35] F. Iannone, F. Ambrosino, G. Bracco, M. De Rosa, A. Funel, G. Guarnieri, S. Migliori, F. Palombi, G. Ponti, G. Santomauro, and P. Procacci. Cresco enea hpc clusters: a working example of a multifabric gpfs spectrum scale layout. In *2019 International Conference on High Performance Computing Simulation (HPCS)*, pages 1051–1052, 2019.

# NEUTRONIC ANALYSES OF PLASMA FACING COMPONENTS OF DEMO DIVERTOR

Simone Noce<sup>\*1</sup>, Davide Flammini<sup>2</sup>, Giovanni Mariano<sup>2,3</sup>, Giuseppe Mazzone<sup>2</sup>, Fabio Moro<sup>2</sup>, Francesco Romanelli<sup>1</sup>, Rosaria Villari<sup>2</sup> and Jeong-Ha You<sup>4</sup>

<sup>1</sup>*University of Rome Tor Vergata, Industrial Engineering Department, I-00133, Rome, Italy<sup>1</sup>*

<sup>2</sup>*ENEA, Department of Fusion and Nuclear Safety Technology, I-00044 Frascati (Rome), Italy*

<sup>3</sup>*University LA Sapienza of Rome, DIAEE department, I-00186 Rome, Italy*

<sup>4</sup>*Max-Planck-Institut für Plasmaphysik, 85748 Garching, Germany*

**ABSTRACT.** Three-dimensional nuclear analyses with MCNP5 Monte Carlo and Fispact II codes have been performed to assess the nuclear loads and main activity quantities on the Plasma Facing Components (PFCs) of European DEMO divertor. The aim of this work is to provide significant outcomes in the PFCs design and concept selection, considering the two principal configurations currently under evaluation: ITER-Like and Chromium low activation concepts.

## 1 Introduction

In the frame of this activity within EUROfusion WPDIV-2 program, three-dimensional nuclear analyses with MCNP5 code [1] and JEFF3.2 nuclear data library [2] have been performed in CRESCO cluster to support design of the PFCs of the DEMO divertor [3]. Two configurations have been analysed: ITER-Like (IL) [4] and a low-activation Chromium-based alternative concept [5], focusing on the evaluation of the nuclear loads (damage, nuclear heating and helium production) on them sub-components. Moreover, activation analysis has been performed with FISPACT II inventory codes [6] to assess their specific activity, decay heat and contact dose at the end of DEMO-1 operations. This work presents detailed neutronics results using heterogeneous description and actual geometry of the two PFC concepts. The computing resources and the related technical support used for this work have been provided by CRESCO/ENEAGRID High Performance Computing infrastructure and its staff [7]. CRESCO/ENEAGRID High Performance Computing infrastructure is funded by ENEA, the Italian National Agency for New Technologies, Energy and Sustainable Economic Development and by Italian and European research programmes, see <http://www.cresco.enea.it/english> for information.

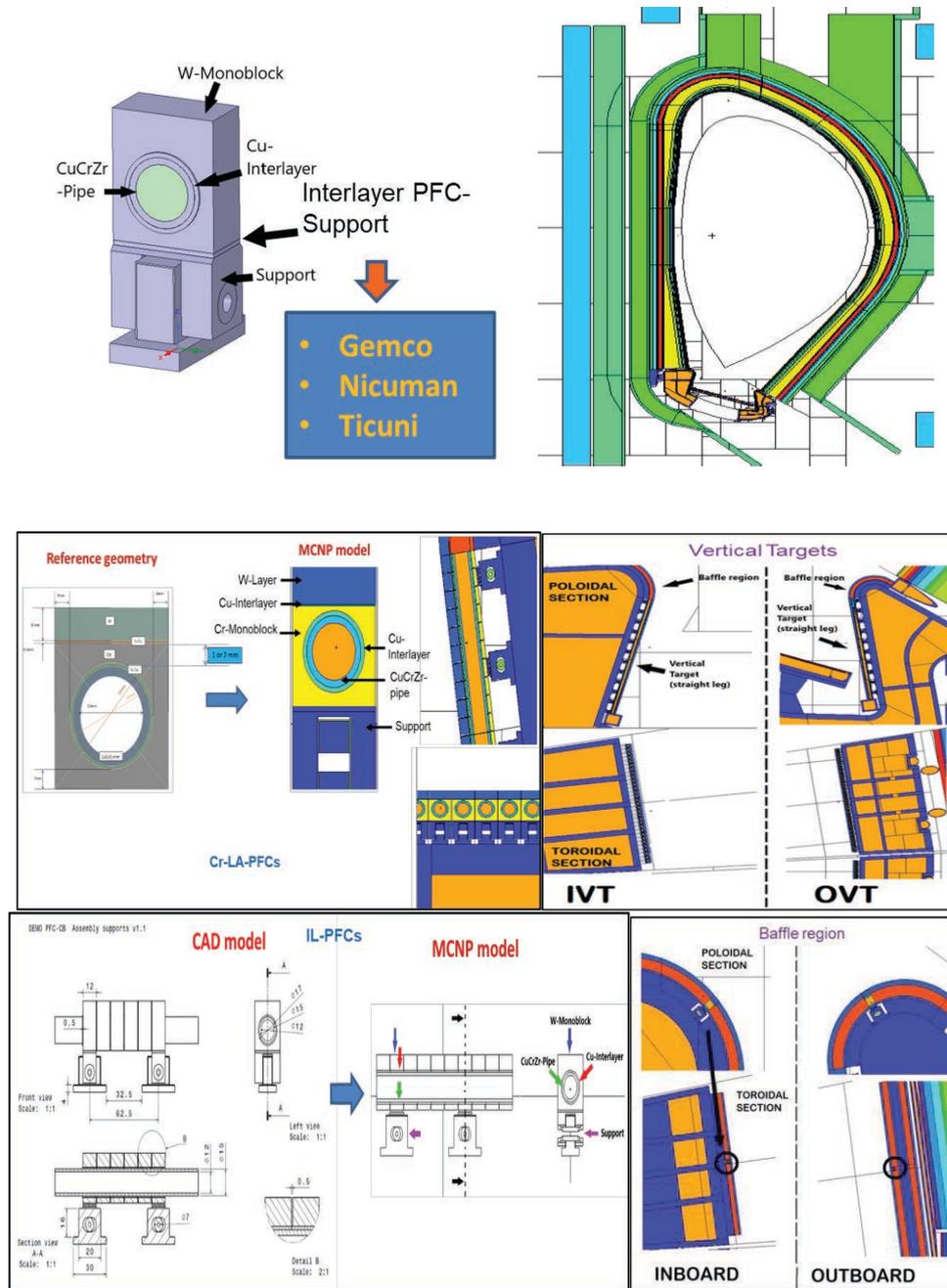
## 2 Divertor PFCs MCNP models description and integration

Starting from the reference CAD models of both the configurations [8,5], detailed MCNP models (including also an accurate PFC-CB supports representation) have been created and integrated in the

---

<sup>1</sup> Corresponding author. E-mail: [Simone.Noce@uniroma2.it](mailto:Simone.Noce@uniroma2.it)

Divertor CB-2018 MCNP model [9], with WCLL “layered” (Heterogeneous) Blanket representation [10]. In Figure 1 is shown a more detailed description of the reference geometries of the PFCs concepts and MCNP Model used for this study.



**Fig. 1:** Left column: representation of a single unit cad of the ITER-Like PFC concept, with a list of the materials candidate for the PFC-support interlayer (top), cad and MCNP models of the Chromium Low Activation (LA) concept (middle) and of the ITER-Like (IL) concept (bottom). Right column: DEMO17 MCNP Model with Single Module WCLL Blanket (top), divertor PFC Inner and Outer Vertical Targets (IVT and OVT) (middle), and Baffle region (bottom) MCNP model description.

### 3 Results

The Inner and Outer Vertical Targets (IVT & OVT) straight leg of the DEMO divertor heterogeneous model have been replaced with the detailed MCNP geometry of the IL and Chromium low activation PFCs and integrated into the WCLL DEMO MCNP model.

For both the concept, the neutronics model of the baffle regions PFCs consists of:

- homogenized layers, representing the actual volume fraction of the constituent materials.
- A single detailed PFC element (including support) integrated in the expected most irradiated region for evaluating the maximum loads (see Figure 1, lower part of the right column).

The most significant nuclear quantities evaluated in this analysis are the following:

- Maximum nuclear heating density, Helium production and damage on the baffle region for ITER-LIKE and Chromium LA concept PFCs (see Tables 1 and 2).
- Activity, Decay Heat, Contact Dose rate on W (IL) [11] and Cr (Chr. LA) [12] with the aim of comparing the activation performance of both materials.
- Activity, Decay Heat, Contact Dose rate on two candidates for the PFC-support interlayer brazing alloys (Nicuman [13] and Gemco [14]) and comparison.

The maximum nuclear loads on the ITER-Like (IL) PFC sub-component placed on the expected most irradiated zone (baffle region) are:

- W-Monoblock: 20.30 W/cm<sup>3</sup> (nuclear heating), 1.80 appm/FPY (He-production), 1.99 dpa/FPY (nuclear damage), where FPY is Full Power Year.
- Cu-Interlayer: 8.01 W/cm<sup>3</sup>, 51.96 appm/FPY, 8.04 dpa/FPY.
- CuCrZr-Pipe: 8.04 W/cm<sup>3</sup>, 52.5 appm/FPY, 8.04 dpa/FPY.

The Maximum nuclear loads on the Chromium Low Activation (LA) PFC sub-component placed on the expected most irradiated zone (baffle region) are:

- W-Layer: 20.80 W/cm<sup>3</sup>, 2.07 appm/FPY, 2.22 dpa/FPY.
- Copper Interlayers: 8.50 W/cm<sup>3</sup>, 60.0 appm/FPY, 8.90 dpa/FPY.
- Cr-Monoblock: 6.69 W/cm<sup>3</sup>, 41.3 appm/FPY, 4.31 dpa/FPY.
- CuCrZr-Pipe: 8.76 W/cm<sup>3</sup>, 52.1 appm/FPY, 7.96 dpa/FPY.

The nuclear heating in Tungsten is three times than in Chromium. Conversely the He-production in Chromium is more than a factor 20 higher than W and the damage is twice and slightly higher values are observed in the nuclear loads of the other sub-components.

Furthermore, specific activity, contact dose rate and decay heat have been evaluated for the comparison between W and Cr in the monoblock and between Gemco, Nicumans as material options for the PFC-support interlayer (where brazing alloys are necessary for the welding). The irradiation history corresponds to 5.2 years operational scenario [15].

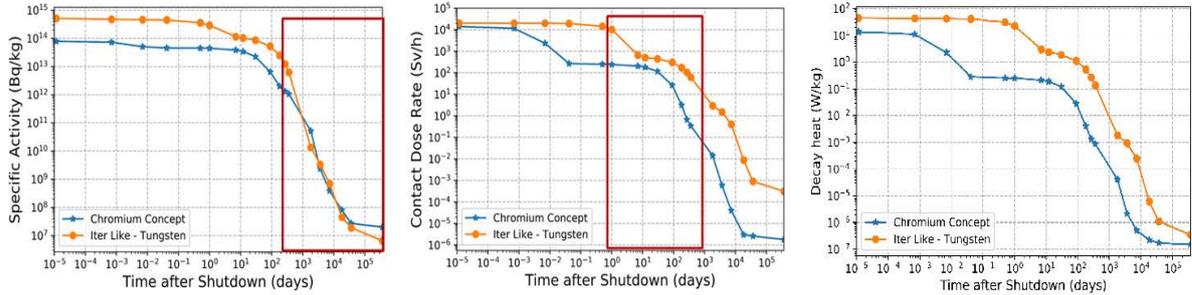
As far as the activation analyses is concerned, the following main points are highlighted:

Cr vs W (see Figure 2):

- Cr shows lower specific activity (Bq/kg) than W within the first year after the shutdown (at the shutdown activity in Cr is  $8 \times 10^{13}$  Bq/kg and in W is  $5 \times 10^{14}$  Bq/kg). After 5 years from the shutdown the induced activity on both the materials decreases, showing an overlapping of the trends and highlighting a similar behaviour in Cr and W for long times.
- Cr contact dose rate (Sv/h) values are similar to the W at the shutdown, with a faster drop after few days. The contact dose rate on Cr falls below 10  $\mu$ Sv/h after more of 20 years.
- At the shutdown the decay heat (W/kg) values are 40 W/kg and 14 W/kg for W and Cr, respectively. The Cr always exhibits lower values than W.

The Chromium confirms to be favourable compared to tungsten in terms of activation for short-medium cooling times (1 year), for longer times the differences with W are less significant due to the impurities. Regarding Gemco and Nicuman Interlayer comparison: the trends are similar for activity, contact dose rate and decay heat: Nicuman shows greater values than Gemco within the first day after

shutdown, for medium cooling times the Gemco shows generally lower values than Nicuman, at long cooling times the Nicuman and Gemco have a similar behaviour. Nicuman is generally not recommended because of higher Nickel content than Gemco.



**Fig.2:** Comparison between W and Cr in terms of specific activity, contact dose rate and decay heat, after plasma shutdown.

## References

- [1] X-5 Monte Carlo Team: MCNP - A General Monte Carlo N-Particle Transport Code, Version 5, Los Alamos National Laboratory, Los Alamos, New Mexico, USA, April 2003.
- [2] JEFF3.2 nuclear data library, [http://www.oecdnea.org/dbforms/data/eva/evatapex/jeff\\_32/](http://www.oecdnea.org/dbforms/data/eva/evatapex/jeff_32/).
- [3] EFDA\_D\_2NQWJV v1.0 - DIV-1-T006-D015-AWP 2019-Neutronic analyses of PFC of DEMO Divertor, 2NQWJV, [https://idm.euro-fusion.org/?uid=2NQWJV&action=get\\_document](https://idm.euro-fusion.org/?uid=2NQWJV&action=get_document)
- [4] [https://idm.euro-fusion.org/?uid=2NVYGP&action=get\\_document](https://idm.euro-fusion.org/?uid=2NVYGP&action=get_document)
- [5] DIV-2.4.5-T001-D001\_Chromium\_concept\_2nd\_2MF5J3\_v1\_3
- [6] The FISPACT-II User Manual UKAEA-R(18)001 February 2018, <https://fispact.ukaea.uk/documentation-2/reports/>
- [7] F. Iannone et al., "CRESCO ENEA HPC clusters: a working example of a multifabric GPFS Spectrum Scale layout," 2019 International Conference on High Performance Computing & Simulation (HPCS), Dublin, Ireland, 2019, pp. 1051-1052, doi: 10.1109/HPCS48598.2019.9188135
- [8] F4E-OMF-567 – Annex B – Technical Specification (F4E\_D\_27MKGH v.1.5)
- [9] P. Frosi, et al., Loads Specification (LS) for Divertor Cassette 2018 (Incl. Neutronics, EM Analysis) Report of DIV-1-T005-D005 WP 2018 (2NLWLE V1.0)
- [10] EFDA\_D\_2NUPDT - INTERNAL DELIVERABLE BB-3.2.1-T005-D001: WCLL DESIGN REPORT 2018, [HTTPS://IDM.EURO-FUSION.ORG/?UID=2NUPDT](https://idm.euro-fusion.org/?uid=2NUPDT)
- [11] Technical Report Design Criteria, Codes and Standards ENS-1.1.5.0-NT-00-R10: Material compositions for PPPT neutronics and activation analyses, EFDA\_D\_2MM3A6 v1.1 (2018)
- [12] <https://www.plansee.com/en/materials/chromium.html>
- [13] <http://www.matweb.com/search/datasheet.aspx?matguid=f57629c960bb4c21816943db19e32173&ckck=1>
- [14] <http://www.morganbrazealloys.com/en-gb/datasheets/material-datasheets/>
- [15] EFDA\_D\_2L8TR9 - PMI-3.3-T014-D001 - Guidelines for neutronic analyses <https://idm.euro-fusion.org/?uid=2L8TR9>

# NEUTRONICS ANALYSES IN SUPPORT OF THE CONCEPTUAL DESIGN OF THE DTT TOKAMAK RADIAL NEUTRON CAMERA

Barbara Caiffi<sup>1\*</sup>, Maurizio Angelone<sup>1</sup>, Andrea Colangeli<sup>1</sup>, Davide Flammini<sup>1</sup>, Nicola Fonnesu<sup>1</sup>, Giovanni Mariano<sup>1</sup>, Fabio Moro<sup>1</sup>, Rosaria Villari<sup>1</sup>

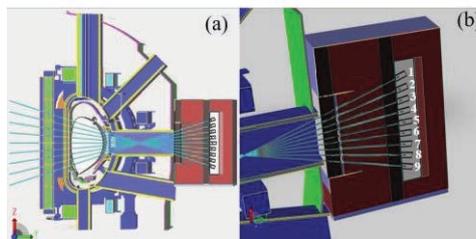
<sup>1</sup>ENEA, Department of Fusion and Nuclear Safety Technology, Frascati, Rome, Italy

**ABSTRACT.** In this report, the neutronics analysis performed in support of the Radial Neutron Camera (RNC) design of the Divertor Tokamak Test (DTT) [1] is presented. The layout of this diagnostic was optimized using a detailed 3D model and simulating the performances by means of the MCNP [3] Monte Carlo code in its MPI parallel version on CRESCO6 infrastructure. The outcomes of this analysis provide the detectors requirements and guidelines for the development of the above-mentioned diagnostics, investigating its feasibility and suitability with the neutron emissivity foreseen for the DTT operational scenarios.

## 1 Introduction

This work is a preliminary neutronics analyses performed to support the Divertor Tokamak Test DTT [1] neutron camera design. A detailed MCNP model of the diagnostic system with detectors assembly has been developed and integrated in the reference neutronics model [2] representing a 20° sector of the DTT machine. Three-dimensional neutron transport simulations have been carried out by means of the MCNP Monte Carlo code [3] coupled with the FENDL 2.1 nuclear data libraries [4]. In particular, the parallel MPI version of the MCNP was used on the CRESCO6 infrastructure (cresco\_62h24 queue with ~600 cores). The activities required several 24h runs (at least 3/4 for each configuration, plus some debugging CPU time), with an efficiency of about 10-50 particle/(core\*s). The computing resources and the related technical support used for this work have been provided by CRESCO/ENEAGRID High Performance Computing infrastructure and its staff [7]. CRESCO/ENEAGRID High Performance Computing infrastructure is funded by ENEA, the Italian National Agency for New Technologies, Energy and Sustainable Economic Development and by Italian and European research programmes, see <http://www.cresco.enea.it/english> for information

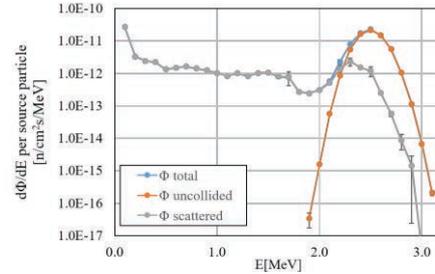
## 2 MCNP model in baseline configuration



**Figure 1:** (a) DTT MCNP model including the Radial Neutron Camera; (b) detail of the RNC

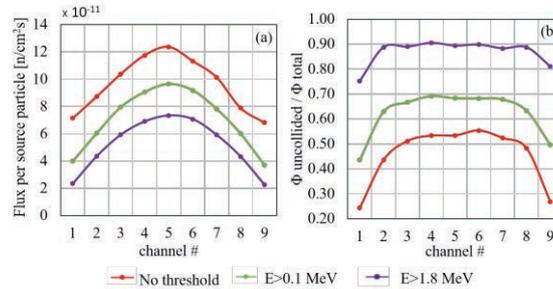
Starting from the reference DTT 20° model [2], a preliminary layout of a radial neutron camera was included (see Fig.1). The main components of this diagnostic are collimators, i.e cylindrical holes passing through the concrete shielding block, and the NE213 detectors, defined as cylinders (1 cm<sup>2</sup> of surface and 1 cm thickness) placed inside cylindrical steel envelopes in the detector box. In the

baseline configuration, the radius of the collimators is 1 cm and their length spans from 140 cm to 147 cm, depending on the inclination. As neutron source, the thermal emissivity profile of a DD plasma at 11 keV of temperature was considered. The energy distribution follows a Gaussian shape, centred at 2.45 MeV and with  $\text{FWHM}=82.6 \sqrt{T_i} \approx 274 \text{ keV}$ , where  $T_i=11 \text{ keV}$  is the ion temperature. Using the above-mentioned model, the neutron flux and energy spectrum were evaluated in the region of interest for each channel, i.e. where the detectors are located. The neutron flux spectrum for the central channel, shown as an example in Figure 2, exhibits a Gaussian-shaped uncollided component, coming directly from the plasma, which represents the signal in RNC measurements, plus a scattered component which extends down to thermal energies and represents the background.

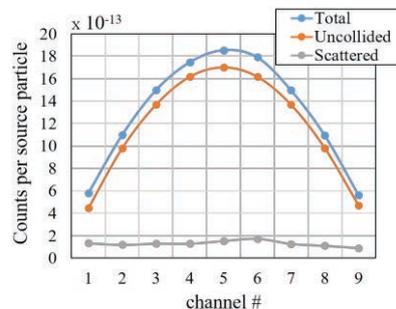


**Figure 2:** Energy spectrum of the neutron flux impinging on the detector of channel 5

In Figure 3 (a), the total fluxes per source particle for all the channels are presented, without and with energy thresholds of 0.1 MeV and 1.8 MeV, again showing the uncollided and collided contributions separately. About one neutron out of  $10^{10}$  emitted by the plasma reaches a RNC detector. This corresponds to an incident flux of  $10^7 \text{ n/cm}^2\text{s}$  assuming the highest neutron yield expected for DTT ( $1.5 \times 10^{17} \text{ n/s}$  [2]). The neutron fluxes and spectra were used to estimate the detectors counting rate and the range of measurable neutron emissivity. In particular, the spectra were convoluted with the NE213 response functions [5] to obtain the Pulse Height Spectrum (PHS). The counts per source particle in each detector were calculated by integrating the PHS, starting from an energy threshold of 1.8 MeV, whose value comes from the optimization of signal to noise ratio.



**Figure 3:** (a) Neutron flux per source particle and (b) uncollided/total flux ratio in each channel.



**Figure 4:** Counts per source particle in each channel.

The expected counting rate per source particle, shown in Figure 4, varies from  $6 \times 10^{-13}$  n/s to  $2 \times 10^{-12}$  n/s, which corresponds to 80 kHz for the most peripheral channels 1 and 9 and 280 kHz for the central one, with a neutron yield of  $1.5 \cdot 10^{17}$  n/s. To have at least 10% accuracy in 25 ms, as required in ITER for DD operations [6], the minimum acceptable neutron yield is  $7 \times 10^{15}$  n/s, while to have the same accuracy in 100 ms, a neutron yield four times lower, i.e.  $2 \times 10^{15}$  n/s, can be accepted. In both cases, this is not suitable for the first year operation scenario [2], in which a maximum neutron yield of about  $5 \times 10^{14}$  n/s is foreseen. To improve the RNC performances, different strategies can be adopted. A lower threshold can be used, but this methodology has the drawback of worsening the signal to noise ratio, already critical for the peripheral channels. Shorter collimators with larger radius could also be a possible solution, which would also allow to reduce the whole dimension of the neutron camera. In order to identify the most promising solutions, a geometrical optimization study was performed, whose results will be presented in the next section.

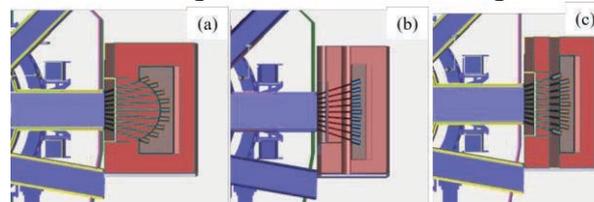
### 3 MCNP model optimization

To enhance the counting rate, in particular in peripheral channels, affecting as little as possible the signal to noise ratio, other configurations were considered and an optimization study was performed. This optimization was focused mainly on the collimators, which are the key element of this diagnostics, and in particular on their size (length and radius).

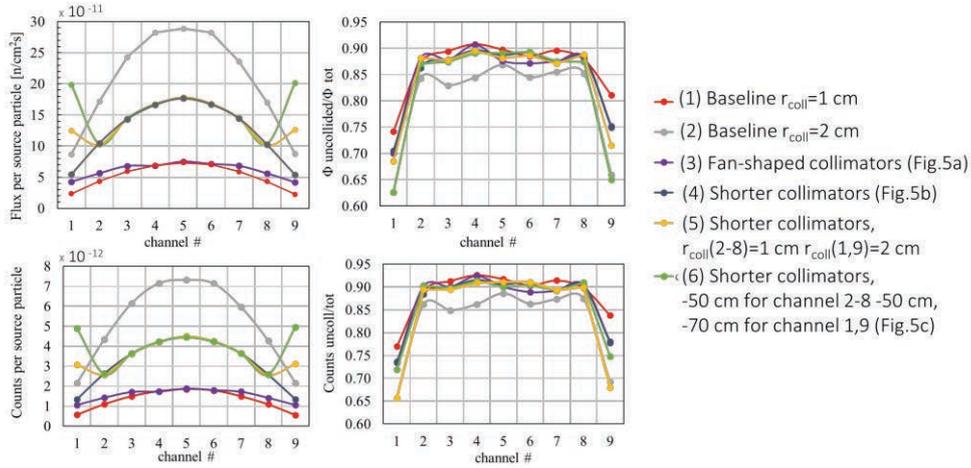
The configuration considered are:

- Baseline, with collimators radius  $r_{\text{coll}} = 1$  cm;
- Baseline with larger collimators,  $r_{\text{coll}} = 2$  cm;
- Fan-shaped collimators, to homogenize their length (Fig.5a);
- Collimators 50 cm shorter (Fig.5b);
- Collimators 50 cm shorter, with larger radius ( $r_{\text{coll}} = 2$  cm) for peripheral channels #1 and #9 and baseline configuration ( $r_{\text{coll}} = 1$  cm) for the others;
- Collimators 70 cm shorter for peripheral channels #1 and #9 and 50 cm shorter for the others (Fig.5c).

The results of the above mentioned configurations are shown in Figure 6 and resumed in Tab. I



**Figure 5:** Variant models studied in the optimization. (a): elliptical fan shaped collimators; (b): collimators 50 cm shorter than baseline; (c): collimators of channels 2-8 50 cm shorter than baseline, collimators of channels 1 and 9 are 70 cm shorter than baseline.



**Figure 6:** Neutron flux (a) and detector counts (c) per source particle with their corresponding direct neutrons fraction (b) (d) as a function of the channel assuming an energy threshold of 1.8 MeV.

**Table 1:** Results for the models considered (count per source particle, max. rate and least detectable neutron yield)

Model	Counts per source particle		Max. Rate [kHz]*		Min. Neutron Yield[n/s]**	
	CH1	CH5	CH1	CH5	CH1	CH5
(1)	$5.8 \cdot 10^{-13}$	$1.9 \cdot 10^{-12}$	87	285	$3.4 \cdot 10^{15}$	$1.1 \cdot 10^{15}$
(2)	$2.1 \cdot 10^{-12}$	$7.3 \cdot 10^{-12}$	315	1100	$9.5 \cdot 10^{14}$	$2.7 \cdot 10^{14}$
(3)	$1.1 \cdot 10^{-12}$	$1.9 \cdot 10^{-12}$	165	285	$1.8 \cdot 10^{15}$	$1.1 \cdot 10^{15}$
(4)	$1.3 \cdot 10^{-12}$	$4.5 \cdot 10^{-12}$	195	675	$1.5 \cdot 10^{15}$	$4.4 \cdot 10^{15}$
(5)	$4.9 \cdot 10^{-12}$	$4.5 \cdot 10^{-12}$	735	675	$4.1 \cdot 10^{15}$	$4.4 \cdot 10^{15}$
(6)	$3.1 \cdot 10^{-12}$	$4.5 \cdot 10^{-12}$	465	675	$6.5 \cdot 10^{15}$	$4.4 \cdot 10^{15}$

\* Max. neutron yield= $1.5 \cdot 10^{17}$  n/s [1].

\*\* assuming 10% accuracy in 50 ms, i.e. assuming as a minimum rate 2 kHz.

With larger collimator radius,  $r_{\text{coll}} = 2$  cm,  $5 \times 10^{14}$  n/s is sufficient as a neutron yield to have 10% accuracy in 100 ms, making this diagnostics suitable also for the first year operation scenario. In this configuration, a counting rate of about 1 MHz is expected for high performance scenario (neutron yield of  $1.5 \times 10^{17}$  n/s). This is a quite high rate though still manageable with the available electronics. The counting rate can be increased also by reducing the collimator length, with the additional benefit of making the diagnostic smaller and lighter. The flux in the channels can be homogenized by using shorter or larger collimators for the most peripheral channels (i.e. configuration 5 and 6). In the last two configurations considered, the minimum counting rate per source particle is  $3 \times 10^{-12}$  count/s, which allows a 10% accuracy in 50 ms.

## 4 Conclusions

In this paper, a preliminary study in support of the design of the radial neutron camera for the DTT machine is presented. The geometry of this diagnostics was included in the 20° model of the tokamak and three-dimensional neutron transport simulations have been carried out by means of the MCNP Monte Carlo code. The neutron fluxes impinging on the detectors placed at the end of each of the nine collimated lines of sight were calculated, as well as the corresponding spectra and the fraction of uncollided neutrons. Furthermore, using the response function of the NE213 scintillator, the counting rate was estimated as a function of the neutron yield. Starting from the baseline model, different geometrical optimizations were performed, aimed at increasing the counting rate, reducing the size of

the diagnostics and improving the signal to noise ratio. As the physics scenarios are still not fully characterized, none of the models studied can be considered as the best so far. Though, this analysis demonstrates the feasibility of the layout proposed and its flexibility to a range of neutron emissivity of almost three orders of magnitude (neutron yield from  $5 \times 10^{14}$  to  $10^{17}$  n/s). Furthermore, this work also proves that effective optimisation solutions can be adopted, according to the requirements that will be defined for this diagnostics.

## References

- [1] DTT Divertor Tokamak Test facility - Interim Design Report. Ed by ENEA. ISBN 978-88-8286-378-4. April 2019 ("Green Book")
- [2] R. Villari et al., Nuclear design of Divertor Tokamak Test (DTT) facility, this conference
- [3] X-5 Monte Carlo Team: MCNP - A General Monte Carlo N-Particle Transport Code, Version 5, LANL Report LACP-03-0245, (2005).
- [4] FENDL-2.1: Fusion Evaluated Nuclear Data Library (Dec. 2004) <https://wwwnds.iaea.org/fendl21/>
- [5] A. Zimbal, et al., Compact NE213 neutron spectrometer with high energy resolution for fusion applications, *Rev. Sci. Instrum.* 75 (2004) 3553. <http://dx.doi.org/10.1063/1.1787935>.
- [6] D. Marocco et al., Neutron Measurements in ITER using the Radial Neutron Camera, *Journal of Instrumentation* 7(03) (2012) DOI: 10.1088/1748-0221/7/03/C03033 6903807, 1030-1033509(A), Ottawa (2013).
- [7] F. Iannone et al., "CRESCO ENEA HPC clusters: a working example of a multifabric GPFS Spectrum Scale layout," 2019 International Conference on High Performance Computing & Simulation (HPCS), Dublin, Ireland, 2019, pp. 1051-1052, doi: 10.1109/HPCS48598.2019.9188135

# THERMAL ENHANCEMENT OF DATA CENTER WITH ARTFUL WORKLOAD ASSIGNMENT

Davide De Chiara<sup>1</sup>, Marta Chinnici<sup>2</sup>, Ah-Lian Kor<sup>3</sup>, Eric Rondeau<sup>4</sup>, Jean-Philippe Georges<sup>4</sup>,  
Anastasiia Grishina<sup>5</sup>, Guido Guarnieri<sup>1</sup>

<sup>1</sup>*ENEA-ICT Division, C.R. Portici Piazzale Enrico Fermi 1, Napoli 80055, Italy  
{davide.dechiara, guido.guarnieri}@enea.it*

<sup>2</sup>*ENEA-ICT Division, C.R. Casaccia Via Anguillarese 301, Roma 00123, Italy  
marta.chinnici@enea.it*

<sup>3</sup>*Leeds Beckett University, Leeds, UK  
a.kor@leedsbeckett.ac.uk*

<sup>4</sup>*CNRS, CRAN, F-54000 Université de Lorraine Vandoeuvre-lès-Nancy, France  
{eric.rondeau, jean-philippe.georges}@univ-lorraine.fr*

<sup>5</sup>*Department of Mathematics and Computer Science, Eindhoven University of Technology, 5612 AZ Eindhoven,  
The Netherlands;  
a.grishina@tue.nl*

**ABSTRACT.** Power consumption of Data Center has negative impacts on environments. Researchers are adopting practical solutions to render Data Center “green” while maintaining a desirable level of quality of service or meeting customers’ Service Level Agreements (SLAs). Energy savings could be made in two noteworthy areas, namely: compute systems and cooling systems. A reliable cooling system is necessary to produce a persistent flow of cold air to cool servers which heat up due to increasing computational loads. Servers’ dissipated waste heat phenomena a critical on the cooling systems. So, it is crucial to identify hotspots that recurrently occur in the server zones. This is facilitated through the application of machine learning techniques to an available big dataset for thermal characteristics for High-Performance Computing ENEA Data Center, namely CRESCO6. We present an algorithm with the DC thermal characteristics as input parameters and aims to snapshot of the DC, try to reduce large thermal-gradient due to uneven distribution of server dissipated waste heat and increasing cooling effectiveness. The work presented comes from the Anastasiia Grishina’s master’ thesis.

**KEYWORDS.** Data Center, HPC, Machine Learning, Big Data, Thermal, Hotspot, Cooling, Thermal management, workload assignment.

## 1 Introduction

A big chunk of worldwide produced electricity is through hydrocarbon combustion. Thus, this causes a rise in carbon emission and other Green House Gasses (GHG) in the environment, contributing to global warming. Data Centers (DC) worldwide were estimated to have spent about 271 billion kWh of electricity in the year 2010 [1] and in 2017, US DC used 90 billion kilowatt-hours of electricity [14]. According to [2], without any provision to reduce energy consumption, carbon emission is estimated to increase from 307 million tons in 2007 to 358 million tons in 2020. Servers in DCs consume energy that is related to processing assigned computing loads, and almost all of the energy

input is being dissipated as waste heat energy. Cooling systems are deployed to maintain the temperature of the computing servers at the vendor specified temperature for consistent and reliable performance. Koomey [1] emphasises that a DC energy input is primarily consumed by cooling and compute systems (including servers in chassis and racks). Thus, these two systems have been critical targets to affect energy savings.

DC cooling involves the installation of cooling systems and useful hot/cold aisle configurations. Poor thermal management in a DC could be the primary contributor to IT infrastructure inefficiency due to thermal degradation. Server CPUs are the primary energy consumers and waste heat dissipators [4]. Generally, existing DC air-cooling systems are not sufficiently efficient to cope with the vast amount of waste heat generated by high performance-oriented microprocessors. Thus, it is necessary to disperse the dissipated waste heat so that there will be an even distribution of waste heat within a premise to avoid overheating. Undeniably, a more effective energy savings strategy is necessary to reduce energy consumed by a cooling system and yet efficient in cooling the servers (in the computing system). One known method is thermal-aware scheduling where a computational workload scheduling is based on expected waste heat dissipation. Thermal-aware schedulers adopt different thermal-aware approaches (e.g. system-level for work placements) [16]; execute ‘hot’ jobs on ‘cold’ compute nodes (Wang, et. al., 2009); predictive model for job schedule selection [17]; ranked node queue based on thermal characteristics of rack layouts (Liang, et. al., 2011) and optimisation (e.g. optimal setpoints for workload distribution and supply temperature of the cooling system (Van Damme, et. al., 2019). Heat modelling provides a model that links server energy consumption and their associated waste heat. Thermal-aware monitoring acts as a thermal-eye for the scheduling process and entails recording and evaluation of heat distribution within DCs. Thermal profiling (Chaudhry, et. al., 2015) is based on useful monitoring information on workload-related heat emission and is helpful to predict the DC heat distribution.

Our analysis investigates the correlation between thermal-aware scheduling and computer workload scheduling for a better distribution of heat within a DC to avoid hotspots and cold spots. It is noteworthy that hotspots are identified through user (and or server) thermal profiling via unsupervised learning technique viewing the variability of thermal data and uncertainties in defining temperature thresholds.

The novel contribution of the research presented in this work is thermal characteristics analysis using real big thermal characteristics monitoring dataset for the ENEA High-Performance Computing (HPC) CRESCO6 compute nodes [22]. Research conducted include the following: hotspots localisation [20, 21]; users categorisations based on submitted jobs to CRESCO6 cluster; compute nodes categorisation based on both internal servers and surrounding air temperatures [20, 21] due to workload related waste heat dissipation.

Thermal characteristics data (i.e. exhaust air temperature, CPUs temperatures), are collected using thermal sensors. They serve as inputs to the clustering algorithm. Next, a series of clustering results are intersected to unravel nodes (identified by IDs) that frequently fall into high-temperature areas. The paper is organised as follows: Introduction in Section 1, Background with related research in section 2; Section 3 discusses an action plan with a description of methodology; in Section 4, we discuss the impact of recommended actions; Section 5 for conclusions and future work.

## **2 Background: Research done**

In the HPC-DC (High-Performance Computing Data Center) field it is crucial to satisfy service level agreements with minimal energy utilisation. DC energy efficiency has been a long-standing challenge due to a combination of multiple factors that determine DC energy efficiency [21]. Existing research strives to explore the trade-offs between performance (i.e. productivity) and sustainability (i.e. energy efficiency). One of the thermal equipment-related challenges is raising the

setpoint of the cooling system or lowering the speed of CRAC (Computer Room Air Conditioning) fans to save energy, and reduce ambient temperatures (note: prolonged high temperatures could decrease IT systems reliability due to thermal degradation). To date, research to investigate on the relationship (if any) between optimal the cooling system energy consumption and long-term IT system reliability is yet to be conducted [8]. Another long-standing challenge is IT resource over-provisioning that causes energy waste due to underutilised or idle servers. Reviewed research explores an optimal allocation of PDUs (Power Distribution Units) for servers, multi-step algorithms for power monitoring, and on-demand provisioning [8].

As far as metrics are concerned, one standard and widely accepted DC energy efficiency metric is Power Usage Effectiveness (PUE) proposed by Green Grid Consortium [2]. It shows the ratio of total DC energy utilisation with respect to the energy consumed solely by IT equipment. In general, DC energy efficiency metrics encompasses the following: thermal characteristics; the ratio of renewable energy use; energy productivity of various IT system components, and etc. The scientific community is yet to rigorously explore a holistic framework that would comprehensively characterise DCs using a fixed set of metrics and reveal potential pitfalls in DC operations. Though some current research work has made such attempts but to date, we are yet to have a standardised framework [9] [13]. To reiterate, the thermal characteristics of the IT system ought to be the primary focus of an energy efficiency framework because it is the main contributing factor to energy consumption within a DC. Several kinds of research have been conducted to address this issue. In [11], the authors propose an ambient temperature-aware capping to maximize power efficiency while minimising overheating. Their research includes an analysis of the composition of energy consumed by a cloud-based DC. Their findings for the design of DC energy consumption are roughly 45% for computing systems; 40% for refrigeration-based air conditioning; remaining 15% for storage and power distribution systems. Additionally, about 50% of DC energy is spent on non-computing devices. In [6], the authors present an analytical model that describes DC resources with heat transfer properties and workloads with thermal features. Thermal modelling and temperature estimation based on thermal sensors collected and analysed data ought to consider the emergence of server hotspots and heat islands within the premise due to the increase in inlet air temperature, inappropriate positioning of a rack, ineffective airflow, or even inadequate room ventilation.

However, there is no reference to heat-modelling or thermal-monitoring cum profiling. Kong et al. [4] emphasise important notions of thermal-aware profiling, thermal-aware monitoring, and thermal-aware scheduling. Thermal-aware techniques are linked to the minimisation of waste heat production, heat convection around server cores, task migrations, and thermal-gradient across the microprocessor chip and microprocessor power consumption. Finally, Dynamic thermal management (DTM) techniques in microprocessors ncompass the following: Dynamic Voltage and Frequency Scaling (DVFS), Clock gating, task migration, and Operating System (OS) based DTM and scheduling. In [5], Parolini et al. propose a heat model; provide a brief overview of power and thermal efficiency from microprocessor micro-level to DC macro-level. In this paper, the authors incorporate thermal-aware scheduling, heat modelling, thermal aware monitoring and thermal profiling. This experiment involves measurement, quantification, and thermal analysis of compute nodes and cooling systems. The analysis aims to uncover underlying causes that cause temperatures to rise that leads to the emergence of thermal hotspots. Overall, effective DC management requires energy use monitoring, significantly, energy input, IT energy consumption, monitoring of supply air temperature and humidity at room level (i.e. granularity level 0 in the context of this research), monitoring of air temperature at a higher granularity level (i.e. Computer Room Air Conditioning/Computer Room Air Handler (CRAC/CRAH) unit level, granularity level 1). DC energy efficiency metrics will not be critically reviewed in this work and the subsequent discussion focuses on thermal guidelines from the American Society of Heating, Refrigerating And AC Engineers (ASHRAE) [7].

### 3 Action Plan

To avoid hotspots and optimise cooling effects, we must detect server nodes that often occur in the hotspot zones through the application of a machine learning clustering algorithm on the workload management platform. The significant thermal characteristics dataset of ENEA Portici CRESCO6 computing cluster is employed for an analysis which includes 24 measured values (or features) for each single calculation node (see Table 1) and comprises measurements for the whole year 2019 of 434 calculation nodes and 20832 cores. The power consumption massive computing workloads amount to a maximum of 190 kW.

**Table 1.** Thermal Characteristics Dataset – description of features.

<b>Feature</b>	<b>Description</b>
<b>Nodename</b>	server ID, integer from 1 to 434
<b>Timestamp</b>	timestamp of a measurement
<b>System, CPU, Memory Power</b>	Node instantaneous power, memory power, CPU power use in three corresponding columns (expressed in Watt)
<b>Fan 1a, Fan1b, ..., Fan 5a, Fan 5b</b>	Speed of a cooling fan installed inside the node (expressed in RPM - revolutions per minute)
<b>System, CPU, Memory, I/O utilisation</b>	Percentage values of CPU, RAM memories and I/O utilization
<b>Inlet, CPU1, CPU2, Exhaust temperature</b>	Temperature at the front, inside (on CPU1 and CPU2) and at the rear of every single node (expressed in Celsius)
<b>SysAirFlow</b>	Speed of air traversing the node expressed in CFM (cubic foot per minute)
<b>DC Energy</b>	Meter of total energy used by the node, updated at corresponding timestamp and expressed in kWh

#### 3.1 Energy Saving Method

Thermal-aware DC scheduling is designed based on real monitoring data of active cluster nodes in an actual physical DC. Data collected are related to inlet air temperature, the internal temperature of each node, the energy consumption of CPU, RAM, memory, etc; environmental parameters (e.g. air temperatures and humidity in both the hot and cold aisles); cooling system related parameters and finally, individual users who submit their jobs to a cluster node. We explore the relationship between dynamic workload assignment and energy consumption of compute as well as cooling systems [22]. The constraint is that each arrived job must be assigned irrevocably to a particular server without any information about impending incoming jobs. Once the job has been assigned, no pre-emption or migration is allowed, which is a typical rule observed for HPC applications due to high data reallocation incurred costs. In particular, we create a plan for advanced physical as well as static job placement in compute nodes based on rack slot availability within the DC. The map forms a matrix comprising computing units with specific characteristics and certain resource availability level at a given time  $t$ . The primary goal is to create a list of candidate nodes to deliver “calculation performance” required by a user’s job. When choosing the candidate nodes, the job-scheduler will evaluate the suitability of the thermally cooler nodes (at the instant  $t$ ) based on their capability to satisfy a user’s SLA. To improve job scheduler-related decision making, it is essential to know in

advance, the type of job a user will submit to a node(s) for computation. Such insight is provided by several years' worth of historical data and advanced data analytics using machine learning algorithms. Through Platform Load Sharing Facility (LSF) accounting data, we code user profiles into four macro-categories:

1. CPU\_intensive
2. MEMORY\_intensive
3. CPU&MEMORY\_intensive
4. CPU&MEMORY\_not intensive

Related terms are as follows: 1) CPU-intensive - applications that are computation intensive; 2) Memory-intensive-a significant portion of these applications require RAM processing and disk operations.

This classification facilitates energy savings and better task allocation to cluster nodes with the primary goal of overall node temperatures reduction. Additionally, when job allocation is evenly distributed, emergence of thermal hotspots and cold spots could be avoided. The temperatures of the calculation nodes could be evened out, thus, resulting in a more even distribution of heat across the cluster.

### **3.2 Profiled log for users and workloads comprehension**

The main reasons for understanding users' behaviour are as follows.

- Identify parameters based on the diversity of submitted jobs for user profiling;
- Analyse the predictability of various resources (e.g. CPU, Memory, I/O); and
- Identify their time-based usage patterns.

Such useful information will feed into building predictive models for estimating future CPU and memory usage based on historical data carried out in the LSF platform. The abstraction of behavioural patterns in the job submission and its associated resource consumption is necessary to predict future resource requirements. The user profile is created based on submitted job-related information. A crosstab of the accounting data provided by the LSF platform and resource consumption data help guide the calculation of relevant thresholds that code jobs into several distinct utilisation categories. The procedure is straightforward and provide rough estimates to draw appropriate conclusion such as: "if the CPU load is high (e.g., over 90%) during almost 60% of the job running time for an application, then the job can be considered as a CPU-intensive one. The objective for the job-scheduler is the optimisation of task scheduling activities when a job with the same AppID (i.e. the same type of job) or the same username is re-submitted to a cluster. In case of a match with the previous AppID or username, related utilisation stats from the profiled log are retrieved. Based on the utilisation patterns, this particular user/application will be placed into one of the four previously discussed categories. Once a job is categorised, a thermally suitable node is selected to satisfy the task calculation requirements. A task with high CPU and memory requirement will not be immediately processed until the node temperature is well under a safe temperature threshold. Node temperature refers to the difference between the node's outlet exhaust air and inlet air temperatures (note: this generally corresponds to the air temperature in the aisles cooled by the air conditioners).

### **3.3 Machine learning for hotspot localisation and thermal follow-up study [20]**

Typical statistical analysis of temperature measurements could not uncover specific nodes or groups of nodes which cause rack hotspots. We employ the methodology in [20, 21] to apply machine learning techniques for node clustering to localise hotspot. Locating hotspots on CRESCO6 group of nodes is accomplished through clustering of sequential sets of nodes into clusters with higher or

lower hot aisle and internal server temperature. Therefore, a Machine Learning clustering technique is selected for deeper analysis of hotspots location and applied to the dataset of nodes temperature measurements. The phases of data analysis involve data cleansing and dataset organization, three data analysis substages which lead to results exploitation in the form of recommendations for the DC. Sequential clustering firstly begins with determining the optimal number of clusters made with the use of two indices, actual clustering of servers into groups with low, medium and high surrounding air temperature ranges and consolidation of results to achieve the most frequently occurred cluster label for each server. [20]

In this paper, K-Means algorithm is selected for clustering the nodes for numerous reasons:

- the number of features used for clustering is small, therefore, formulated clustering problem is simple and does not require more complex algorithms;
- K-Means has linear computational complexity which makes it fast to use for the type of problem in question. While the formulation of the problem is simple, it requires several thousands of repetitions of clustering for each set of N nodes. From this point of view, the speed of the algorithm becomes an influential factor;
- K-Means' weak point, namely random choice of initial centroids, which can lead to different results when different random generators are used, is not assumed as an issue in the current case, since the nodes are clustered several times based on sets of measurements taken at different timestamps and minor differences brought by randomness will be outperformed by repetition of clustering procedure.

The number of clusters K is an unknown parameter which is estimated for each of the three combinations separately using two metrics: 1-average Silhouette Coefficient and 2-Within Cluster Sum of Squares (WCSS) metric [18], [19].

In brief, Silhouette coefficient is computed for each clustered sample and shows how much clusters are isolated from each other or the quality of clustering. The +1 value of Silhouette index for a specific number of clusters K indicates the high density of clusters, -1 shows incorrect clustering and 0 stands for overlapping clusters, therefore, we will focus on local maxima of this coefficient. WCSS is used in the Elbow method of determining the number of clusters and is used here to support the decision obtained from Silhouette coefficient estimation. It measures the compactness of clusters, and the optimal value of K is the one that results in the "turning point" or an "elbow" of the WCSS(K) graph. In other words, increasing the number of clusters after reaching the elbow point will not result in significant improvement of clusters compactness. Although it could be argued that other indices could be additionally used for determining the number of clusters, the combination of the two aforementioned methods has converged on the same values of K, which is assumed to be sufficient for this current research. [20]

Once the optimal number of clusters is obtained, actual clustering is performed for the chosen bases. For every cluster base we further examine how frequently every node is assigned to each cluster and deduce the final cluster label and corresponding sets of nodes. Subsequently, sets of nodes in the hot range for every cluster base are intersected to unravel nodes that are clustered to be in "danger" or hot zone with the highest frequency by three clustering algorithms. Sequential clustering is further performed for each set of N=434 samples based on three combinations of available thermal data: exhaust (base 1), CPU (base 2), exhaust and CPU temperature measurements. The full dataset consists of M=80000 sets of temperature monitoring data where each set consists of 434 node samples with data from sensors installed in different locations. Two metrics are computed for random sets to be clustered.

A number of clusters could be determined using several approaches that are currently widespread among data scientists. However, none of them is considered accurate as they all provide an approximate value. This work utilizes two methods: 1-Within Cluster Sum of Squares (WCSS) or an

elbow method, and 2-average Silhouette Index. The method is based on the idea that increasing the number of clusters after the turning point or an elbow is not meaningful.

In the Average Silhouette method, a Silhouette index is computed for every data point or every member of every cluster and then is averaged over all data points. It estimates consistency of the data within clusters and should be maximized for better separation of the clusters. During sequential clustering each node has been labelled with a certain temperature range cluster. Since clustering is repeated for each set of measurements grouped by time label  $t_i$ , every node is clustered several times and marked with different labels while the algorithm is in progress.

This information implicitly implies the duration of a particular node exists in a certain temperature range. Here, the nodes most frequently occur in the medium temperature range for all cluster bases. However, some nodes remain in the hot range for more than 50% of clustering cases. When nodes persist in the hot range for an extended period or often fall in this range, it implies that they are subject to reduced reliability and accessibility as they are tuned to switch to lower power mode when overheated. Therefore, we continue with the analysis to identify the actual node IDs that have most frequently been clustered within the hot ranges. The main result of this analysis is detection of the hot range node. Such models provide an overview of the whole IT room air temperature distribution, whereas the area of interest is limited by racks and their immediate proximity. Therefore, with less computational power (and thus energy consumption) analysis techniques of this phase have brought about sufficient information to incentivize improvement of thermal conditions.

### **3.4 Real-time workload management based on Thermal awareness: Cluster evaluation**

It is beneficial to have an overview of relevant thermal parameters (e.g. temperatures of each component in the calculation nodes) for each cluster to facilitate efficient job allocation by the job-scheduler. Generally, a snapshot is obtained through direct interrogation of the nodes and installed sensors in their vicinity, or inside the calculation nodes. For each individual node, all available parameters are evaluated, though the highly prioritised parameter is the difference between the node's inlet and exhaust air temperatures. If there is an evident difference, then we deduce that the node is very busy and the same for each calculation node until we have virtually stored relevant information in a matrix that represents the state of the entire cluster. For new job allocation, the scheduling algorithm will choose a node based on its states depicted in the matrix (e.g. recency or Euclidean distance, its internal temperature, etc...). Through this, generated waste heat is evenly distributed over the entire "matrix" of calculation nodes so that hotspots could be significantly reduced. Moreover, a user profile is an equally important criterion for resource allocation. This is due to the fact that user profiles provide insights into user consumption patterns and the type of submitted jobs and their associated parameters. For example, if we know that a user will perform CPU-intensive jobs for 24 hours, we will allocate the job in a "cell" (calculation node) or a group of cells (when the number of resources requires many calculation nodes) that are physically well distributed or with antipodal locations. This selection strategy aims to evenly spread out the high-density nodes followed by the necessary cooling needs. This will help minimise DC hotspots and ascertain efficient cooling with reduction in cooling-related energy consumption.

## **4 Impact of action and discussion**

As previously discussed, we have created user profiles based on submitted job-related information. Undeniably, these profiles are dynamic because they are constantly revised based on user resource consumption behaviour.

The following design principles guide the design and implementation of the job scheduler:

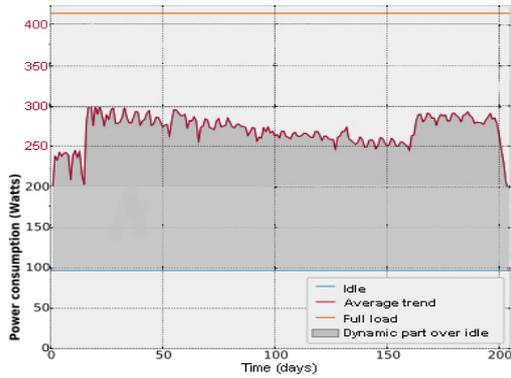
- 1) Job categories - assign an incoming job to one of these 4 categories: CPU-intensive, memory-intensive, neither CPU nor memory-intensive, and both CPU and memory-intensive tasks;
- 2) Utilization monitoring - monitoring CPU and memory utilization while making scheduling decisions;
- 3) Redline temperature control - ensure operating CPUs and memory under threshold temperatures;
- 4) Average temperatures maintenance - monitor average CPU and memory temperatures in a node and manage an average exhaust air temperature across a cluster.

To reiterate, user profile categorisation is facilitated by maintaining a log profile of both CPU and memory utilization for every job (with an associated user) processed in the cluster. A log file contains the following user-related information: (1) user ID; (2) Application identification; (3) the number of submitted jobs; (4) CPU utilization; (5) memory utilization.

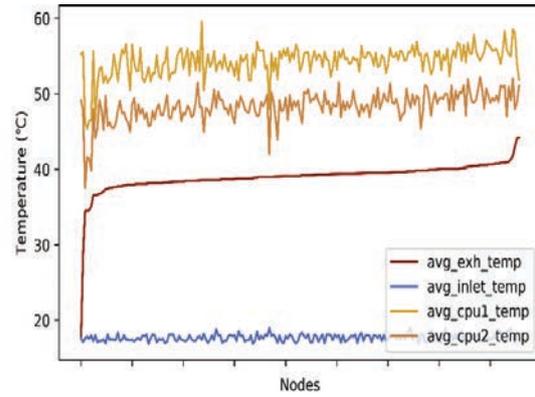
#### **4.1 Baseline of analysis**

By applying the discussed evaluation criteria, we obtain four associated user categories. The algorithm always makes a comparison between a job just submitted by a user and the time series (if any) of the same user. If the application launched or the type of submitted job remains the same, then the user will be grouped into one of the 4 categories. During each job execution, the temperature variations of the CPUs and memories are recorded at pre-established time intervals. Finally, it continuously refines the user behaviour based on the average length of time the user uses for the job. This will provide a more accurate user (and job) profile because it allows for reliable information on the type of job processed in a calculation node and its total processing time. The job scheduler will exploit such information for better job placement within an ideal array of calculation nodes in the cluster. A preliminary study is conducted to provide insight into the functioning of the clusters. For over 1 year, we have observed the power consumption (Fig. 1) and temperature (Fig. 2) profiles of the nodes with running workloads. We have depicted energy consumed by the various server components (CPU, memory, other) in Figure 3 and presented a graph that highlights the difference in energy consumption between idle and active nodes (Fig. 4).

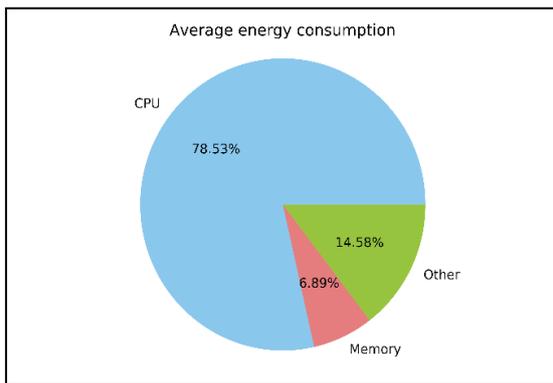
It is observed that for each node, an increase in load effects an increase in temperature difference between inlet and exhaust air for that particular node. Figure 5 [20] depicts the average observed inlet air temperature (blue segment, and in the cold aisle), and exhaust air temperature at their rear side) amaranth segment, in the hot aisle). Note that the temperature measurements are also taken two CPUs adjacent to every node. The setpoints of the cooling system are approximately 18°C at the output and 24°C at the input of the cooling system – as respectively shown in Figure 5 as blue and red vertical lines. However, it appears that the lower setpoint is variable (supply air at 15-18°C) while the higher setpoint varies from 24-26°C. As observed from the graph, the cold aisle maintains the setpoint temperature at the inlet of the node, which affirms the efficient design of the cold aisle (i.e. due to the use of plastic panels to isolating the cold aisle from other spaces in the IT room). However, the exhaust air temperature has registered on average, 10°C higher level than the hot aisle setpoint. Notably, exhaust temperature sensors are directly located at the rear of the node (i.e. in the hottest parts of the hot aisle).



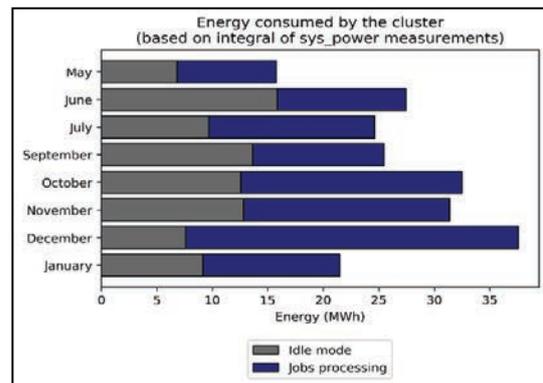
**Fig 1.** The representative shape of Power profile portion on average for all available nodes. Power consumption dataset for a subset of 200 days [20].



**Fig 2.** Temperature profiles (subset of 1 month) on average for all available nodes. Nodes are sorted in the order of exhaust air temperature increase [20].



**Fig 3.** Average energy partitioning for all nodes of cluster CRESCO6 [20]

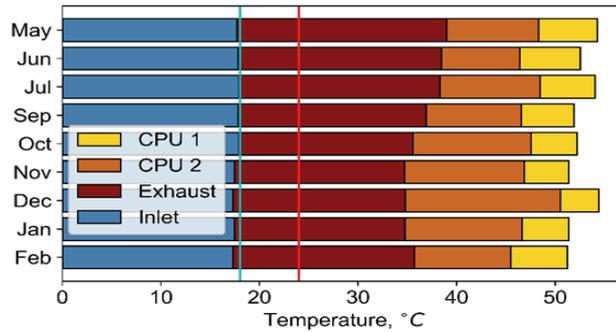


**Fig 4.** Energy consumption in idle and active nodes (subset of 8 months) [20]

Therefore, it is observed that hotspots are detected at the back of server racks, while the hot aisle air is chilled down to the 24-26°C. This is due the ineffective cooling system (i.e. CRAC system) which results in hot air intake, air circulation and cold-hot air mix in the hot aisle.

Meanwhile, the previously mentioned temperature difference of 10°C between the hotspots and the ambient temperature unravels the cooling system weak points because it could not directly cool the hotspots. Cold air flows through the node and is measured at the inlet, then at CPU 2 and CPU 1 locations (directly on the CPUs) and finally, at the exhaust point of the server. The differences between observed temperature ranges in these locations are averaged for all the nodes. An investigation on the observed temperature distribution contributes to the overall understanding of the thermal characteristics, as it provides an overview of the prevailing temperatures shown in Figure 5 and Figure 6. For every type of thermal sensors, the temperature values are recorded as an integer number, so the percentage of occurrences of each value is calculated. The inlet air temperature is registered around 18°C in the majority of cases and has risen up to 28°C in around 0.0001% of cases. It could be concluded that the cold aisle temperature remains around the 15-18°C setpoint for most of

the monitored period. Ranges of the exhaust temperature and those of CPUs 1 and 2 are in the range 20-60°C with most frequently monitored values in the intervals of 18-50°C. Although these observations might incur measurement errors, they reveal servers that are at risks of frequent overheating when benchmarked with manufacturer’s recommendation data sheets. [20]



**Fig 5.** Temperature observed on average in all nodes during 9 months with vertical lines corresponding to cold and hot aisle setpoints. [20]

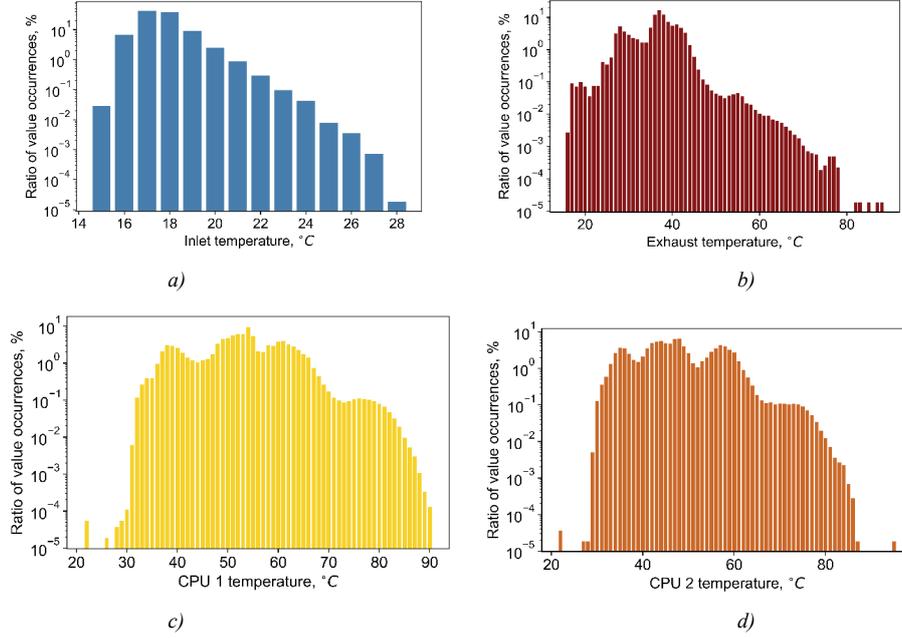
Additionally, this study focuses on variation between subsequent thermal measurements with the aim of exploring temperature stability around the nodes. All temperature types have distinct peaks of zero variation which decreases symmetrically and assumes a Gaussian distribution. It could be concluded that temperature tends to be stable in the majority of monitored cases. However, the graphs for exhaust and CPUs 1 and 2 temperature variation (Figure 6 reveal that less than 0.001% of the recorded measurements show an amplitude of air temperature changes of 20°C or more occurring at corresponding locations).

Further investigation is needed to uncover causes of abrupt temperature changes so that appropriate measures could be undertaken by DC operators to maintain prolonged periods of constantly favourable conditions. We propose a scheduler upgrade which aims to optimise CPU and memories-related resource allocation, as well as exhaust air temperatures without relying on profile information. Prescribed targets for the proposed job scheduler are shown in Table 2.

**Table 2.** Schema with prefixed target for improved job scheduler.

	<b>Proposed job scheduler</b>
<b>Strategy</b>	<i>Schedules job based on utilization and temperature information gathered at run-time</i>
<b>Job Assignment</b>	Assigns a job to the coolest node in a cluster at any point in time
<b>Job Scheduling</b>	Schedules a job on the coolest node in a cluster
<b>Temperature control</b>	Maintains uniform temperate across a cluster
<b>Node Activity</b>	At least 50% are active nodes at any given time in a cluster
<b>Pros</b>	Works better with a large cluster
<b>Cons</b>	Overhead of communication of temperature and utilization information

The proposed job schedule attempts to differentiate the following: between CPU-intensive tasks and memory-intensive tasks; consider CPU and memory utilization during the scheduling process; maintain CPU and memory temperatures under the threshold redline temperatures; minimise the



**Fig 6.** Distribution of monitored temperature values taken for all nodes and months [20].

average exhaust air temperature of nodes to reduce cooling cost. It takes into account information on hotspots derived from machine learning procedures and feedback of node status through queried Confluent platform [15] (monitoring software installed on each node). When all the nodes are busy, the job scheduler will manage the temperatures, embarks on a load balancing procedure by keeping track of the coolest nodes in the cluster. In doing so, the scheduler continues job executions even in hot yet undamaging conditions. The job scheduler maintains the average cluster CPU and memory utilization represented by  $U_{\text{CPUavg}}$  and  $U_{\text{MEMavg}}$ , CPU and memory temperatures represented by  $T_{\text{CPUavg}}$ ,  $T_{\text{MEMavg}}$ , respectively. The goal of our enhanced job scheduler is to maximize the COP (coefficient of performance). Below are the 7 constraints (at nodes level) for our enhanced scheduler:

1. *check constraint  $T_{\text{CPU}}^i (\text{instant } i) < T_{\text{CPUAvg}}$*
2. *otherwise, check constraint  $T_{\text{Mem}}^i < T_{\text{Memavg}}$*
3.  *$T_{\text{Memavg}} < T_{\text{MemMax}}$  and  $T_{\text{CPUavg}} < T_{\text{CPUMax}}$*
4.  *$T_{\text{out}}^i \leq (\sum_{i=1}^N T_{\text{out}}^i) / N$*
5. *Each job is assigned to utmost one node*
6. *Minimise response time of job*

With the first and second constraints being satisfied, ensure that the memory and CPU temperatures remain below the threshold temperatures. If a cluster's nodes exceed the redline threshold, then optimize the temperature by assigning jobs to the coolest node in the cluster, as far away as possible from a hotspot. The third constraint specifies that if the average temperature of memory or CPU rises above the maximum temperature, then the scheduler should stop scheduling tasks as it might encounter hardware failures. The fourth constraint states that the exhaust air temperature of a node should be the same or less than the average exhaust air temperature of the cluster (taking into consideration N number of nodes). The fifth constraint ensures that a node gets utmost one job at a single point in time.

The following is the description of our algorithm:

```

****matrix of node with position r-ow and c-olumn****
Cluster= matrix[r,c]
Hotspot_list= hotspot_list(Cluster) //from machine learning procedure
user=getUSERfromSubmittedJob_in_LSF
Jobtype= getJobProfile(user)

****push the values of utilization and temperature for cpu and memory into matrix****
for (i=0; i=number_of_node;i++) do
    nodename = getnodeName(i)
    Ucpu = getCPU_Utilization(nodename)
    Umemory = getMEMORY_Utilization(nodename)
    Tcpu = getCPU_Temperature(nodename)
    Tmemory = getMEMORY_Temperature(nodename)
End for

*****if a user is not profiled *****
if Jobtype= null then
    *****try to understand job type at run time*****
    if (Ucpu <= U_threshold_cpu) && (Umemory <= U_threshold_memory) then
        Jobtype=easyJob
    else if (Ucpu>U_threshold_cpu) && (Umemory < U_threshold_memory) then
        Jobtype=CPUintensiveJob
    else if (Ucpu<U_threshold_cpu) && (Umemory > U_threshold_memory) then
        Jobtype=MEMORYintensiveJob
    else
        Jobtype=CPU&MEMORYintensiveJob
    end if
end if

***** I try to find the candidate nodes for each type of job*****
avgTempCluster= avgTemp(Cluster)
minT_nodename= getTempNodename(minTemp(Cluster))
maxT_nodename=getTempNodename(maxTemp(Cluster))

*****intervals of temperatures for candidate nodes*****
bestCPUIntensiveNode=getNode(minT_nodename, minT_nodename+25%, Hotspot_list)
bestMEMORYIntensiveNode= getNode(minT_nodename+50%, minT_nodename+75%, Hotspot_list)
bestCPU&MEMORYIntensiveNode= getNode(minT_nodename+25%, minT_nodename+50%, Hotspot_list)
bestEasyJob= getNode(maxT_nodename, maxT_nodename-25%, hotspot_list )

*****job assignments*****
if Jobtype= CPUintensiveJob then
    assignJob (bestCPUIntensiveNode)
else if Jobtype= MEMORYintensiveJob then
    assignJob (bestMemoryIntensiveNode)
else if Jobtype= CPU&MEMORYintensiveJob then
    assignJob(bestCPU&MEMORYIntensiveNode)
else
    assignJob(bestEasyJob)
end if

```

The algorithm takes the node matrix and lists of hotspots by considering the physical arrangement of every single node inside the racks. Firstly, obtain the profile of the user who submits a resources request. This is done by retrieving the user's profile from a list of stored profiles. If the user profile does not exist, then when a user executes a job for the first time, the algorithm calculates a profile instantaneously. All the indicated threshold values are operating values calculated for each cluster configuration and are periodically recalculated and revised according to the use of the cluster nodes. Subsequently, some temperature calculations are made from the current state of the cluster (through a snapshot of the thermal profile). Finally, the last step is to assign the job to the node based on the expected type of job. Through this, the algorithm helps avert the emergence of hotspots and cold spots by uniformly distributing the jobs in the cluster.

## 5 Conclusion and Future Work

Energy efficiency ought to be the ultimate goal for a DC with a sizeable high-performance computing facility to support the development of sustainability targets. This work primarily focuses on two significant aspects: IT equipment energy productivity and thermal characteristics of an IT room and its infrastructure. The findings of this research are based on the analysis of available monitored thermal characteristics-related data for CRESCO6. These findings feed into recommendations for enhanced thermal design and better load management. In this research, clustering performed on big datasets for CRESCO6 IT room temperature measurements has grouped nodes into clusters based on their thermal ranges followed by uncovering the clusters they frequently subsume during the observation period. Additionally, a machine-learning algorithm has been employed to detect better the hotspots and about 8% of the nodes have been frequently placed in the hot range category (thus labelled as hotspots). Some measures to mitigate risks associated with the issue of hotspots have been recommended: more efficient directional cooling, load management, and continuous monitoring of the IT room thermal conditions. This research brings about two positive effects in terms of DC energy efficiency. Firstly, being a thermal design pitfall, hotspots pose as a risk of local overheating and servers thermal degradation due to prolonged exposure to high temperatures. Undeniably, information of hotspots localisation could facilitate better thermal management of the IT room where waste heat is evenly distributed. Thus, it ought to be the focus of enhanced thermal management in the future. Secondly, we discussed ways to avert hotspots through thermal-aware resource allocation (i.e. select the coolest node for a new incoming job), and selection of nodes (for a particular job) that are physically distributed throughout the IT room. This work targets on a scheduling algorithm aiming at allocating workload to physical hosts of Data Centers in such a way that the target host will not be overloaded or over-heated. This means that we schedule calculation nodes with respect to the temperature and CPU utilization of processors.

Under normal operational conditions of a Data Center, therefore, with a non-HPC platform, our scheduler is easily deployable which can bring about estimated savings of approximately 7% of ENEA Data Center electricity consumption. It is obvious that for each Data Center, the estimate varies due to other factors. Generally, with new generation equipment (calculation nodes+air conditioners) the savings value could be easily determined. However, older equipment would have to consider the impact of equipment obsolescence of equipment on energy efficiency. Our scheduler is beneficial if a Data Center focuses on the trade-offs between sustainability and performance because the scheduler primarily reviews the nature of the workload reallocation process in terms thermal characteristics rather than speed and execution time. In our research, we have not explored the impact of virtualisation on energy efficiency because that was beyond the research scope. However, we have validated the novel scheduler on a simulation environment and compared our results with several other scheduling schemes. The experimental results clearly evidence our scheduler's sustainability-related benefit.

For our future research work, we will address on the optimal trade-off between performance and sustainability for our proposed scheduler. We would like to conduct machine learning for data that comes from LSF (load sharing platform) to better understand Data Center user behaviour. Non-optimized algorithms will provide invaluable insights into correct execution of tasks and resources waste. Our future work will also aim to improve the models for calculating processor temperature in order to support the study on multi-core servers.

**Acknowledgments:** This work was supported by EMJMD PERCCOM Project [23]. During the work on the project, the author A. Grishina<sup>4</sup> was affiliated with ENEA-R.C. Casaccia, Rome, Italy. The computing resources and the related technical support used for this work have been provided by CRESCO/ENEAGRID High Performance Computing infrastructure and its staff [24].

CRESCO/ENEAGRID High Performance Computing infrastructure is funded by ENEA, the Italian National Agency for New Technologies, Energy and Sustainable Economic Development and by Italian and European research programmes, see <http://www.cresco.enea.it/english> for information. Moreover, the authors extend their gratitude to the research HPC group at the ENEA-R.C. Portici for the discussion of ENEA-Data Center control and modelling.

## References

- [1]. J. Koomey: Growth in Data Center Electricity use 2005 to 2010. Analytics Press., 1–24. <https://doi.org/10.1088/1748-9326/3/3/034008>, 2011.
- [2]. Greenpeace: How Dirty Is Your Data? A Look at the Energy Choices That Power Cloud Computing, 2011.
- [3]. V. D. Reddy, et al: Metrics for Sustainable Data Centers, *IEEE Trans. Sustain. Comput.*, vol. 2, no. 3, pp. 290–303, Jul. 2017.
- [4]. J Kong, S. W. Chung, k. Skadron: Recent thermal management techniques for microprocessors. *ACM Computing Surveys*, vol. 44, no 3. <https://doi.org/10.1145/2187671.2187675>, 2012.
- [5]. L. Parolini, B. Sinopoli, B. H. Krogh, Z. Wang: A Cyber-Physical Systems Approach to Data Center Modeling and Control for Energy Efficiency. *Proceedings of the IEEE*. 100. 254-268. [10.1109/JPROC.2011.2161244](https://doi.org/10.1109/JPROC.2011.2161244), 2012.
- [6]. L. Wang, et al: Thermal aware workload placement with task-temperature profiles in a datacenter. *J. Supercomput.* 2012, 61, 780–803, doi:10.1007/s11227-011-0635-z, 2012.
- [7]. ASHRAE Technical Committee 9.9, “Thermal Guidelines for Data Processing Environments – Expanded Data Center Classes and Usage Guidance,” 2011.
- [8]. X. Jin, F. Zhang, A. V. Vasilakos, and Z. Liu: Green Data Centers: A Survey, Perspectives, and Future Directions,” *arXiv*, vol. 1608, no. 00687, 2016.
- [9]. M. Chinnici, A. Capozzoli, and G. Serale: Measuring energy efficiency in data centers. In *Pervasive Computing Next Generation Platforms for Intelligent Data Collection*, Chapter 10, pp. 299-351, 2016.
- [10]. M. Chinnici, et al: Data Center, a Cyber-Physical System: Improving Energy Efficiency Through the Power Management,” in *2017 IEEE 15th Intl Conf on Dependable, Autonomic and Secure Computing, 15th Intl Conf (DASC/PiCom/DataCom/CyberSciTech)*, 2017, pp. 269–272.
- [11]. S. Yeo, et al.: ATAC: Ambient Temperature-Aware Capping for Power Efficient Datacenters, in *Proceedings of the 5th ACM Symposium on Cloud Computing*, Seattle, WA, USA, doi:10.1145/2670979.2670966, 2014.
- [12]. A. Capozzoli, et al: Thermal Metrics for Data Centers: A Critical Review,” *Energy Procedia*, vol. 62, pp. 391–400, Jan. 2014.
- [13]. A. Capozzoli, et al: Review on performance metrics for energy efficiency in data center: The role of thermal management,” *LNC*, vol. 8945, pp. 135–151, 2015.
- [14]. Vxchnge site, <https://www.vxchnge.com/>, last accessed 2020/3/27.
- [15]. Confluent site, <https://docs.confluent.io/platform.html>, last accessed 2020/3/27.
- [16]. T. Van Damme, C. De Persis and P. Tesi: Optimized Thermal-Aware Job Scheduling and Control of Data Centers, in *IEEE Transactions on Control Systems Technology*, vol. 27, no. 2, pp. 760-771, March 2019, doi: 10.1109/TCST.2017.2783366
- [17]. G. Varsamopoulos, et. al.: Energy Efficiency of Thermal-Aware Job Scheduling Algorithms under Various Cooling Models, In S. Ranka et al. (Eds.): *IC3 2009, Communications in Computer and Information Science*, vol 40. Springer, Berlin, Heidelberg, pp. 568–580, 2009, doi: [https://doi.org/10.1007/978-3-642-03547-0\\_54](https://doi.org/10.1007/978-3-642-03547-0_54)
- [18]. L. Kaufman and P. J. Rousseeuw, *Finding Groups in Data: An Introduction to Cluster Analysis*. Hoboken, NJ, USA: John Wiley & Sons, Inc., 1990.

- [19]. A. Kassambara, "Determining The Optimal Number Of Clusters: 3 Must Know Methods." [Online]. Available: <https://www.datanovia.com/en/lessons/determining-the-optimal-number-of-clusters-3-must-know-methods/>. [Accessed: 06-May-2019].
- [20]. A. Grishina, "Data Center Energy Efficiency Assessment Based on Real Data Analysis," Master Thesis, School of Engineering Science, PERCCOM, Lappeenranta University of Technology, pp. 1-105, 2019, url: <http://urn.fi/URN:NBN:fi-fe2019090326485> [accessed 15-Oct-2020]
- [21]. A. Grishina, M. Chinnici, A.-L. Kor, E. Rondeau, and J.-P. Georges, "A Machine Learning Solution for Data Center Thermal Characteristics Analysis," *Energies*, vol. 13, no. 17, p. 4378, Aug. 2020.
- [22]. D. De Chiara, M. Chinnici, and A.-L. Kor, "Data Mining for Big Dataset-Related Thermal Analysis of High Performance Computing (HPC) Data Center," in *ICCS 2020: Computational Science – ICCS 2020*, 2020, pp. 367–381.
- [23] Klimova, A.; Rondeau, E.; Andersson, K.; Porras, J.; Rybin, A.; Zaslavsky, A. An international Master's program in green ICT as a contribution to sustainable development. *J. Clean. Prod.* 2016, 135, 223–239.
- [24] F. Iannone et al., "CRESCO ENEA HPC clusters: a working example of a multifabric GPFS Spectrum Scale layout," 2019 International Conference on High Performance Computing & Simulation (HPCS), Dublin, Ireland, 2019, pp. 1051-1052, doi: 10.1109/HPCS48598.2019.9188135

# BENCHMARK OF THE LAMMPS CODE ON CRESCO AND SUMMIT HPC SYSTEMS

M. Lupo Pasini<sup>1\*</sup>, S. Giusepponi<sup>2</sup>, F. Buonocore<sup>2</sup>, M. Celino<sup>2</sup>, A. Frattolillo<sup>3</sup> and S. Migliori<sup>2</sup>

<sup>1</sup>ORNL-CSE Division Oak Ridge, TN, 37831, USA

<sup>2</sup>ENEA-TERIN-ICT Division, C.R. Casaccia, via Anguillarese 301, 00123 Rome, Italy

<sup>3</sup>ENEA-FSN-FUSTEC Division, C.R. Frascati, via E. Fermi 45, 00044 Frascati (RM), Italy

**ABSTRACT.** We present the benchmark results of a numerical model where the solid Deuterium grows under gas pressure. The two-phase molecular Deuterium systems (gas + solid) is modeled using classical molecular dynamics approach and the Large-scale Atomic/Molecular Massively Parallel Simulator (LAMMPS) code. The performance between different modern supercomputers are analyzed and compared. In addition, the numerical simulations reach high accuracy to produce results that match with the data collected through laboratory experiments.

*This manuscript has been authored in part by UT-Battelle, LLC, under contract DE-AC05-00OR22725 with the US Department of Energy (DOE). The US government retains and the publisher, by accepting the article for publication, acknowledges that the US government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this manuscript, or allow others to do so, for US government purposes. DOE will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan (<http://energy.gov/downloads/doe-public-access-plan>). This work used resources of the Oak Ridge Leadership Computing Facility (OLCF), which is a DOE Office of Science User Facility supported under Contract DE-AC05-00OR22725.*

## 1 Introduction

A numerical model to describe the growth of the solid Deuterium under gas pressure was set up in our previous work [1]. In this model a two-phase molecular Deuterium ( $D_2$ ) system (gas + solid) was analyzed and characterized at very extreme conditions of temperature and pressure below the triple point:  $T = 18.71$  K;  $P = 171.3$  mbar. This study was necessary in order to address technological issues related to the production of high-quality pellets for high-speed pipe-gun injectors [2, 3, 4, 5, 6, 7, 8, 9]. In fact, the numerical simulations reach high accuracy to produce results that match with the data collected through laboratory experiments, as de-sublimation temperature and the quality of the cryogenic pellets at different thermodynamics conditions. The system consists in an initial configuration of 128000 Deuterium  $D_2$  particles; 64000 particles in solid phase + 64000 particles in gas phase. It is used to evaluate and compare the HPC performance of the supercomputer CRESCO6 at the Agenzia nazionale per le nuove tecnologie, l'energia e lo sviluppo economico sostenibile (ENEA) and SUMMIT of the Oak Ridge Leading Computing Facility (OLCF) at Oak Ridge National Laboratory (ORNL).

---

\*Corresponding author. E-mail: lupopasini@ornl.gov.

## 2 Results and discussion

### 2.1 Description of the hardware

- **CRESCO6 HPC system at ENEA**

CRESCO6 is the main HPC system in ENEA, it consists of 434 compute nodes for a total of 20832 cores based on a Lenovo ThinkSystem SD530 platform. Each node is equipped with two Intel Xeon Platinum 8160 CPUs (24 cores), with a clock-frequency equal to 2.10 GHz, a RAM of 192 GB, and a low-latency Intel Omni-Path 100 Series Single-port PCIe 3.0 x16 HFA network interface. The nodes are interconnected by an Intel Omni-Path network with 21 Intel Edge switches 100 series of 48 ports each, bandwidth equal to 100 GB/s, and latency equal to 100ns. The connections between the nodes have 2 tier 2:1 no-blocking tapered fat-tree topology [10]. The High-Performance Linpack (HPL) test (1.4 Pflops peak rate) allowed CRESCO6 to rank 420<sup>th</sup> in the TOP 500 Nov.2018 list;

- **SUMMIT HPC system at ORNL**

SUMMIT is the newest HPC system in ORNL, it consists of 4608 compute nodes with hybrid architecture, and is based on 4608 IBM Power System AC922. Each node contains two IBM POWER9 CPUs and six NVIDIA Volta GPUs V100 all connected together with NVIDIAs high-speed NVLink and has a nodal performance of 42 Tflops. Moreover, each node has 512GB DDR4 + 96GB HBM2 of coherent memory addressable by all CPUs and GPUs plus 800 GB of non-volatile RAM that can be used as a burst buffer or as extended memory. To provide a high rate of input/output throughput, the nodes are connected in a non-blocking fat-tree using a dual-rail Mellanox EDR 100 InfiniBand interconnect [11]. The HPL test (200 Pflops peak rate) allowed SUMMIT to rank at the top of the TOP 500 in 2018 and 2019.

### 2.2 Test case

The test case models two-phase molecular Deuterium ( $D_2$ ) system with 128000 particles: 64000 particles in solid phase + 64000 particles in gas phase at  $T = 10$  K. The initial configuration is made up of a solid slab with a thickness of about  $50 \text{ \AA}$  and two free surfaces, each one with an area of  $200 \times 200 \text{ \AA}^2$ , in coexistence with a gas at  $P = 25$  mbar. The simulation cell has  $L_x = L_y = 200 \text{ \AA}$  and  $L_z = 10^5 \text{ \AA}$ . This system is used as starting configuration for molecular dynamics (MD) simulations with LAMMPS code [12], at constant temperature  $T = 10$  K and constant pressure  $P$ . The pressure along  $x$  and  $y$  directions is kept fixed at 0 mbar, instead along the  $z$  direction is kept fixed at 25 mbar. To evaluate the performance, the average number of timesteps per seconds was considered for a MD simulation with  $10^5$  overall steps.

### 2.3 Results on CRESCO6

The initial study has been conducted on CRESCO6 to assess the scaling of the code with respect to a different partitioning of the 3D spatial domain of the system. Results are reported in Table 1 and Figure 1a-c). LAMMPS splits the workload by spatially partitioning the domain across the number of processes, say  $P$ . A 3D grid is defined with  $P_x \times P_y \times P_z = P$ , where  $P_x$ ,  $P_y$ , and  $P_z$  are the partitioning of the domain along the  $x$ ,  $y$  and  $z$  directions, respectively. The default setting (black symbols) generates a uniform partitioning of the domain that is extremely inefficient for strongly inhomogeneous systems

Table 1: Benchmark results on CRESCO6 for increasing number of cores (up to 4800) with respect to a different partitioning of the 3D spatial domain of the system.

Number of cores	Timestep/s			
	Uniform partitioning of	Manual tuning of	Automatic tuning of	Automatic tuning of
	$P_x P_y P_z$ (default)	$P_x P_y$ and $P_z = 1$	$P_x P_y$ and $P_z = 1$	$P_x P_y$ and $P_z = 2$
48	13	234	234	221
96	13	372	371	361
144	13	459	353	453
192	13	526	430	516
240	12	562	444	553
480	-	582	523	662
960	-	602	320	462
1920	-	606	460	349
3840	-	43	111	375
4800	-	226	255	87

like the ones we are focusing on in this work. In fact, our system is characterized by a confined region occupied by  $D_2$  at solid state (high density) and another more extended region occupied by  $D_2$  at gaseous state (low density), and the system is inhomogeneous along the  $z$  direction. Therefore, the performance is significantly deteriorated ( $\sim 13$  timesteps/s) with the default setting because of an unbalanced workload across the processes. The effect of this strongly unbalanced workload leads to not obtaining any benefit with an increase of the number of cores. A manual tuning of the parameters  $P_x$  and  $P_y$  by fixing  $P_z = 1$  (red symbols) significantly improves the performance increasing the timesteps per second from 12 up to 562 for the 240 cores case. This choice produces a speedup of about 50 and good scalability is achieved up to 144/192 cores. Another option is to have the code automatically tune  $P_x$  e  $P_y$  by fixing  $P_z = 1$  (green symbols) and  $P_z = 2$  (blue symbols), that leads to a similar, albeit slightly inferior, performance with respect to the manual tuning. In panels b) and c) we report the calculated values for the speedup and the efficiency in parallelization. We limited the graphs up to 480 cores because results with greater number of cores are not significant.

## 2.4 Results on SUMMIT

The same system has been scaled on SUMMIT as well and results are reported in Table 2 and panels d-f) of Figure 1. The presence of more powerful nodes on SUMMIT allows us to increase the number of timesteps performed per second with respect to what obtained with CRESCO6 by a factor up to 2. In particular, the more complex architecture of each node allows SUMMIT to attain an improvement with respect to CRESCO6 for a lower total number of central processing unit (CPU) cores especially when all the available graphics processing units (GPUs) are used. For example, the full use of a single compute node with all 6 GPUs increases the performance up to 610 timesteps per second. Value that represents the limit for CRESCO6.

However, the size of the system seems to be too small to benefit from the GPUs. In fact, cases with

Table 2: Benchmark results on SUMMIT for increasing number of cores with respect to a different GPUs.

Nodes	# used cores per node	# used GPUs per node	Total number of cores	Total GPUs	3D Grid	Timestep/s
1	42	0	42	0	default	13
1	42	0	42	0	6×7×1	233
2	42	0	84	0	12×7×1	444
4	42	0	168	0	12×14×1	792
8	42	0	336	0	24×14×1	1180
1	42	1	42	1	6×7×1	377
2	42	1	84	2	12×7×1	439
4	42	1	168	4	12×14×1	793
8	42	1	336	8	24×14×1	691
1	42	6	42	6	6×7×1	610
2	42	6	84	12	12×7×1	851
4	42	6	168	24	12×14×1	1042
8	42	6	336	48	24×14×1	1108

GPUs have worse speedup and efficiency than that in which only CPUs are used. This is mainly due to a frequent data movement between CPUs and GPUs, causing a communication overhead that overcomes the computation. As a final observation, we notice that the default 3D spatial domain again produces 13 timesteps per second even with SUMMIT (see the black symbol in Figure 1a).

### 3 Conclusions

We demonstrated that modeling a system of molecular deuterium ( $D_2$ ) using LAMMPS on state-of-the-art supercomputers is possible with a perfect linear scaling, provided that the size of the system is large enough to ensure that the communication does not dominate over the computation. Moreover, OLCF facilities have demonstrated a relevant benefit using GPUs for a hybrid CPU-GPU parallelization that leads to a speed-up of  $2.0\times$  with respect to using only CPU cores. Future directions are to use this benchmark results to increase the size of the systems up to millions of atoms. This is necessary to simulate wider surfaces and characterize thicker reconstructed solid Deuterium layers.

### Acknowledgments

The authors would like to thank Jan Michael Carrillo and Bobby Sumpter from Oak Ridge National Laboratory for the valuable advice provided.

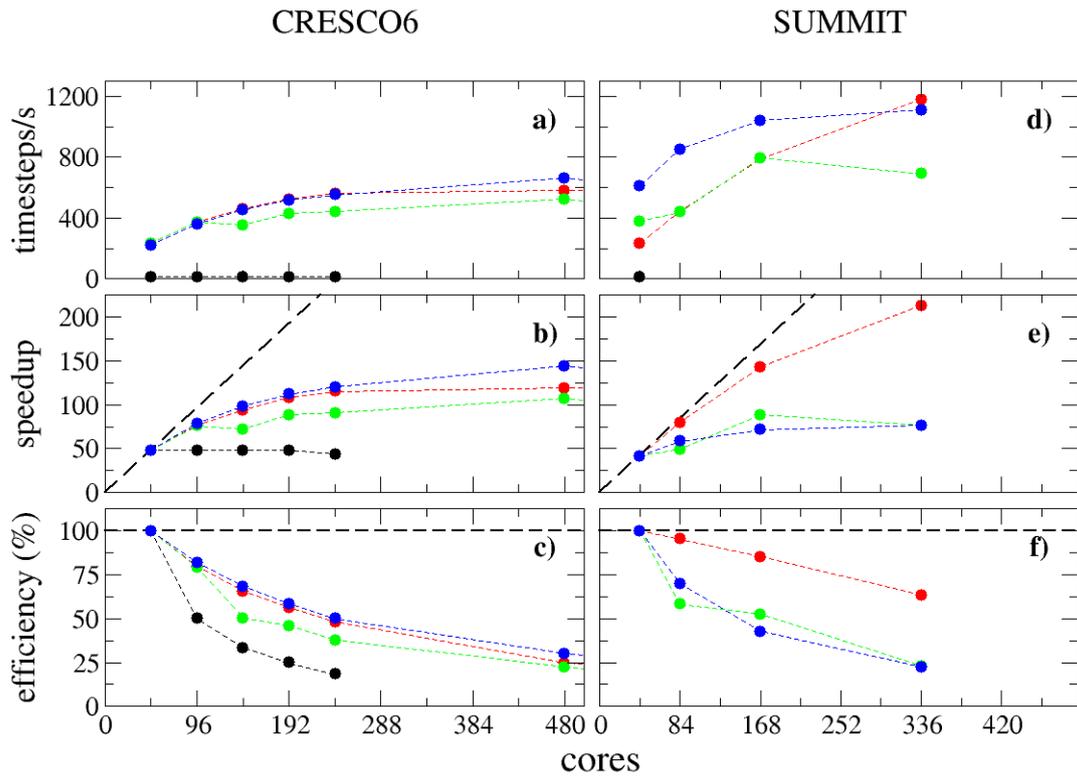


Figure 1: Results for CRESCO6: (a) timesteps per second vs. number of cores, the corresponding speedup (b) and efficiency (c). Black symbols refer to default uniform partitioning  $P_x$ ,  $P_y$ , and  $P_z$  of the 3D spatial domain; red symbols refer to manual tuning of  $P_x$ ,  $P_y$  with  $P_z = 1$ ; green symbols refer to automatic tuning of  $P_x$  and  $P_y$  with  $P_z = 1$ ; blue symbols refer to automatic tuning of  $P_x$  and  $P_y$  with  $P_z = 2$ . Results for SUMMIT: (d) timesteps per second vs. number of cores, the corresponding speedup (e) and efficiency (f). Red symbols refer to simulation without GPUs usage; green symbols refer to simulation with one GPU per node, and blue symbols refer to simulation with six GPUs per node. Black dashed lines refer to ideal cases.

## References

- [1] S. Giusepponi *et al.* Study of solid molecular deuterium  $D_2$  growth for high speed pipegun pellet injector. *Volume edito da ENEA: "High Performance Computing on CRESCO infrastructure: research activities and results 2018"*, ISBN: 978-88-8286-390-6, pages 202–205, 2019.
- [2] S. K. Combs *et al.* Experimental study of the propellant gas load required for pellet injection with ITER-relevant operating parameters. *Fusion Science and Technology*, 68:319–325, 2015.
- [3] F. Bombarda *et al.* State of the art and perspective of high-speed pellet injection technology. *Fusion Engineering and Design*, 124:797–800, 2017.
- [4] A. Frattolillo *et al.* An innovative approach for DEMO core fuelling by inboard injection of high-speed pellets. *Fusion Engineering and Design*, 124:846–849, 2017.
- [5] S. K. Combs *et al.* Gas gun model and comparison to experimental performance of pipe guns operating with light propellant gases and large cryogenic pellets. *Fusion Science and Technology*,

72:404–415, 2017.

- [6] S. K. Combs and L. R. Baylor. Pellet-injector technology. brief history and key developments in the last 25 years. *Fusion Science and Technology*, 73:493–518, 2018.
- [7] A. Frattolillo *et al.* Core fueling of DEMO by direct line injection of high-speed pellets from the HFS. *IEEE Transactions on plasma science*, 46:1429–1435, 2018.
- [8] A. Frattolillo *et al.* Addressing the feasibility of inboard direct-line injection of high-speed pellets, for core fueling of DEMO. *Fusion Engineering and Design*, 146:2426–2429, 2019.
- [9] P. T. Lang *et al.* Optimizing the EU-DEMO pellet fuelling scheme. *Fusion Engineering and Design*, 156:111591, 2020.
- [10] F. Iannone *et al.* CRESCO ENEA HPC clusters: a working example of a multifabric GPFS spectrum scale layout. *2019 International Conference on High Performance Computing & Simulation (HPCS), Dublin, Ireland*, pages 1051–1052, 2019.
- [11] <https://www.olcf.ornl.gov/olcf-resources/compute-systems/summit/>.
- [12] S. Plimpton. Fast parallel algorithms for short-range molecular dynamics. *J. Comp. Phys.*, 177:1–19, 1995.

# VALIDATION AND DEVELOPMENT OF THE FNG SOURCE ROUTINE ON FLUKA

Alessandro Calamida<sup>1,2\*</sup>, Andrea Colangeli<sup>2†</sup>, Mario Pillon<sup>2‡</sup> and Salvatore Fiore<sup>1,2§</sup>

<sup>1</sup>Università degli studi di Roma “La Sapienza”, Physics Department, Rome, Italy.

<sup>2</sup>ENEA Research Centre, Frascati (RM), Italy.

## ABSTRACT.

Creating a source routine able to reproduce the fusion neutron spectrum emitted at the FNG (Frascati Neutron Generator) is an important goal to study the feasibility of future upgrades for the facility. Now it was ported also on FLUKA from the MCNP code. Achieve fusion reactions with the standard routines of FLUKA and MCNP is not possible because they miss the deuterium fusion cross sections. So, first the source routine was converted from MCNP to FLUKA, then, after its validation, it was expanded to represent the whole FNG fusion spectrum. This work is a part of the PhD work of the student Alessandro Calamida.

## 1 Validation of the FNG source on FLUKA

After the conversion in FLUKA, the source code needed to be validated for the new software and to this aim we made a series of simulations.

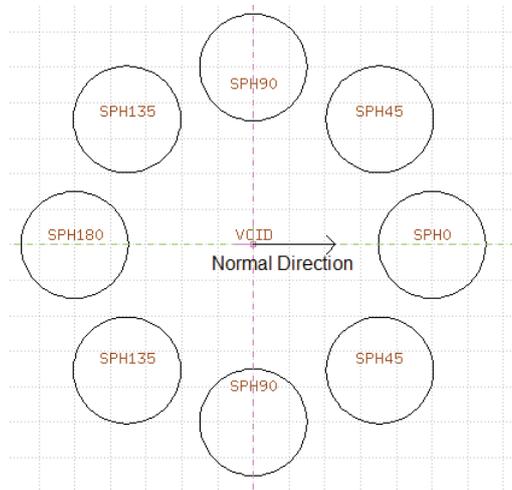


Figure 1: Geometry for the spheres simulations.

The geometry of the simulation is shown in Figure 1. We placed a sphere of radius 15 cm every 45°.

\*E-mail: alessandro.calamida@enea.it.

†E-mail: andrea.colangeli@enea.it.

‡E-mail: mario.pillon@enea.it.

§E-mail: salvatore.fiore@enea.it.

They are placed at 50 cm from the point of origin of the source. The material is vacuum for both the surrounding and the spheres.



Figure 2: USRTRACK scoring for the spheres simulations.

For the scoring for FLUKA we used the USRTRACK detector, a track detector that scores the track density left by a particle in a given volume. It makes the following calculation:

$$\dot{\Phi}(v)dt = n(v)vdt = \frac{dN(v)}{dV} \frac{dv}{dl} = \lim_{\Delta V \rightarrow 0} \frac{\sum_i l_i(v)}{\Delta V}, \quad (1)$$

Where  $v$  is the particle velocity,  $V$  the volume of the region and  $l_i$  the length of the particle track inside the region. The final results are normalized for primary and for the volume of region. It is also differentiated by the energy bin width in the following way:

$$\frac{d\Phi}{dE}, \quad (2)$$

For MCNP the detector used was the F4. This is also a tracking detector that scores the track length density inside the volume of the region. The scoring executes the following calculation:

$$F4 = \frac{1}{V} \int dE \int dt \int dV \int d\Omega \psi(\vec{r}, \hat{\Omega}, E, t). \quad (3)$$

To make a correct validation for the FNG source code in FLUKA we put the same settings in both detectors\*.

For the energy binning FLUKA has a set of fixed group for neutrons below 20 MeV. The width of the energy group is variable and it is based on the ENDF/B-VI.8, ENDF/B-VII.0, Jendl-3.3, Jendl-3.4, Jeff-3.1, etc., data libraries, for more details see Ref. [1]. For the DT reaction we set the range  $\sim 12.5 \div 20$  MeV, instead for the DD one we set  $\sim 1 \div 4$  MeV. So, for MCNP we set the energy bin width to match the ones of FLUKA.

The results of the simulation are plot in the Figure 3. These show that both software give the same output. Making the integral of the spectra, the difference between them are below 0.01%. So, the source is perfectly validated in FLUKA.

---

\*For detector here we do not intend the real detectors but the scoring card that we set in the simulation code. This will be valid for the rest of the essays when we will talk about detectors in simulations.

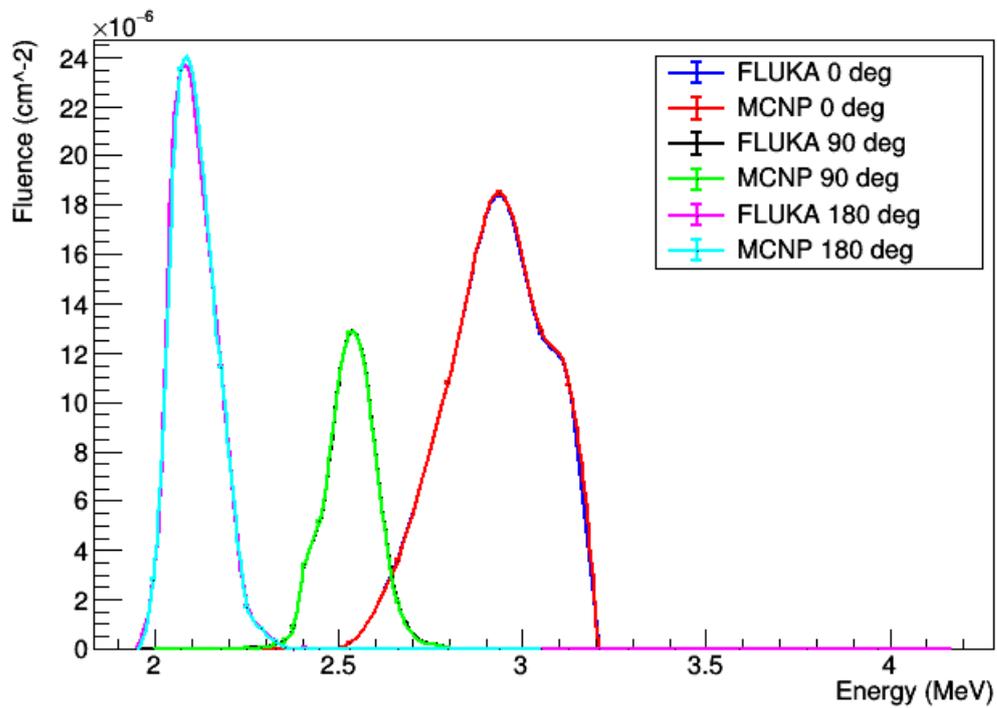
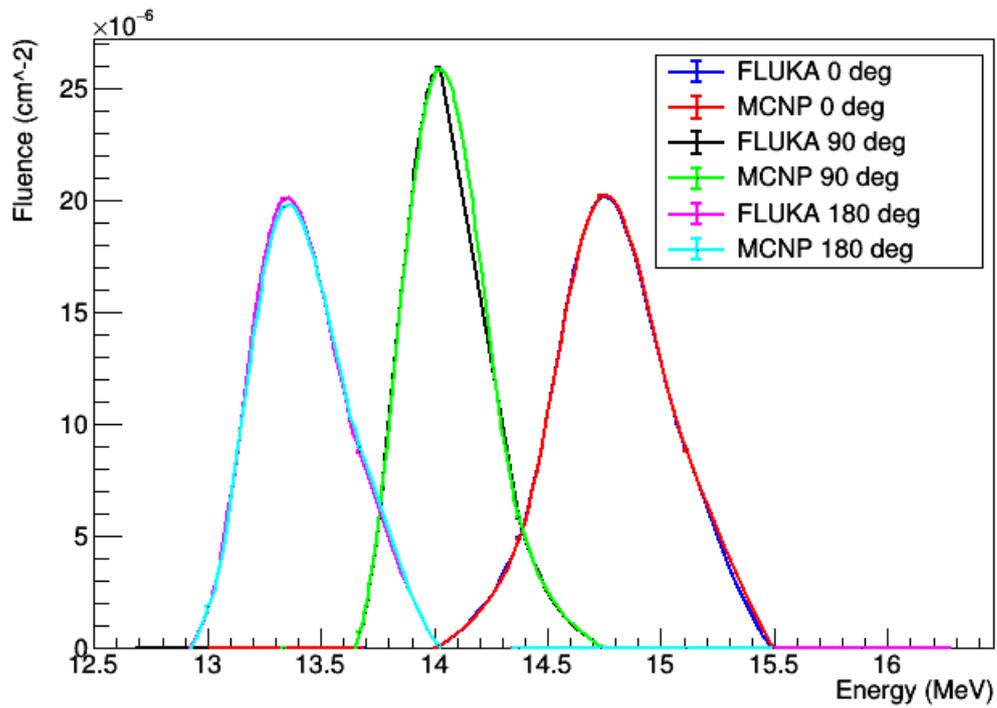


Figure 3: Comparison of the neutron energy spectrum for some angles simulated with both FLUKA and MCNP. The geometry configuration of the simulations is the spheres one in vacuum. The spectra calculated by the two software are the same with no difference between each others. Up there is the DT reaction and below the DD one.

## 2 Development of the FNG source on FLUKA

The original source was designed to reproduce only the spectrum of the neutrons. This was due to the fact that only them were needed at the beginning. A prototype of a source that produce the  $\alpha$  particles spectrum was initialized in the past but it was never fully developed. Moreover, until now the source was only producing one particle per cycle. To study the correlation between  $\alpha$  particles and neutrons, it was necessary a source capable to emit both particles. To achieve this, we used an option inside the source. At every history the routine initializes a variable called NPFLKA. Creating a cycle that increases the value of the variable allows to send more than one particle to the main FLUKA input. Then, it simulates both of them, i.e. a source that emits both the fusion reaction products.

```
WHASOU(1)=deuteron beam energy
WHASOU(2)=Tritium/Titanium or Deterium/Titanium atomic ratio
WHASOU(3)= x coordinate
WHASOU(4)= y coordinate
WHASOU(5)= z coordinate
WHASOU(6)=deuteron beam width
WHASOU(7)=flag, 1 for D-T reaction, 2 for D-D solid reaction
WHASOU(8)=starting cell
WHASOU(9)=flag, 0 for neutron, 1 for positive elons (alpha or helium-3)
WHASOU(10)=flag, 0 for single particle simulation, 1 for dual one, if DT neutron and alpha and if DD neutron and helium-3
SOURCE
      #1 0.26
      #4 0.001
      #7 1
      #10 1
      #13
      #16
      #0 1.4
      #5 0
      #8 1
      #11
      #14
      #17
      #3 0
      #6 0.5
      #9 0
      #12
      #15
      #18
```

Figure 4: The SOURCE card in the FLUKA input. The last option added allows us to chose between the emission of both fusion products or of only one of them. The positive ions produced in the fusion reactions are  $\alpha$  particles for the DT one and the  $^3\text{He}$  for the DD.

In Figure 4 the new SOURCE card is shown with the new option to activate the simultaneous emission of the fusion reaction products. If the flag is activated the source performs two cycle. In the first one the execution runs as in the first source, plus the  $\alpha/^3\text{He}$  kinematic. At the first cycle the  $\alpha/^3\text{He}$  energy, direction and position are passed to the stack of the NPFLKA array. In the second one all the calculations are skipped and the routine directly passes the neutron energy, direction and position previously calculated to the NPFLKA array.

The point of origin of both particles is fixed between one another. The directions are not specular because the reaction products are not emitted back to back. The momentum of the deuterium ion is not zero, so, the sum of the polar angle of the two particle will not be  $180^\circ$ . To reproduce this we created another subroutine modifying the *rot* one and calling it *rotsim*. This keeps trace of the neutrons and  $\alpha/^3\text{He}$  different emission angles.

In Figure 5 the spectra of both of the beam particle is plotted. The neutrons spectra are identical at the ones of Figure 3. The  $\alpha/^3\text{He}$  ones also are equal to the ones expected for the fusion reaction spectra. This validate the simultaneous source code.

To reduce the CPU time for the simulations we still maintain the single source inside the code. In this way, the user can decide from the situation to simulate the neutrons or the  $\alpha/^3\text{He}$  or both.

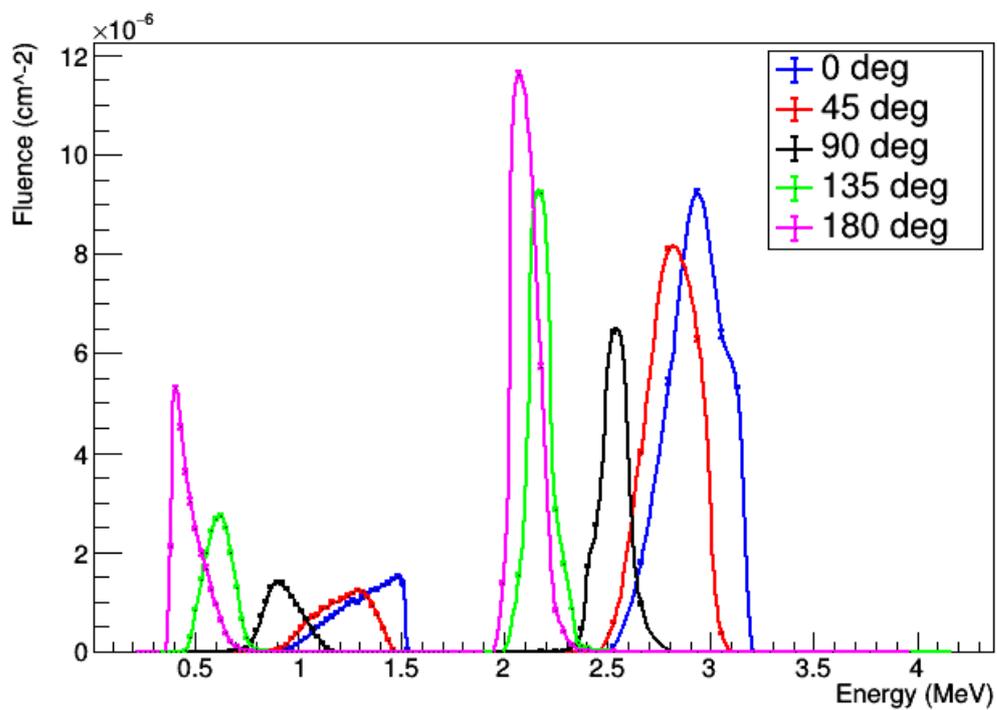
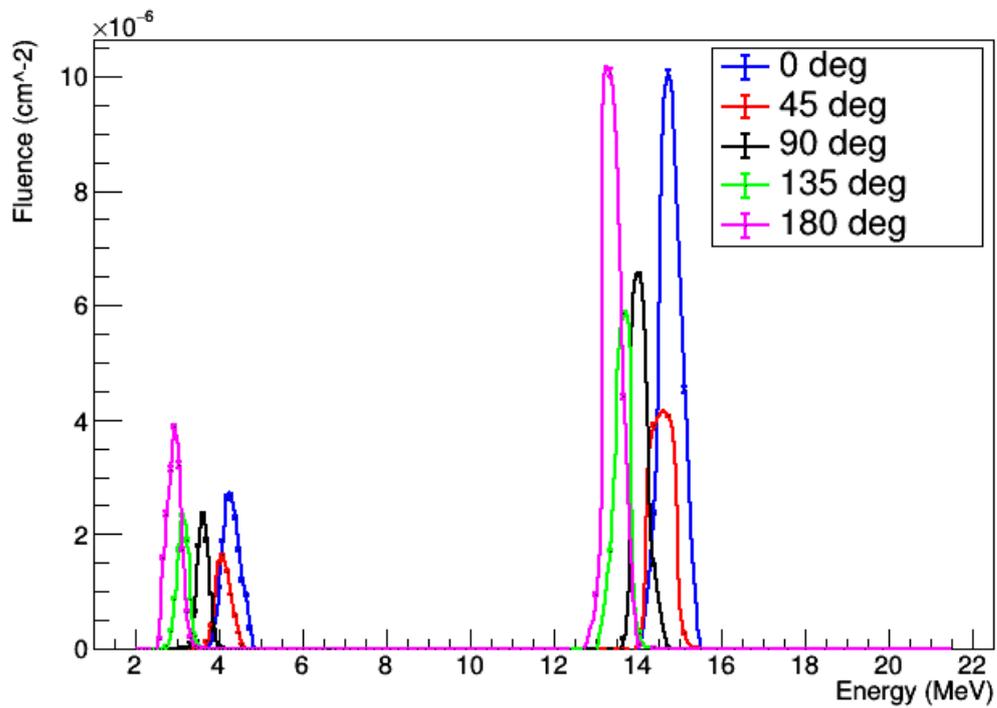


Figure 5: Energy spectrum of neutrons and  $\alpha/\text{He}^3$  produce simultaneously and simulated by the FNG source.

## Acknowledgements

The computing resources and the related technical support used for this work have been provided by CRESCO/ENEAGRID High Performance Computing infrastructure and its staff [2]. CRESCO/ENEAGRID High Performance Computing infrastructure is funded by ENEA, the Italian National Agency for New Technologies, Energy and Sustainable Economic Development and by Italian and European research programmes, see <http://www.cresco.enea.it/english> for information.

## References

- [1] *FLUKA: a multiparticle transport code*, CERN-2005-010, INFN TC 05/11, SLAC-R-773, 2005.
- [2] F. Iannone et al., *CRESCO ENEA HPC clusters: a working example of a multifabric GPFS Spectrum Scale layout*, 2019 International Conference on High Performance Computing & Simulation (HPCS), Dublin, Ireland, 2019, pp. 1051-1052, doi: 10.1109/HPCS48598.2019.9188135.

# NONEQUILIBRIUM ALCHEMY IN SAMPL7 CHALLENGE

Piero Procacci<sup>1\*</sup>, Guido Guarnieri<sup>2</sup>

<sup>1</sup>*Dipartimento di Chimica, Università di Firenze, Via della Lastruccia 3, I-50019 Sesto Fiorentino, Italy*

<sup>2</sup>*ENEA, Portici Research Centre, DTE-ICT-HPC, P.le E. Fermi, 1, I-80055 Portici (NA) (Italy)*

**ABSTRACT.** In the context of the SAMPL7 challenge, we computed, by means of a non-equilibrium (NE) alchemical technique, the standard binding free energy of two series of host-guest systems, involving as a host the Isaac’s TrimerTrip, a Cucurbituril-like open cavitand, and the beta-Cyclodextrin derivatives designed by Gilson and coworkers. The adopted NE alchemy combines enhanced sampling molecular dynamics simulations with driven fast out-of-equilibrium alchemical trajectories to recover the free energy via NE theorems. Performances were good, confirming the reliability of the computational approach and exposing, in some cases, some important deficiencies of the GAFF2 non-polarizable force field.

## 1 Introduction

The SAMPL initiative [1, 2, 3] periodically proposes community-wide blind challenges aimed at assessing computational techniques as standard predictive tools in rational drug design. The SAMPL systems generally consists of a series of host-guest pairs for which the standard binding free energy must be predicted, given the chemical structure of the partners and the experimental conditions (pH, temperature and pressure) used in the measurements. In the SAMPL7 challenge,[4] the organizers included three host systems, namely the Triptycene walled glycoluril trimer (codename TrimerTrip), various mono-3-substituted  $\beta$ -cyclodextrin analogues(codename CD), and the Gibb Deep Cavity Cavitands or Octa-acids (codename GDCC). In SAMPL7, only *one* ranked submission for each of the host-guest systems was allowed. This strict rule was introduced to avoid the practice, largely adopted in the previous SAMPL6 challenge, of filing multiple submissions with small variants in the hope of hitting the target with “multiple shots on goal“.

In this study we present our ranked predictions for the TrimerTrip and CD host-guest systems, done using a classical molecular dynamics (MD) approach, based on enhanced sampling (ES) of the fully coupled end-states followed by a swarm of nonequilibrium alchemical trajectories with fast switching (FS) of the guest-environment interaction by way of a driven alchemical coupling parameter. From the resulting work distribution, the free energy is recovered exploiting the well known Crooks or Jarzynski NE theorems. The adopted computational protocol is identical to that described in Ref. [5], termed as fast-switching double annihilation method (FSDAM). FSDAM is tailored for CPU-based HPC systems such as CRESCO6[6], relying on an efficient OpenMP/MPI hybrid algorithm whereby the nonequilibrium FS trajectories or equilibrium ES trajectories (each parallelized on the OpenMP layer using a force decomposition strong scaling scheme) are assigned to the MPI layer.

---

\*Corresponding author. E-mail: procacci@unifi.it.

## 2 Methods

Structural details of the host-guest systems for the TrimerTrip and CD SAMPL7 challenges can be found at Ref. [4]. All simulations were performed with the ORAC6.1 program[7] using FSDAM as described in Ref. [5] Briefly, in FSDAM the dissociation free energy is determined in two consecutive stages: in the first stage, the bound and unbound state are sampled at full coupling using Replica Exchange with Solute Tempering (REST).[7] In the second stage (fast switching, FS), starting from the canonical sampling at full coupling, we launch a swarm of independent and concurrent non-equilibrium trajectories where the ligand-environment alchemical coupling parameter is rapidly switched off to zero coupling (ligand in the gas phase). The annihilation free energy in the two branches of the alchemical thermodynamic cycle is recovered from the work distribution exploiting the Crooks and Jarzynski theorems.

The Force Field (FF) parameters of the host and guests molecules were prepared using the Primadorac interface.[8] Each system was solvated in about 1000 OPC3[9] water molecules in a cubic box of  $\simeq 31$  side-length.

Production MD simulations (both REST and FS) were run in the isothermal-isobaric ensemble, with temperature control ( $T=298$  K) using Nose-Hoover thermostat and pressure control ( $P=1$  atm) as described in Ref. [10] Lennard-Jones non-bonded interactions were truncated with a 13.0 Angstrom cutoff, whereas long-range electrostatics were handled with the PME[11] method using  $\alpha = 0.37 \text{ \AA}^{-1}$ , 1  $\text{\AA}$  spacing for the gridded charge array and a 4-th order B-spline interpolation. SHAKE constraints were applied to bonds involving hydrogen atoms, and the simulation was integrated using a five step RESPA integrator.[10]

In the REST stage, we scaled the torsional and 14 non bonded interactions of the solute (host+guest) up to 0.1 (i.e. 3000 K) using eight replicas with the scaling protocol described in Ref. [7]. Eight replicates of REST simulations were simultaneously launched on 64 MPI processors for 4 ns on each state, sampling 480 *target state* configurations for the bound host-guest system and for the free guest in bulk. In the bound state, a weak COM-COM harmonic tethering potential ( $k=0.052 \text{ kcal mol}^{-1} \text{ Angs}^{-2}$ ) was imposed to prevent the guest from drifting off in the solvent. The CRESCO6 job for the REST stage involved ten nodes running 64 MPI processes each using 6 threads on the OpenMP layer for a total of 384 cores. A REST job for a host-guest pair was completed in about 4 hours.

In the FS stage, the guest, in the bound and unbound state, was annihilated in 0.36 and 0.24 ns, respectively, in 480 independent trajectories (corresponding to 480 MPI processes) starting from the corresponding points sampled in the REST stage. The bound and unbound annihilation protocol stipulates that the guest atomic charges are first switched off, followed by the annihilation of Lennard-Jones (LJ) interactions. Each of the 480 MPI processes used 6 OpenMP threads, for a total of 2880 cores. A FS job for a host-guest pair was completed in less than 20 minutes.

Annihilation free energy estimates (bound and unbound states) are based on the work distribution (WD) produced in the FS stage. If the WD passed the Anderson-Darling (AD) and the Jarque-Bera (JB) normality tests, the annihilation free energy is calculated using the Gaussian estimator.[5] Otherwise, the statistically boosted Jarzynski average is used, exploiting the decorrelation between discharging work and Lennard-Jones annihilation work.[5] Error on annihilation free energy estimates was computed by bootstrapping with re-sampling in all cases. A finite size correction to the dissociation free energy due to net charges on the ligand has been calculated as described in Ref. [5]. The standard state correction to the dissociation free energy for translational restraint is given by  $\Delta G_{\text{SSC}} = RT \ln(V_{\text{site}}/V_0)$ , where  $V_{\text{site}}$  is the binding site volume.[12]  $V_{\text{site}}$  is computed from the variance of the guest-host COM-COM

distances monitored during the REST stage as  $V_{\text{site}} = 4\pi(2\sigma)^3/3$ .

### 3 Results

Three ranked submissions for the TrimerTrip were filed, namely the present submission (GAFF2/FSDAM), a prediction set using the AMOEBA polarizable force field and FEP, and a submission using a mixed approach, with sampling via standard MD and binding free energy calculation using the semi-empirical tight-binding xtb-GNF program develop by S. Grimme. Correlation plots for the ranked submissions are reported in Figure 1(a) and quality metrics of the corresponding predictions are shown in Table 1. In the Figure 1(a) and in the Table 1 we also report an (unsubmitted) calculation of binding free energies using Autodock4,[13] assuming full flexibility of the ligand and rigidity of the host (using the conformation provided by the organizers at Ref. [4]). Outliers, differing by more than 4 kcal/mol with respect to the experimental value, are marked in red color. While AMOEBA/FEP appears to be the best correlated set to the experimental measures according to Pearson coefficient and to the Kendall rank coefficient  $\tau$ , mean unsigned errors (MUE) are surprisingly minimal for the Autodock set, with GAFF2/FSDAM and AMOEBA/FEP exhibiting similar MUEs. The prediction set based on a mixed MD/QM (semi-empirical) approach is consistently the worst for all quality metrics. GAFF2/FSDAM and AMOEBA/FEP do not have outliers in common. Very likely, discrepancies in g18 and g10 for GAFF2/FSDAM should be ascribed to force field deficiencies, related to the fixed charge approach. In the case of g18, the AM1/BCC charges could underestimate the polarization induced by the host's carboxy groups in the bound state, leading to an extra charge accumulation on the aromatic nitrogen. The neglect of this likely polarization effect can lead to underestimation of the electrostatic contribution to the g18 decoupling in the bound state and hence to the binding affinity. For the diamondoid derivative g10, each amino hydrogen bears a fixed charge of  $0.31 e$ , probably leading in this case to a systematic overestimation of the carboxy-amino electrostatic interactions in the bound state and hence to an overestimation of the binding affinity. Purging g10 and g18 from the GAFF2/FSDAM data-set results in  $R = 0.67$  and  $\text{MUE} = 1.61$ .

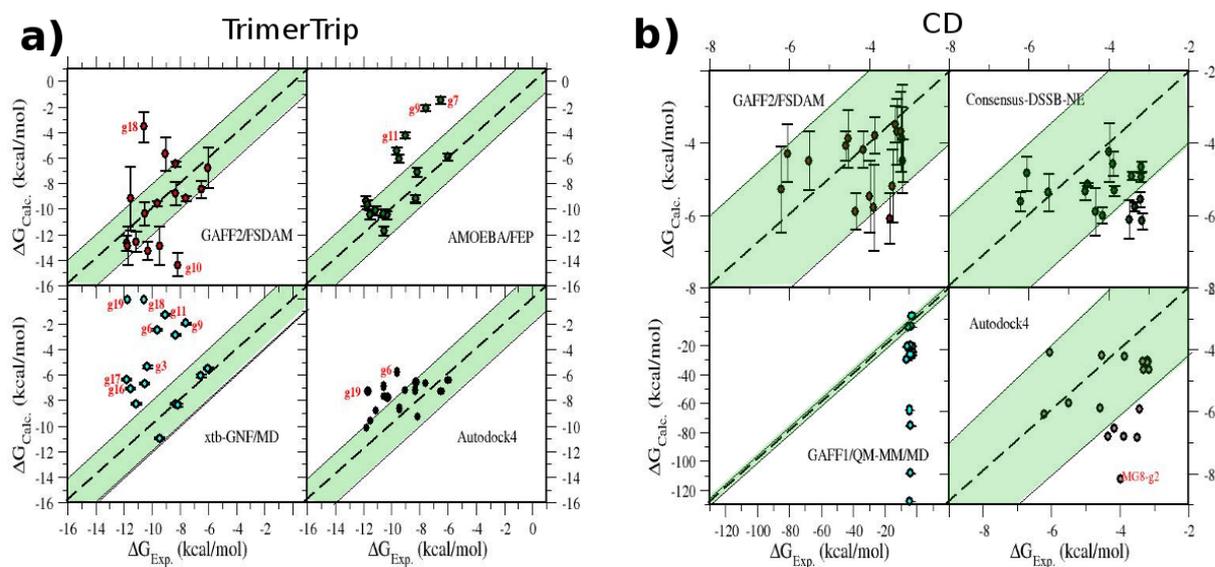


Figure 1: a) TrimerTrip and b) CD correlation plots. Data in the shaded region are within 2.0 kcal/mol of experimental counterpart

In Figure 1(b) we show the correlation plots for the three ranked submissions in the CD host-guest challenge, namely that based on our GAFF2/FSDAM approach, a prediction set relying again on NE alchemical technology using the so-called double-system-single-box approach[14] (DSSB) and finally predicted binding free energies using a QM/MM approach. Also in this case, we provide the plot corresponding to the Autodock4 unsubmitted prediction. The experimental data are clustered in a range of less than 3 kcal/mol and both GAFF2/FSDAM and DSSB correctly and remarkably predicts binding affinities within approximately the same range with no outlier. Given this small experimental range, and given that the systematic uncertainty in fixed charges force fields for solvation free energies are of the order of 2 kcal/mol,[3] MUE appears to be the most meaningful metrics, with GAFF2/FSDAM resulting the best performing method. While the QM/MM ranked prediction set is, quite expectedly,[3, 2] totally off-the-mark (note the expanded scale in the left-bottom correlation plot of Figure 1(b)), the Autodock4 calculation is again in decent agreement with the experimental data, performing only slightly worse than the MD-based NE approaches do (see Table 1).

<b>TrimerTrip</b>					
Method	$R_{xy}$	a	b	MUE	$\tau$
Autodock	0.50	0.35	-4.45	2.00	0.38
FEP/AMOEBA	0.71	1.24	3.94	2.10	0.47
GAFF2/FSDAM	0.35	0.61	-4.06	2.23	0.23
xtb-GNF/MD	-0.06	-0.10	-6.07	4.49	-0.05

<b>CD</b>					
Method	$R^2$	a	b	MUE	$\tau$
GAFF2/FSDAM	0.19	0.17	-3.87	1.01	0.22
DSSB/NE	0.13	0.13	-4.61	1.43	0.02
Autodock	0.13	0.17	-4.78	1.66	0.07
QM/MM	0.10	3.68	-20.36	32.00	0.22

Table 1: Salient data metrics for the assessment of the ranked submissions and of the Autodock prediction set (not submitted) in the TrimerTrip and CD SAMPL7 challenges:  $R^2$ : Pearson’s coefficient;  $a$ : slope of the regression line;  $b$ ; intercept of the regression line; MUE: mean unsigned error;  $\tau$ : Kendall’s rank coefficient.

## 4 Conclusion

In this contribution, we have presented our ranked prediction sets for the TrimerTrip and CD systems in the context of the latest SAMPL7 challenge, using FSDAM, a nonequilibrium alchemical approach combined with enhanced sampling end-state simulations. The performances of our MD-based technique, that uses conventional fixed-charge force fields, is in line with our previous submissions in the SAMPL6 host-guest challenge[3] (done using the very same technology) yielding binding free energies estimates within 2 kcal/mol in most of the cases. Outliers are rare and likely to be ascribed to structural deficiencies of the force field due to the neglect of important polarization effects in the anisotropic environment of some host-guest complexes.

Concerning specifically the SAMPL7 challenge ranked submissions, in the TrimerTrip only one much more computationally demanding approach (based on a polarizable force field) did better than FSDAM. We finally must honestly point out the surprising good results for both the TrimerTrip and CD obtained

using a simple and inexpensive docking approach. While MD simulations certainly provide valuable information on entropic and conformational effects in ligand-receptor association that docking cannot simply deliver by design, modern docking score functions, such as those provided by the efficient Autodock4 software, appear to be remarkably predictive in host-guest systems.

## References

- [1] Hari S. Muddana, Andrew T. Fenley, David L. Mobley, and Michael K. Gilson. The sampl4 host-guest blind prediction challenge: an overview. *J. Comput Aided Mol. Des.*, 28(4):305–317, 2014.
- [2] Jian Yin, Niel M. Henriksen, David R. Slochower, Michael R. Shirts, Michael W. Chiu, David L. Mobley, and Michael K. Gilson. Overview of the sampl5 host-guest challenge: Are we doing better? *J. of Comput. Aided Mol. Des.*, pages 1–19, 2016.
- [3] Andrea Rizzi, Steven Murkli, John N. McNeill, Wei Yao, Matthew Sullivan, Michael K. Gilson, Michael W. Chiu, Lyle Isaacs, Bruce C. Gibb, David L. Mobley, and John D. Chodera. Overview of the sampl6 host-guest binding affinity prediction challenge. *J. Comput. Aided Mol. Des.*, 32(10):937–963, Oct 2018.
- [4] <https://samplchallenges.github.io/roadmap/submissions/>, accessed 23 June 2020.
- [5] Piero Procacci, Massimiliano Guarrasi, and Guido Guarnieri. Sampl6 host-guest blind predictions using a non equilibrium alchemical approach. *J. Comput. Aided Mol. Des.*, 32(10):965–982, Aug 2018.
- [6] F. Iannone, F. Ambrosino, G. Bracco, M. De Rosa, A. Funel, G. Guarnieri, S. Migliori, F. Palombi, G. Ponti, G. Santomauro, and P. Procacci. Cresco enea hpc clusters: a working example of a multifabric gpus spectrum scale layout. In *2019 International Conference on High Performance Computing Simulation (HPCS)*, pages 1051–1052, 2019.
- [7] Piero Procacci. Hybrid MPI/OpenMP Implementation of the ORAC Molecular Dynamics Program for Generalized Ensemble and Fast Switching Alchemical Simulations. *J. Chem. Inf. Model.*, 56(6):1117–1121, 2016.
- [8] Piero Procacci. Primadorac: A free web interface for the assignment of partial charges, chemical topology, and bonded parameters in organic or drug molecules. *J. Chem. Inf. Model.*, 57(6):1240–1245, 2017.
- [9] S. Izadi and A. V. Onufriev. Accuracy limit of rigid 3-point water models. *J. Chem. Phys.*, 145(7):074501, 2016.
- [10] M. Marchi and P. Procacci. Coordinates scaling and multiple time step algorithms for simulation of solvated proteins in the npt ensemble. *J. Chem. Phys.*, 109:5194–5202, 1998.
- [11] U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, and L. G. Pedersen. A smooth particle mesh ewald method. *J. Chem. Phys.*, 103:8577–8593, 1995.
- [12] Piero Procacci and Riccardo Chelli. Statistical Mechanics of Ligand-Receptor Noncovalent Association, Revisited: Binding Site and Standard State Volumes in Modern Alchemical Theories. *J. Chem. Theory Comput.*, 13(5):1924–1933, 2017.

- [13] Garrett M. Morris, Ruth Huey, William Lindstrom, Michel F. Sanner, Richard K. Belew, David S. Goodsell, and Arthur J. Olson. Autodock4 and autodocktools4: Automated docking with selective receptor flexibility. *J. Comput. Chem.*, 30(16):2785–2791, 2009.
- [14] V. Gapsys, S. Michielssens, J.H. Peters, B.L. de Groot, and H. Leonov. *Molecular Modeling of Protein*, chapter Calculation of Binding Free Energies, pages 173–209. Humana Press, 2015.

# LATTICE ATOMISTIC KINETIC MONTE CARLO SIMULATION WITH NVIDIA GPUS

Marco D.Mazzeo<sup>1\*</sup>, Simonetta Pagnutti<sup>2</sup>

<sup>1</sup>*European Centre for Medium-Range Weather Forecasts, Reading, UK*

<sup>2</sup>*ENEA, TERIN-ICT Division*

**ABSTRACT.** This report presents a parallel version of the simulation of the diffusion of point defects within a non-irradiated Fe-2%Cu alloys through a kinetic Monte Carlo (KMC) approach. The use of KMC applications is growing in very different simulation fields, from epidemiology to radiation damage in materials for nuclear, since they provide an easy implementation of the dynamical evolution of complex systems. For this reason, a great effort has been put in KMC parallelization, which remains a very challenging problem for the time dependence of the method that is sequential in character. It belongs to the class of parallel discrete-event simulations, which are widely applied in computer science, economics as well as in engineering and in physics. Thanks to a novel parallel particle-centric approach, a simplified version of a code simulating the precipitation of Cu in bcc Fe, under the driving force of an augmented population of vacancies [1], was modified and adapted to run on NVIDIA GPUs (Graphics Processing Unit).

## 1 KMC simulation model

We consider a non-irradiated material where vacancies are the sole point defects, albeit in a small concentration. Any insoluble specie precipitates towards areas reach of its atoms because of the movements of the vacancies and the specific interatomic interactions. The vacancy population changes for thermally activated jumps simulated through energy evaluations and migratory frequencies. KMC simulates the dynamics of the system by evaluating and randomly selecting thermally activated events. Vacancy movements are directly linked to their migration event frequencies that are evaluated, in the code considered, thanks to a properly trained neural network. The crystal is assumed to be a BCC system hosting a smaller number of Cu atoms and rare vacancies and subject to periodic boundary conditions (PBCs). The evolution of the system is simulated as a state-to-state dynamics corresponding to a Markov walk. The system propagation corresponds to a first-order process with exponential decay statistics:

$$p_{survival}(t) = \exp(-\Gamma_{tot}t)$$

Where  $p_{survival}(t)$  is the probability of the system to already survive in state  $i$ ,  $\Gamma_{tot} = \sum_j \Gamma_{ij}$  is the total escape rate from state  $i$  and  $\Gamma_{ij}$  is the rate constant for escape from state  $i$  to state  $j$

$$\Gamma_{ij} = \Gamma_0 \exp(-E_m(i \rightarrow j) / k_B T)$$

$E_m$  is the local interatomic potential-dependent migration energy,  $\Gamma_0$  is approximatively constant, and  $k_B$  is the Boltzmann's constant. The stochastic algorithm used to propagate the system from state to state is the BKL (Bortz, Kalos and Lebowitz) [2].

At each transition, the time of first escape is incremented by

$$\Delta t = -\ln(r) / \Gamma_{tot}$$

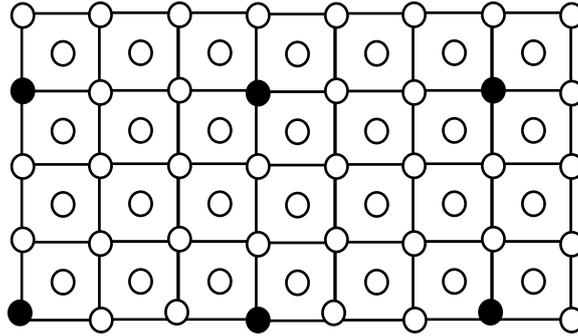
---

<sup>1</sup> Corresponding author. E-mail: [m.mazzeo@ecmwf.int](mailto:m.mazzeo@ecmwf.int), [md.mazzeo@gmail.com](mailto:md.mazzeo@gmail.com)

where  $r$  is a random number uniformly distributed in the  $]0,1]$  interval.

A parallelization of the algorithm is not straightforward since KMC approach has an intrinsic serial character. The scheme adopted is the “null-event” algorithm [3,4].

The defect distribution is initially sparse and rarefied but defects tend to aggregate due to system dynamic. This renders parallel domain decomposition techniques, relying on processor clusters, inefficient because of the unfavourable ratio between the local pre-processor computation and the inter-processor communications; moreover, vacancies can group in a few small regions thus leaving idle processors without vacancies to process. For this reason we have adopted a parallel *particle-centric* approach suited for focussing the computation where it is needed the most, i.e. around the defects.



**Fig.1:** Two-dimensional representation of parallel multi-colour sub-lattice. The sites of one of the  $3 \times 3 \times 2$  interleaved sub-lattices are marked with black filled circles. Vacancies, which can be moved in parallel at each MC step, are those positioned on the currently selected sub-lattice (black filled circles).

The simulation lattice is rethought as comprised of 54 ( $3 \times 3 \times 3 \times 2$ ) distinct sub-lattices of different colours, each featuring the same resolution (Figure 1). The sub-lattice constant is chosen in such a way to guarantee independence among the Local Atomic Environments around the sub-lattices points. Vacancies located on the sites of a chosen sub-lattice are processed in parallel by using the synchronous, parallel generalization of the BKL algorithm proposed by Martinez & al. By assuming that the number of sub-lattice points equals that of the Processing Elements (PE), following Lobacevsky [5], we can classify the Martinez algorithm as “supersaturated” and the new approach as “unsaturated”, that turns to a “saturated” one in the not so probable case that all the sub-lattice points host a vacancy.

A simplified pseudo-code of the core simulation code is the following:

- Randomly select one colour among all the 54 coloured sub-lattices
- Identify the  $N_v$  vacancies residing in the current sub-lattice, store their IDs in an array and communicate this array to the GPU.
- Compute in parallel the migratory frequencies corresponding to the events associated with the selected vacancies  $k_{tot,v} = \sum_j k_{ij}$ ,  $v = 1, N_v$  and copy them from the GPU to the CPU;
- find with the CPU the maximal migratory frequency  $K_{max} = \max\{k_{tot,v}\}$ ,  $v = 1, \dots, N_v$  and communicate it to the GPU;
- In the GPU, for all the vacancies, in parallel, an event is carried out on the basis of the migration frequencies, including null events:  $k_{0,v} = K_{max} - k_{tot,v}$ .  
For this step, we must ensure independency of random numbers sequences.
- the new vacancies coordinates are copied from the GPU to the CPU;
- the CPU updates the number of computed events and sample the time increment for all the vacancies:  
 $\Delta t = \frac{\ln(r)}{K_{max}}$ ; where  $r \in ]0,1]$  is a suitable random number.

A CPU-based stage identifies the vacancies featuring the currently selected colour and their coordinates, thanks to fast bit-wise operations, and compactly store their indices. Then, a CPU-to-GPU memory copy, of vacancies identifiers, is carried out with a CUDA (Compute Unified Device Architecture) function. A GPU kernel calculates the migratory frequencies corresponding to the selected vacancies and copy them into an array stored in the CPU's memory, which computes the maximal value. Then, in parallel, a GPU function selects the events to handle and updates the needed lattice data and global memory parts. The new vacancies coordinates are sent to the CPU (in a compressed format - they are required for determining the vacancy colours in the successive parallel stage), which also updates the number of computed events and the simulation time.

The chosen parallel approach employs the many small cores available within a GPU and delegates to them the calculation of the data of the events, their random selection and the lattice data updating through the GPU scheduling of numerous threads.

## 2 Running Code

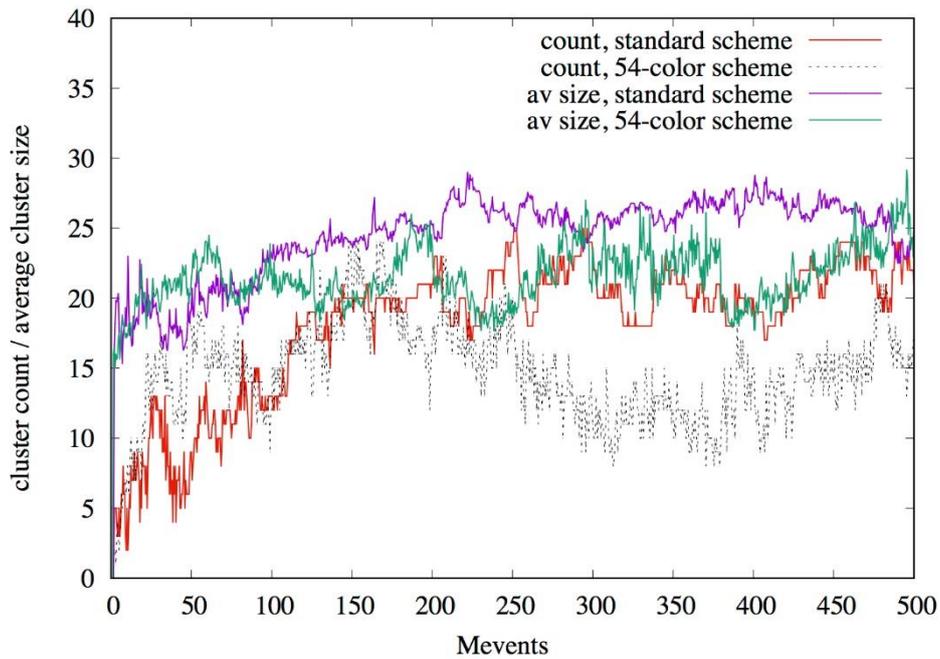
### 2.1 Compilation and execution

The adopted NVIDIA GPUs are part of the ENEA's platform CRESCO, which is an integrated computational infrastructure comprising geographically distributed high performance parallel computers [6]. GPUs can be accessed in different manners. One way resorts to FARO (Fast Access to Remote Objects) a ThinLinc application which allows to access to ENEAGRID/CRESCO [7] with a web browser.

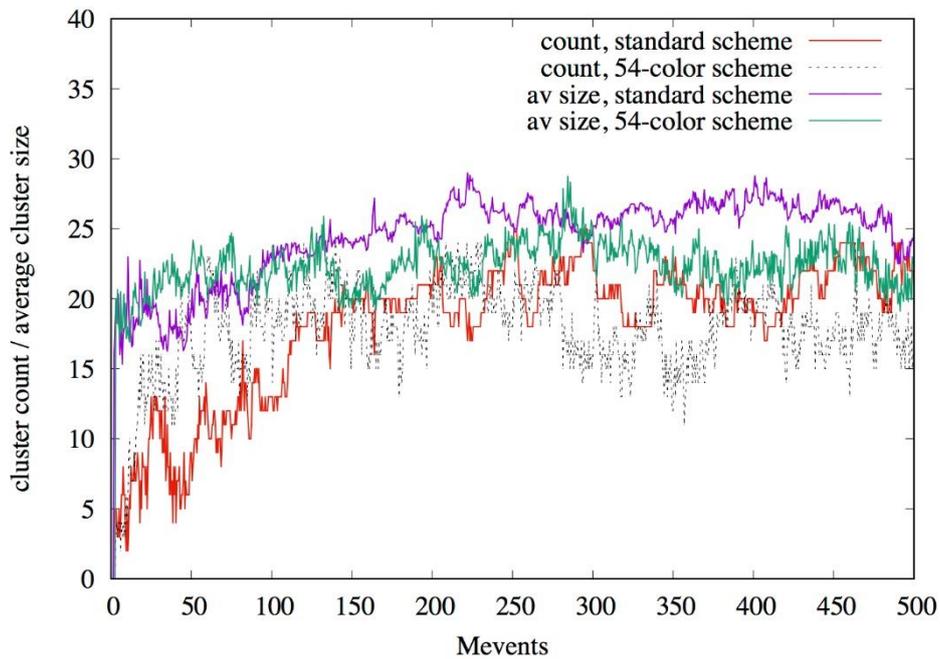
Some CPU code parts employ OpenMP and 20 threads (*ad hoc* for the CPU of CRESCO4 leveraging the Tesla K40 GPU).

### 2.2 Validation

The null-event parallel kinetic MC method was implemented and applied to a particle-centric 54-color scheme. A comparison with the serial approach was carried out to validate the new parallel scheme. The tested system comprises a BCC matrix of  $50 \times 50 \times 50 \times 2$  lattice sites hosting 50 vacancies and 5000 Cu fully immersed in a Fe atoms lattice. The number of clusters and the average cluster size were calculated every 500K events for 500M-event runs. The plots for the standard scheme and for the new approach are given in Figure 2 (null events are counted as the others). By comparing the red line with the black dashed one, it seems that the two different simulation methods produce similar cluster counts, especially at the beginning of the simulation and at its end. By comparing the other two lines, the two methods also produce similarly sized clusters and hence the new 54-color scheme seems statistically efficient and correctly implemented. The results would present an even less discrepancy if the null events are not reported in the graph (Figure 3).



**Fig.2:** Number of clusters ("count") and average cluster size ("av size") obtained with the sequential approach and with the new parallel algorithm as a function of MC events (null events count as the others) [BCC 50x50x50x2-site matrix with 50 vacancies, 5000 Cu atoms].



**Fig.3:** Results of Figure 2 without reporting the null events.

### 2.3 Performance

For systems comprising up to a few hundreds of vacancies, the performance of the new parallel CUDA code, in terms of events computed per second, is similar to that of an OpenMP version running on 4+ CPU cores since most of the cores of a GPU likely remain idle if it is equipped with several tens or hundreds of cores.

However, if the GPU has many hundreds of cores and the simulated system has many hundreds of vacancies, the GPU code performance may be more than 10 times faster than that running on a multi-core CPU. A simulation with  $400 \times 400 \times 400 \times 2$  lattice sites and 25600 vacancies shows that the CPU version, running on 10 cores of a CRESCO4 node hosting the Tesla K40 GPU and featuring multi-threading, performs 1.4 Gevents/h whereas the parallel CUDA code leveraging the Tesla K40 GPU performs 12 Gevents/h – a 8.6-fold speedup. The cited CUDA and GPU-based code performance is such that a 100 billion-event run (needed for well exploring the behaviour of said large systems) can be executed in a few hours only; thereby rendering possible lengthy,  $O(10)$  seconds simulations.

## Conclusion

A kinetic MC simulator of the defects diffusion, in a non-irradiated Fe-2%Cu alloy, was examined and modified to handle multiple independent defects in parallel. This was possible through the combination of a particle-centric method with a multi-colour-based adaptation of the null-event MC scheme. The code was modified to run on NVIDIA GPUs by means of the programming model CUDA that enables the good use of NVIDIA GPUs at a low programming overhead. The number of events processed per hour on a high performance GPU may be of the order of  $10^{10}$  for large systems; this performance surpasses that of a multi-core CPU implementation by various orders of magnitude and renders viable the simulation of very large systems.

## Acknowledgements

This work has significantly benefitted from the interaction with ENEA researchers Simonetta Pagnutti, Gianclaudio Ferro and Marcello Galli.

## References

- [1] N. Castin, L. Malerba, Calculation of proper energy barriers for atomistic kinetic Monte Carlo simulations on rigid lattice with chemical and strain field long-range effects using artificial neural networks, *J. Chem. Phys.* 132, 074507, 2010.
- [2] A.B. Bortz, M.H.Kalos, J.L. Leibowitz, *Jour. Comput. Physics*, 17, 10-18, 1975.
- [3] A. Chatterjee, D. G. Vlachos, An overview of spatial microscopic and accelerated kinetic Monte Carlo methods. *J. Computer-Aided Mater Des* 14, 253-308, 2008.
- [4] E. Martínez, P. R. Monasterio, J. Marian, Billion-atom synchronous parallel kinetic Monte Carlo simulations of critical 3D Ising systems. *J. Comput. Phys.* 230, 1359-1369, 2011.
- [5] B. Lubachevsky, *Jour. Comput. Physics*, 75, 103-122, 1988.
- [6] <https://www.eneagrid.enea.it/>; <https://www.eneagrid.enea.it/CRESCOportal/>
- [7] F. Iannone et al., "CRESCO ENEA HPC clusters: a working example of a multifabric GPFS Spectrum Scale layout," 2019 International Conference on High Performance Computing & Simulation (HPCS), Dublin, Ireland, 2019, pp. 1051-1052, doi: 10.1109/HPCS48598.2019.9188135

## AUTHOR INDEX

1. Abraham, Alex, 143
2. Adani, Mario, 87
3. Albino, Andrea, 110
4. Alderuccio, Daniela, 155
5. Almedral, Juan A., 46
6. Alonzo, Massimo, 101
7. Ambrosino, Fiorenzo, 7, 29
8. Anav, Alessandro, 62
9. Angelone, Maurizio, 205
10. Arcidiacono, Nunzio, 137
11. Bianco, Simona, 143
12. Boccaletti, Stefano, 46
13. Bracco, Giovanni, 7
14. Briganti, Gino, 87
15. Briganti, Matteo, 110
16. Briguglio, Sergio, 183
17. Buonocore, Francesco, 110, 126, 177, 225
18. Burgio, Nunzio, 42
19. Burrese, Emiliana, 67
20. Caiffi, Barbara, 205
21. Calabrese, Rolando, 71
22. Calamida, Alessandro, 231
23. Cappelletti, Andrea, 87
24. Caputo, Stefano, 34, 75, 171
25. Carillo, Adriana, 62
26. Cecere, Donato, 137
27. Celino, Massimo, 7, 110, 126, 133, 177, 225
28. Cherubini, Nadia, 92, 163

29. Chiariello, Andrea Maria, 143
30. Chinnici, Marta, 210
31. Chiovaro, Pierluigi, 80
32. Ciancarella, Luisella, 87
33. Colangeli, Andrea, 205, 231
34. Console Camprini, Patrizio, 117
35. Conte, Mattia, 143
36. D'Elia, Ilaria, 87
37. D'Isidoro, Massimo, 87
38. De Chiara, Davide, 210
39. De Marzi, Gianluca, 21
40. De Nicola, Antonio, 34, 75, 171
41. De Rosa, Matteo, 7
42. Del Nevo, Alessandro, 80
43. Di Troia, Claudio, 183
44. Donati, Greta, 34, 75, 171
45. Esposito, Andrea, 143
46. Falconi, Mattia, 105
47. Fiore, Salvatore, 231
48. Fiorillo, Luca, 143
49. Flammini, Davide, 159, 201, 205
50. Fogaccia, Giuliana, 183
51. Fonnesu, Nicola, 205
52. Formisano, Giovanni, 7
53. Frattolillo, Antonio, 225
54. Funel, Agostino, 7
55. Fusco, Valeria, 183
56. Gandolfo, Giada, 92, 163
57. Garcia-Muller, Pablo Luis, 126
58. Georges, Jean Philippe, 210
59. Giacomazzi, Eugenio, 137
60. Giusepponi, Simone, 126, 225

61. Gori, Paola, 177
62. Grishina, Anastasiia, 210
63. Guarnieri, Guido, 7, 87, 210, 237
64. Gutierrez, Ricardo, 58
65. Iacovelli, Federico, 105
66. Iannone, Francesco, 7
67. Ienco, Andrea, 38
68. Kor, Ah-Lian, 210
69. Lepore, Antonio, 29
70. Lepore, Luigi, 92, 163
71. Leyva, Inmaculada, 46
72. Lupo Pasini, Massimiliano, 225
73. Mancini, Giorgio, 51, 133
74. Mariano, Giovanni, 159, 201, 205
75. Marzo, Giuseppe A., 92, 163
76. Massaro, Arianna, 151
77. Mayo-Garcia, Rafael, 126
78. Mazzeo, Marco D., 243
79. Mazzone, Giuseppe, 159, 201
80. Meineri, Carlo, 167
81. Migliori, Silvio, 7, 155, 225
82. Milano, Giuseppe, 34, 171
83. Moro, Fabio, 159, 201, 205
84. Mosca Conte, Adriano, 177
85. Munaò, Gianmarco, 34, 171
86. Munoz Garcia, Ana B., 151
87. Noce, Simone, 159, 201
88. Pagnutti, Simonetta, 243
89. Palma, Massimiliano, 62
90. Palombi, Filippo, 7
91. Palumbo, Biagio, 29
92. Paoli, Paola, 38

93. Pavone, Michele, 151
94. Pecoraro, Salvatore, 7
95. Picchia, Franca Rita, 137
96. Pieranti, Antonio, 87
97. Pillon, Mario, 231
98. Pimpinella, Maria, 101
99. Pizzirusso, Antonio, 34, 75, 171
100. Ponti, Giovanni, 7
101. Procacci, Piero, 237
102. Prosino, Pier Paolo, 177
103. Pulci, Olivia, 177
104. Romanelli, Francesco, 201
105. Romeo, Alice, 105
106. Rondeau, Eric, 210
107. Rossi, Patrizia, 38
108. Sannino, Gianmaria, 62
109. Santagata, Alfonso, 42
110. Santomauro, Giuseppe, 7
111. Scaglione, Salvatore, 122
112. Sendina-Nadal, Irene, 46
113. Simoni, Fabio, 7
114. Sposito, Gianluca, 29
115. Struglia, Maria Vittoria, 62
116. Tapfer, Leander, 67
117. Totti, Federico, 110
118. Turuncoglu, Ufuk Utku, 62
119. Villari, Rosaria, 159, 201, 205
120. Vlad, Gregorio, 183
121. Wang, Tao, 96
122. Wang, Xin, 183
123. You, Jeong-Ha, 201
124. Zhang, Rui, 21

125. Zhao, Ying, 171
126. Zola, Danilo, 122
127. Zollo, Giuseppe, 187



ENEA  
Promotion and Communication Service

*www.enea.it*

Printed at the ENEA Technographic Laboratory – Frascati

December 2020

