
# HPC-EUROPA3 FIRST RESEARCH PROJECTS DIRECTORY

*October 2018*

HPC-Europa3 aims to maintain a persistent and high-quality service offering Transnational Access to the most advanced HPC infrastructures in Europe which are currently available to the scientific research community. This third HPC-Europa programme builds on the success of previous Transnational Access programmes, which operated in Europe for almost two decades.

As would be expected from a Transnational Access programme, HPC-Europa3 offers collaborative visits to research groups in a different country from the one in which the participant currently works. Researchers collaborate with a scientific host in their own area of research, while also receiving mentoring from one of the partner HPC centres, to help them make the best use of the HPC facilities for their research.

HPC-Europa3 therefore represents a powerful mechanism to support researchers to become proficient in the use of HPC and numerical simulation.

HPC-Europa3 is a truly multi-disciplinary programme attracting applications from all computational disciplines. From its beginning in 2017, more than 250 applications were made to the programme.

The present First Projects Directory presents the results of 38 projects that have been completed in the first year of the programme. The Projects Directory series will continue in the next years with similar reports to document scientific projects that will be accomplished during the HPC-Europa3 life time.
Reading the contributions presented in this report, the importance of the scientific results can be seen. This is also evidenced by the number of publications produced, confirming the significant contribution of the HPC-Europa programme to support European researchers in computational science.

# Summary

# Chemistry

# COMPUTATIONAL MODELLING OF THE STABILITY AND GROWTH OF SUPPORTED GOLD NANOPARTICLES ON MAGNESIA

*S. Francis, A. Roldan*

*Cardiff School of Chemistry, Cardiff University, Cardiff, United Kingdom*

## Introduction

The comprehension of how stable nanoparticles form is of continuing importance to increase the efficacy of catalysts utilised in many areas of science and industry. Regarding gold nanoparticles specifically, their effective use in catalysts in a wide range of reactions [1] is dependent on the stability of particles of particular morphologies.

My work currently focuses upon analysis of the formation and stability of small gold clusters (< 2nm) on magnesia support.

Ferrando *et al.* found good structural agreement of computationally acquired 2-3 nm sized gold clusters with experimental findings obtained through atomic scale imaging [2]. This project sought to relate Ferrando's group work with my own to enhance understanding of the formation gold clusters; with a particular interest in the evolution of their shape with size.

## Methods

Low energy Au clusters are initially produced using a global optimisation procedure. The systems are modelled by two potentials, a model derived from the second moment approximation to the tight-binding model (SMATB potential) for the metal-metal interactions ($E_{MM}$) and a second potential fitted to DFT to model the metal surface interaction ($E_{MO}$). The potentials have been described by Verscich *et al*. [3] and Ferrando *et al.* [4]

$$E_{MO} = \sum_{i=1}^{n} E_i(x, y, z + \mathcal{Z})$$

For the metal-surface interaction the above equation is used. The dependence of the energy on the distance from the substrate (z) is reproduce via a Morse-like function, whereas a periodic cosine function is used to model the dependence of the interaction energy on x and y coordinates. Z is the number of nearest neighbours, which is calculated to include all neighbours within 1.25 $r_0$, where $r_0$ is O-O distance.

Once the global optimisation using Monte Carlo methods has been performed, the clusters are optimised using spin polarised periodic planewave DFT, employed in the Vienna Ab Initio Simulation Package (VASP). A PBE functional has been used to calculate exchange and correlation contributions and long-range interactions were calculated using Grimme's dispersion correction DFT-D3. Projector-augmented wave (PAW) pseudopotentials as implemented in VASP were used to describe the core electrons. A kinetic cut-off was set at 450 eV.

## Results

In the first instance, the Monte Carlo procedure was used to determine the structure of gold clusters at sizes of 50 and 70 atoms (1-1.5nm). Subsequent DFT optimisations performed on these clusters showed that the shapes determined were retained. However, with some distortion seen, increasingly with the introduction of dispersive forces into the calculations an example can be seen in Figure 1.

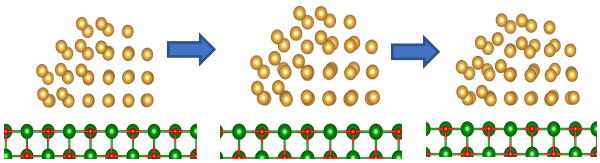


Figure 1 -50 atom Au clusters, initial structure and subsequently after DFT optimisation without and finally with dispersion corrections.

In addition to these larger cluster sizes, smaller clusters (< 20 atoms) were also calculated with the Monte Carlo programme to make comparisons of the lowest energy structures determined previously with DFT alone.

The structures produced from the Monte Carlo approach were not retained in DFT optimisation. An example of this is shown in Figure 2. The perpendicular lowest energy supported structure is 0.277 eV lower in energy than the square based pyramid structure found using Monte Carlo.

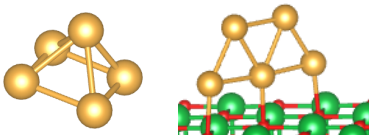


Figure 2- The first image shows the predicted lowest energy structure produced by Monte Carlo and the second image the lowest energy structure found in DFT for a 5 atom cluster.

## Conclusions

The inconsistency between the Monte Carlo approach and the DFT calculations for small cluster sizes was expected. This is owing to the nearest neighbour inclusion into the $E_{MO}$ parameterisation being restricted to 0, 4 or 8, representing a single atom adsorption, epitaxial monolayer or epitaxial bilayer. Smaller clusters will have a greater ratio of atoms with co-ordination numbers varying from these, so the interaction is not accurately modelled by the potential.

The retention of the 50 and 70 atom structures when put into DFT eludes to the conclusion that the Monte Carlo procedure gives a good approximation of cluster structure for larger cluster sizes (> 1 nm). These clusters are currently being used toward a secondary aim, in which the effect of the interfacial epitaxy is analysed.

## References

[1] Duan Z et al, Chem. Phys, 5486–5490, 2016; [2] Ferrando R et al, Phys Chem, 5:131-137, 2014; [3] Vervisch W et al, Phys. Review B, 65(24): 245411, 2002; [4] Ferrando R et al, J. Chem. Phys, 130:174702-174703, 2009.

## Acknowledgements

The work has been performed under the Project HPC-EUROPA3 (INFRAIA-2016-1-730897), with the support of the EC Research Innovation Action under the H2020 Programme; the author gratefully acknowledges the support of R. Ferrando at Università degli Studi di Genova and the computer resources and technical support provided by CINECA.

# EFFECT OF QM-REGION SHAPE AND SIZE ON SURFACE QM/MM EMBEDDED-CLUSTERS

*H. Jenkins[1], M. Kick[2], A. Logsdail[1], H. Oberhofer[2]*

*[1]Cardiff Catalysis Institute, Cardiff University, United Kingdom; [2]Chair for Theoretical Chemistry, TUM, Germany*

## Introduction

The current most popular computational method in modelling surface reactions is periodic planewave density functional theory (DFT) [1]. Periodic DFT is both expansive in its functionality, and easy to use. However, there are still some major limitations. For instance, the preferred basis representation of planewaves is inefficient for high-level hybrid-DFT approaches; large models are needed for isolated catalytic chemistry, with many tens or hundreds of atoms included but redundant in calculations, as all atoms need to be treated quantum mechanically, even if they are far from the relevant "active site". In addition, periodic surface models cannot be used to model charged defects due to technical incompatibilities when trying to include any compensating background charge.

As an alternative, hybrid quantum- and molecular-mechanical (QM/MM) modelling can be coupled with an embedded cluster approach, which combines the high accuracy of QM modelling, with the high efficiency of MM modelling (MM), all in an aperiodic model that reproduces the bulk environment [2].

The most challenging component of a QM/MM simulation is ensuring the correct balance between QM and MM energies and forces, the effect of which is most prominent at the interface between regions. Most obviously, setting up a balanced boundary that ensures the wavefunction of the QM region matches that of a continuous structure. To ensure this accurate electronic representation, it is necessary to use pseudo-potentials or link atoms at the boundary interface, which preserve the bonding nature of the QM region. It is also necessary to choose a shape and size for the QM region that minimises any spurious quantum confinement effects, ideally preserving the electronic structure of the bulk material.

## Methods

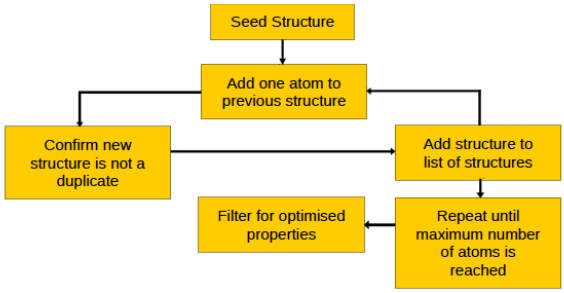
The aim of this work was to develop a Python module that would provide users with suitably-shaped QM regions, for a QM/MM simulation of any solid state network material, based on properties. The properties we have focused on are related to geometric and electronic quantities, such as the surface area, internal bonding, symmetry, dipoles, etc.

The realised software takes an unpartitioned QM/MM cluster as input, along with some options, such as required atoms, a maximum size, and desired properties. The software then returns a selection of QM/MM embedded-cluster models for each size that best optimise the desired geometric/electronic properties. Additionally, the underlying code is sufficiently modular such that it could be repurposed for similar tasks and integrated into larger pieces of software.

The new module has been written in Python, as it is a widely used, understood, and flexible language in the computational chemistry and HPC communities. Additionally, Python is a prudent choice given that it is intended for the code to be integrated into an upcoming Python release of ChemShell [3], Py-ChemShell [4].

The more difficult aspects of the module were twofold; ensuring that it could efficiently process the wide breadth of possible structures for large QM region sizes, and determining which geometric/electronic properties should be considered for optimisation. Taking advantage of symmetry, and controlling how the lattice network is explored kept computational costs reasonable, while the experience of hosts Matthias Kick and Harald Oberhofer was essential for choosing relevant geometric/electronic properties.

## Results




The written code works standalone, in that is not reliant on any non-standard Python modules. Furthermore, the inputs and outputs are both in Py-ChemShell readable formats, which will aid in incorporating the module into other software packages.

## Summary and Future Work

The final step for the developed Python modules is to apply the QM/MM embedded-clusters in calculations so as to determine if the chosen target observables result in the most reliable representation of bulk properties. At this stage, it is expected that minimising the number of bonds connecting the QM region and the MM region will reduce the amount of disrupted electron density, and therefore increase reliability, but this remains to be proven. This will be tested in simulations that will use the Tier-1 facilities both at HLRS and in the UK. In agreement with the visit hosts, it is intended for the developed software to be distributed freely on an OpenSource licence through the Py-ChemShell distribution.

## References

[1] Becke J, Chem. Phys. 140, 18A301 (2014); [2] Groenhof, Methods in Molecular Biology, Ch.3 (2013); [3] Sherwood et al, J. Mol. Struct. (Theochem.) 632, 1 (2003); [4] see www.chemshell.org.

## Acknowledgements

The work has been performed under the Project HPCEUROPA3 (INFRAIA-2016-1-730897), with the support of the EC Research Innovation Action under the H2020 Programme; in particular, the author gratefully acknowledges the support of Matthias Kick and Harald Oberhofer at the Chair for Theoretical Chemistry at TUM and the computer resources and technical support provided by HLRS.

# MECHANISM OF LONG-RANGE ELECTRON TRANSFER IN RESPIRATORY COMPLEX I

*M. Röpke[1], A.P. Gamiz-Hernandez[1], M.P. Johansson[2,3], V.R.I. Kaila[1]*

*[1]Department Chemistry, Technische Universität München, Garching, Germany; [2]Department of Chemistry, University of Helsinki, Helsinki, Finland; [3]Helsinki Institute of Sustainability Science, Helsinki, Finland*

## Abstract

The respiratory complex I is a redox-driven proton pump in aerobic respiratory chains that couples transfer of electrons through a 100 Å FeS wire to proton pumping up to 200 Å away from the terminal electron acceptor site. The aim of this HPC3 project was to probe the molecular mechanism of the long-range electron transfer process in complex I by quantum chemical DFTbased models as well as QM/MM simulations. To this end, we derived microscopic electron transfer parameters to estimate non-adiabatic electron transfer rates, and to simulate the transfer process quantum mechanically. The project showed how the electron transfer process is modulated by the spin energetics and by the surrounding protein dynamics.

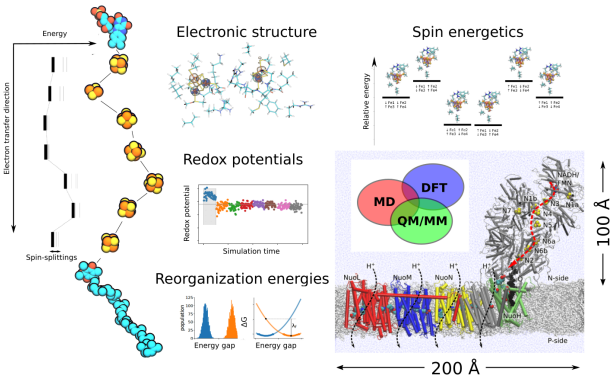
## Introduction

Complex I (NADH:ubiquinone oxidoreductase) functions as the initial electron acceptor in mitochondrial and bacterial electron transport chains. It catalyzes the electron transfer (eT) between NADH and quinone (Q) through a 100 Å chain of iron-sulfur (FeS) centers. At the end of this chain, Q is reduced to quinol (QH2), which the enzyme further couples to proton pumping across its membrane domain, up to 200 Å from the initial site of Q reduction. The structure of complex I has recently been resolved from several different species that has allowed us for the first time to address the molecular mechanisms of this remarkable, 300 Å charge transport process by using multi-scale computational simulations.

## Results

In order to understand the molecular principles behind the initial 100 Å long-range eT process in the FeS chain, we studied in this HPC3-Europe project the structure and energetics of intermediate steps of the charge transfer process in the ironsulfur (FeS) centers in complex I. To this end, we optimized the molecular structures of each FeS center using DFT-based quantum chemical cluster models in combination with multi-state quantum mechanics/classical mechanics (QM/MM)- molecular dynamics (MD) simulations in all possible spin-configurations. Based on these calculations, we derived redox potentials, reorganization energies, and electron couplings to elucidate how the electron transfer process is modulated by the surrounding protein dynamics.

The spin energetics of the FeS centers in complex I is particularly challenging for quantum chemical calculations, as the centers contain between 10 and 18 unpaired electrons that are antiferromagnetically coupled. The spin-coupling was modelled using the broken-symmetry DFT approach to study how the spin energetics affect the electron transfer reactions. Our combined results show the importance of protein dynamics and with our derived microscopic electron transfer properties we shed light into how the electron travels across the ca. 100 Å FeS wire on 100 µs timescales, consistent with EPR experiments.




## References

[1] Kaila VRI (2018) J. R. Soc. Interfaces 15: 20170916. [2] Di Luca A et al, (2017) PNAS 114, E6314–E6321. [3] Gamiz-Hernandez AP et al, (2017) JACS 139:16282-16288. [4] Kaila VRI (2018) in Chem. Biol. 5, “Mechanisms of Primary Energy Transduction in Biology”, Ed. M. Wikström, The Royal Society of Chemistry.

## Acknowledgements

The work has been performed under the Project HPCEUROPA3 (INFRAIA-2016-1- 730897), with the support of the EC Research Innovation Action under the H2020 Programme; VRIK gratefully acknowledges the support of the host, Dr. Mikael P. Johansson at University of Helsinki, and the computer resources and technical support provided by CSC–The Finnish IT Center for Science.

# QUADRUPLEX STRUCTURES AND HALOGEN BONDS

*G. Paragi* [1,2]

[1]*Department of Medicinal Chemistry, University of Szeged, Szeged, Hungary;* [2]*MTA-SZTE Biomimetic Systems Research Group, Hungary.*

## Introduction

Quadruplexes are well-known supramolecular structures and they have important role in many biological processes, as well as in supramolecular chemistry. One of the best-known examples is the guanine (Gu) quadruplex, where four-membered Gu ring quartets are stacked above each other (see Figure1). Previously, we have pointed out that Gu-tetramer contains an extra stabilization effect (cooperativity), which is the consequence of its special hydrogen bond donor-acceptor arrangement [1]. Concerning the similarity between hydrogen and halogen bonds [2], it was demonstrated that when hydrogen bonds are substituted homogenously with halogen bonds in Gu-tetramer (called as N-halo-guanine quartets), the new systems can also support cooperative effect [3]. However, mixed N-halo-guanine quartet systems (see Figure 1, left), which contains hydrogen and halogen bond together or different halogen bonds within the same structure has not been investigated yet. So, hydrogen bonds were systematically substituted for halogen bonds and interaction energies, synergies or charge transfers were calculated. Additionally, solvent effect was also investigated in an implicit way, as the presence of water drastically decrease the cooperativity in the original Gu-tetramer.

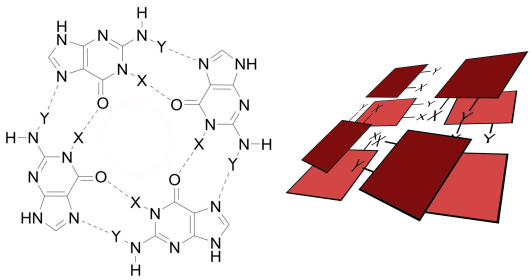


Figure 1 - (left) The investigated tetramers where X, Y can be H, Br, Cl or I. A general quadruplex structure (right) based on the investigated tetramers.

## Methods

All calculations were carried out with the ADF program at BLYP-D3(BJ)/TZ2P level of theory, except the double layer quadruplex optimization. In the latter case DZP basis set was applied during the optimization process, but energies were calculated with TZ2P basis set. VDD charges was taken in charge analyses and interaction energy of the constituents in the tetrad, as well as the synergy of the systems were determined. The synergy of a complex can express the presence of extra stabilization effect in the system and defined as $\Delta E_{syn} = \Delta E_{int} - \Delta E_{sum}$ . Here $\Delta E_{int}$ is the interaction energy of the quartet which we gain when deformed constituents (bases in the geometry taken from the tetramer) form the quartet. The $\Delta E_{sum}$ is the sum of the individual pairwise interactions for all possible base pairs in the quartet. For a detailed description of calculations see the supporting information in [3].

## Results

Regarding tetramers, 16 different complexes could be generated by the systematic substitution of hydrogen atoms in position X and Y. Optimizations were carried out with and without planar restraint. In some of non-restraint cases I obtained systems with near-planar geometry similar to the Gu tetrad, but in most of the cases the final geometry distorted notably. Concerning the interactions, there are many systems with halogen bonds, which provide stronger interaction or larger synergy than the original Gu-tetrad, especially in the presence of iodine. On the other hand, the chloride substitution(s) always weaken the interactions and the cooperativity in the systems. In implicit solvent calculations the interaction energies were decreased significantly and the synergy vanished. This is in line with the results found in the original Gu-tetrad [1]. Considering stacking complexes, the interaction energy or synergy values in vacuum were similar to those found in tetramers. In solvent they were between the tetrameric vacuum and solvent results. So, similar tendencies (decreasing interaction and synergy in water) were found in stacking complexes. Concerning the VDD charges for the H, Cl, Br atoms the charge transfer between adjacent units increased or decreased in line with the change in the interaction energy during the stepwise (1, 2, 3 units taking in the ring geometry) construction of the tetramers. However, this tendency was broken in the presence of iodine, because of charge backdonation occurred via the iodine bond.

## Conclusions

I found that halogen substituted systems behaves very similar to the original Gu-tetrad: it can support cooperative effect which vanish in the presence of solvent. Although single halogen bond usually weaker than a hydrogen bond, stronger interaction could be found in the investigated structures. All these outcomes were not affected by the presence of a second layer.

## References

[1] Fonseca Guerra C et al, Chem. Eur. J., 17:12612, 2011; [2] Wolters LP et al, Chemistry Open, 1:96, 2012; [3] Wolters LP et al, Phys. Chem. Chem. Phys., 17:1585, 2015.

## Acknowledgements

The work has been performed under the Project HPC-EUROPA3 (INFRAIA-2016-1-730897), with the support of the EC Research Innovation Action under the H2020 Programme; in particular, the author gratefully acknowledges the support of Department of Theoretical Chemistry at the Free University of Amsterdam (VU) and the computer resources and technical support provided by SARA center.

# ORGANIC MATERIALS FOR SINGLET FISSION

*E. Siagri*

*University of Trieste, Trieste, Italy; University of Groningen, Groningen, The Netherlands.*

## Introduction

The world's energy demand is increasing because of population growth and industrial evolution; therefore, it is important to find a renewable and inexhaustible source of energy. The resources available in the world are getting depleted and the burning of fossil fuel is damaging the earth ecosystem mainly because of the emission of greenhouse gases. While fossil fuels are destined to be replaced given their finite nature and negative environmental effects alternative eco-friendly sources of energy like solar energy, wind energy, hydropower and geothermal are of great interest. In fact these energy sources can reliably provide a transition to a low-carbon economy which is more environmentally sustainable. In the last decade there has been a significant increase in the production of renewable energy. About 73.5% of the electricity has been generated by nuclear and fossil fuels while 26.5% is produce by renewable source [1]. Solar energy count for 1.9% making it one of the lesser utilized sources of energy. The lack in solar energy production is mainly due to the fact that this technology is not yet fully reliable. Among the disadvantages and limitation there is the problem that the solar energy has intermittency issues. In fact during the night is not possible to generate energy and during cloudy or rainy weather the energy production is bound to decrease. This unpredictability makes solar energy panels a less reliable solution. Moreover, for a continuous supply of electric power, photovoltaic panels require storage batteries which will inevitably increase the investment cost. Despite its disadvantages photovoltaic technology has countless advantages: environmentally friendly, free and unlimited, silent and without complex mechanical parts. Therefore, solar energy could be one of the best options given the fact that it is the most abundant energy source of renewable energy. The sun bathes the earth an average of 1.8 x 1014 kW making it the biggest free source of energy [2]. Solar energy is not harmful for the earth ecosystem and its simplicity and affordability makes it a good solution for fulfilling the global energy demand.

## Singlet Fission

Singlet fission (SF) is a multiple exciton generation process (MEG). MEG is the phenomenon where the absorption of a single photon leads to the excitation of multiple electrons. If this process can be controlled we can get higher efficiencies since multiple excitons may lead to multiple charge carriers after absorption of one photon. Qualitative studies have showed that a solar cell capable of quantitative singlet fission it is possible to increase the Shockley-Queisser limit form about 1/3 to nearly 1/2. Therefore, this phenomenon has high potential use in photovoltaic materials. In particular Singlet Fission (Figure 1) is a process in which a chromophore, in an excited singlet state, shares its excitation energy with a near chromophore molecule in its ground state, causing the conversion of both chromophores in their triplet excited states which then separates and become independent.

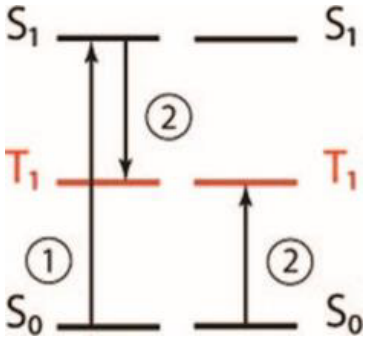



Figure 1 - Singlet fission scheme: First the chromophore on the left gets excited to S1 and then it shares its energy with the adjacent chromophore creating a pair of T1 states [3].

For SF to occur reliably and fast a couple of fundamental conditions needs to be met. First of all it is necessary to have two chromophores because at least two excitation sites are needed to accommodate two triplets excitations. Moreover there are two important energetic condition that must be satisfied; the excitation energy of the first singlet excited state has to be about two times the excitation energy of the first excited triplet, and the excitation energy of the second triplet must be at least two times higher than the excitation energy of the first triplet. The first condition is needed to ensure that the conversion of the singlet excited state to the triplet pair is fast and competitive. SF will be fast if it is slightly exoergic or at least isoergic. If the process is too exoergic then we will have loss of efficiency by heat generation. Slight endoergicity is tolerable since the missing energy can be made up by thermal energy. The second condition is needed to avoid the formation of the S0 + T2 state by making it endoergic and thus very slow which translates in a longer lived T1 state [4]. However, a long triplet lifetime can have its side effects; it will ensure charge separation but also it will increase the opportunity for quenching which must be avoided. The coupling between the chromophores is also an important factor; it needs to be strong enough to ensure a fast SF on the time scale of picoseconds and weak enough to allow the triplets to be nearly degenerate and diffuse apart easily so to permit them to undergo two independent charge separation steps. If the triplets cannot diffuse apart sufficiently rapidly they may destroy each other by triplet-triplet annihilation forming a higher excited singlet or triplet.

## Molecules Investigated

The systems studied in this project are all conjugated organic molecules that show absorption in the visible light range. These molecules are reported in Figure 2.

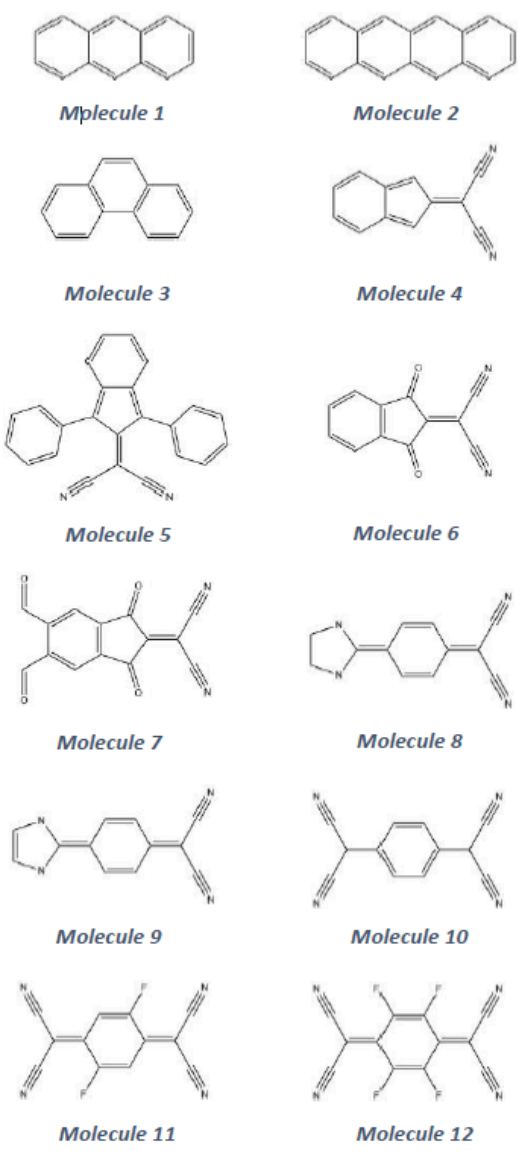



Figure 2 - Organic molecules investigated in this project

## Methods

The geometry optimization procedure was performed using DFT with TZP basis set in the ADF code. To evaluate if the organic molecules studied satisfy the two most important energetic criteria for a fast and competitive SF we performed multiple TDDFT calculations. These calculations were done by varying the basis set and the exchange-correlation functional (BP86, B3LYP, BHandH). Based on these calculations we decided which molecules were the most promising for undergoing SF. After the first exploratory study we performed additional calculations on the most promising systems using ab-initio methods such as Coupled Cluster (CC2), CASSCF/NEVPT2, DDCI and NOCI. The study was then extended at the dimers of these molecules since SF can occur only in the presence of at least 2 chromophores.

## Results

Between molecules 1, 2 and 3 the most suited for SF is tetracene (3) where the difference 2E(T1)-E(S1) indicates that the SF process is slightly endoergic and viable with a little provision of heat. Anthracene (1) and phenanthrene (2) are too endoergic. In the tetracene molecule E(T2) is 1.92 times E(T1) which means that the annihilation to yield S0 + T2 is conveniently slow. Between the three molecules tetracene has been further investigated since it appears to be the best one for our purpose. In particular we studied the dimers of the tetracene crystal. In the dimer the difference between the energy of the S-S and S-T transitions compared to the monomer is about half electron volts. The S0 → S1 is a HOMO-LUMO transition and it has charge transfer character. For this reason it is plausible that TDDFT calculations are not sufficient to get a reasonable understanding of the system.

Between molecules 4, 5, 6 and 7 the ones that seems good candidates for SF are molecule 4 and 5. Molecules 6 and 7 are not interesting for our purpose since their energy difference indicates that SF would be too endoergic. Instead molecules 4 and 5 are slight exoergic and have been further investigated with Coupled Cluster and CASSCF methods. These ab-initio methods have given interesting results. The CC2 calculation on molecule 4 proved that the double excitations contribution is only about 8% and the energy difference 2E(T1)-E(S1) points towards an isoergic SF. Instead the CASSCF/NEVPT2 calculation gives and energy difference slightly endoergic.

Both TDDFT and CASSCF/NEVPT2 calculations for molecules 8 and 9 shows a slightly endoergic SF.

Molecules 10, 11 and 12 are very similar and present the same behavior. TDDFT analisys indicates that SF in these molecules is too exoergic, however CASSCF/NEVPT calculations shows a far less exoergic SF.

## Conclusions

In order to find a suitable material for SF we investigated several conjugated organic molecules with TDDFT and various ab-initio methods. The study has been done on the isolated molecules and on their corresponding dimeric systems where the electronic coupling often complicates the picture. We have gathered a lot of information which will have to be compared with experimental data in order to rationalize the results and validate the methods used.

## References

[1] REN21, Global Status Report 2018: Secretariat Renewable Energy Policy Network for the 21st Century, 2018; [2] Panwar N et al, Renewable and Sustainable Energy Reviews, Role of renewable energy sources in environmental protection, vol. 15, no. 3, pp. 1513-1524, 2011; [3] Smith MB et al, American Chemical Society, Singlet Fission, vol. 110, pp. 6891-6936, 2010; [4] Zdenek Halvas J et al, Journal of Chemistry, Guidance for Mutual Disposition of Chromophores for Singelt Fission, vol. 56, pp 96-106, 2015.

## Acknowledgements

The work has been performed under the Project HPC-EUROPA3 (INFRAIA-2016-1-730897), with the support of the EC Research Innovation Action under the H2020 Programme; the author gratefully acknowledges the support of dr. R.W.A. Havenith, Faculty of Science and Engineering (RUG) and the computer resources and technical support provided by SURFsara.

# EXTENSION OF THE HOMA PARAMETERS FOR A NEW CATEGORIES OF BONDS: BP, BAS, ALN, ALP, ALAS, GAN, GAP AND GAAS

*K.K. Zborowski [1], I. Alkorta[2], J. Elguero[2], T.M. Krygowski[3]*

*[1]Faculty of Chemistry, Jagiellonian University, Kraków, Poland; [2]Instituto de Quimica Médica (CSIC), Madrid, Spain; [3]Department of Chemistry, Warsaw University, Warsaw, Poland.*

## Introduction

Aromaticity is an important property, with a strong influence on stability of chemical compounds. Thus, studying aromatic properties is a very frequent topic of chemical research. There are many methods of the estimation of aromaticity level in chemical species. The HOMA (Harmonic Oscillator Model of Aromaticity) index [1] is one of the most widely used methods for aromaticity level evaluation. Its main advantages are very fast calculations from simple analytic formula and necessary data accessibility. In order to calculate the HOMA value for an aromatic ring one need to know only bond lengths in it. They are often available from experimental data, if not they can be evaluated from theoretical calculations. Thanks to its easy use and the aromaticity importance the HOMA index can be very useful for many scientists if only..., if only HOMA model parameters are known for all types of bonds in the studied aromatic ring. Yes, as almost everything also the HOMA index has its significant disadvantage. The problem is that HOMA parameters must be known for all bond types included in the studied ring. At the beginning rings with the CC bonds only could be studied [1], then extension to the most typical heterocyclic bonds (CC, CN, CO, CP, CS, NN, NO) [2] and BN [3] were made. In order to overcome such a difficulty, we parametrized CSe [4], BC [5] and BB [6] bonds. Because still not all potentially aromatic rings can be investigated by the HOMA model we have recently parametrized CAl, CGa and CAs bonds (unpublished data). In this project we want to open the HOMA model for few other types of bonds.

## Methods

Parameters of the HOMA model can be determined using experimental or theoretical bond lengths. Unfortunately, in case of studied in this work bonds it is difficult to find any reasonable set of experimental data. Thus, parametrization must be done using calculated bond lengths. What more, we have shown recently [7] that the best solution is to use the same level of theory for estimation of studied compounds bond lengths as for calculations of the HOMA model parameters. Due to, we performed calculations using different theoretical approaches. They are two popular DFT functionals (B3LYP and PBE0) and the classical MP2 method. In order to check the influence of dispersion correction, we performed also calculations with and without the D3 dispersion correction for both chosen DFT functionals. This way, other scientists will have a fairly wide set of HOMA parameters and they will be able to choose the best calculation method for their research. We employed 6-311++G** basis set.

## Results

Necessary bond lengths and force constants at chosen levels of theory were determined and HOMA parameters were calculated. In order to see how the results evolve through Periodic Table, the previously parametrized BN bond [3] was included in the study. In the next step, HOMA values were calculated for some hydrocarbons with one of the newly parametrized bond inserted. A representative result for benzenes with a heterocyclic bonds are presented below.

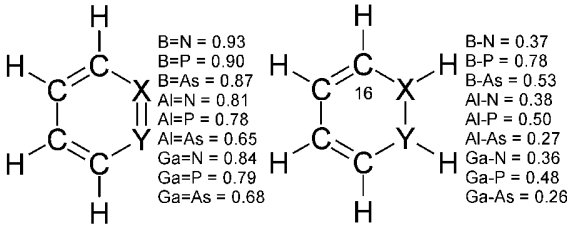



Figure 1 – HOMA index values for heterobenzenes, B3LYP/6311++G** X = B, Al, Ga; Y = N, P, As.

## Conclusions

Calculated HOMA values suggest quite high aromaticity of hydrocarbons with inserted one among studied here heterocyclic bonds. This is observed especially for species containing double boron bonds in the ring (B=N, B=P or B=As). Aromaticities of compounds with other inserted X=Y bonds are moderate while these with X-Y bonds are weakly aromatic. The biggest aromaticity change is noticed between B=N and B-N bonds. Obtained HOMA parameters widely opens possibility of next studies on aromaticity of compounds containing bonds formed between elements from 13th and 15th groups of Periodic Table.

## References

[1] Krygowski TM et al, Tetrahedron Lett., 13:3839-3842, 1972; [2] Krygowski TM, J. Chem. Inf. Comput. Sci., 33:70-78, 1993; [3] Madura ID et al, Tetrahedron, 54:14913-14918, 1998; [4] Zborowski KK et al, Polish J. Chem., 83:477-484, 2009; [5] Zborowski KK et al, Struct. Chem., 23:595-600, 2012; [6] Zborowski KK et al, Struct. Chem., 24:543-548, 2013; [7] Andrzejak M et al, Struct. Chem., 24:1171-1184, 2013.

## Acknowledgements

The work has been performed under the Project HPC-EUROPA3 (INFRAIA-2016-1-730897), with the support of the EC Research Innovation Action under the H2020 Programme; in particular, the authors gratefully acknowledge the support of the IQM (CSIC) in Madrid and the computer resources and technical support provided by BSC.

# Earth Sciences & Environment

# RUNNING A HIGH RESOLUTION EARTH SYSTEM MODEL

*T. Arsouze[1], L.P. Caron[2], M. Castrillo[2]*

*[1]IPSL/ENSTA-ParisTech, Paris, France; [2]Barcelona Supercomputing Center, Barcelona, Spain*

## Introduction

This project has been designed to initiate the collaboration between two very active groups in terms of climate modelling: IPSL/ENSTA on one side with a strong expertise on Mediterranean climate, and BSC on the other side, expert in production of large configuration of Earth system models and on code optimization. A simulation of the Earth System at the ground-breaking resolution of ~12km has been initiated and will be continued in the framework of the European PRIMAVERA project. The dataset produced will be extremely valuable for comparison with regional simulations of the Mediterranean basin produced at IPSL. Also, the new kind of ultra high-resolution simulation aims at preparing the future of climate modelling both for the technical aspects and the scientific analysis. This collaboration will open possibilities for more tools and simulations expertise sharing in the future between these two key players in HPC.

## Methods

Within the framework of the PRIMAVERA project, the BSC is entitled to produce a climatic simulation at 12 km resolution. This cutting-edge effort is both a technical challenge in terms of computational resources involved as well as a scientific tool of a new genre for the community that will prepare the future of climate simulations. The BSC has developed a coupled version of the EC-Earth 3.2 (a European Community Earth-System Model) [1] at a ground-breaking horizontal resolution of 12km in each climate system component (Glob-12km: EC-Earth3_IFS-T1279_NEMO-ORCA12). It offers an un-precedented tool to evaluate the improvement of models due to resolution increase at global scale but also benefits from the added value of regional models (access to finer temporal and geographical scales, better representation of extrema, explicit representation of physical processes, etc). It thus appears as a key link between Global Climate Models and Regional Climate Models, and will provide insight on errors of RCMs models due to GMCs forcing on lateral boundary conditions, influence of better representation of large-scale phenomena on regional climate, influence of regional climate on large scale, etc. BSC has developed a whole environment of tools for climatic simulation: workflow management, post-processing tools, issues handling, etc. An important part of the work consisted in adapting these tools into the supercomputer to the new configuration. Considering the huge amount of data needed in input and generated in output of the model, this type of configuration has demonstrated to be much more difficult to handle than any standard configuration.

## Results

Although a proof of concept of this ultra-high resolution configuration had already been performed, implementing it into a production workflow, using the experimental protocol defined for the HighResMip project for latter distribution to the community has proven to be another challenge. We performed some scalability tests to optimize the distribution of processors to the different components of the model so that the waiting time of each component is minimized. Results are provided in Figure1. The simulation has only been initiated so far, not allowing so far any scientific analysis yet. We provide here a snapshot of vorticity in the ocean after 1 years of simulation in Figure 2.

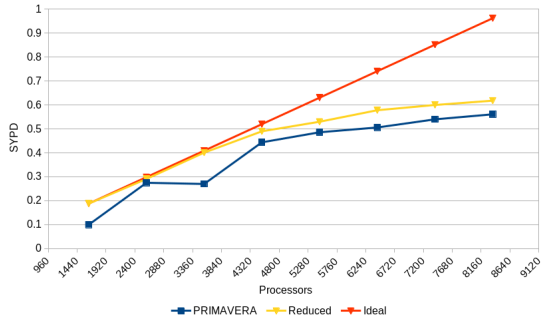



Figure 1 – Scalability test for the Glob-12km. Simulated Years Per Day as a function of the n. of cores used for two different set of outputs

## Conclusions

The initiation of this collaboration between IPSL/ENSTA and BSC has proven to be extremely valuable in terms of expertise exchange on supercomputing resources. The bottlenecks identified during the setup of the ultra-high resolution of the EC-Earth GCM will be extremely valuable information for future development of similar configurations of the RCM in the Mediterranean area at IPSL/ENSTA.

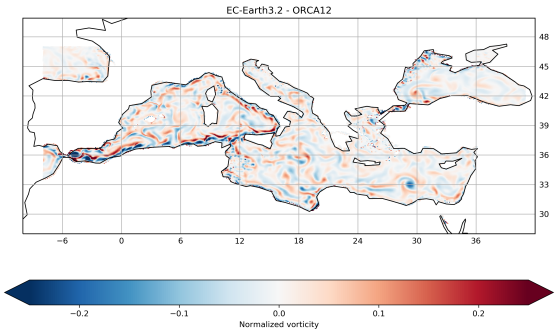



Figure 2 – Oceanic surface potential vorticity normalized by planetary vorticity

As the simulation Glob-12km will be continued during the coming months (some ~150 years of simulation scheduled for the overall project), we will start the comparison with available regional simulations of the Mediterranean region, performed within the context of the MED-CORDEX program [2], as well as with simulations using the same setup but at lower resolution, to identify the added value of this increase in resolution.

## References

[1] http://www.ec-earth.org/; [2] Ruti PM et al, MED-CORDEX initiative for Mediterranean climate studies. Bulletin of the American Meteorological Society 97.7 (2016): 1187-1208.

## Acknowledgements

The work has been performed under the Project HPC-EUROPA3, with the support of the EC Research Innovation Action under the H2020 Programme; in particular, the author gratefully acknowledges the support of the Earth Science Group and the computer resources and technical support provided by BSC.

# INTERNAL REMESHING & REPARTITIONING IN ELMER/ICE

*J. Todd*[1], *T. Zwinger*[2], *P. Raback*[2]

[1]*University of St Andrews, St Andrews, United Kingdom;* [2]*CSC IT Centre for Science, Espoo, Finland*

## Problem Statement

The current implementation of 3D iceberg calving in Elmer/Ice is limited by the requirement for external remeshing using GMSH whenever calving events change the geometry of the domain. Additionally, the external remeshing is limited to 2D, such that the 3D model must first be turned into a representative footprint, written to disk, meshed using GMSH, reread, extruded and deformed to match potentially non-vertical regions of the calving terminus. For reasons of model flexibility, robustness and efficiency, it is desirable to be able to remesh arbitrarily in 3D, with no dependence on vertically structured meshes.

## Aims & Objectives

To address the issues raised above, we aim to interface Elmer with external libraries to handle 3D remeshing and repartitioning. The latter is important to ensure load balance for large parallel simulations. We have the following objectives:

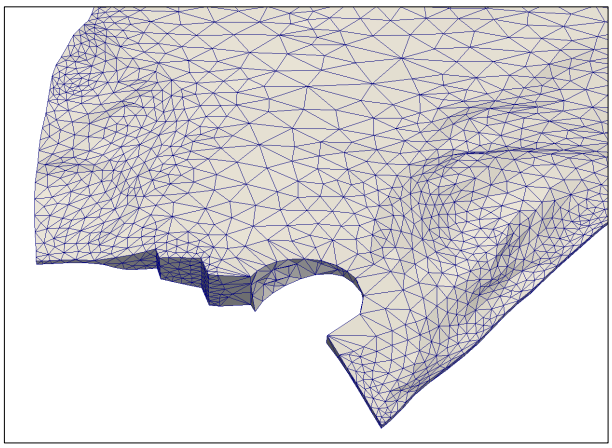
- Interface Elmer with MMG and make use of MMG's 3D remeshing/adaptation routines
- Interface with PT-Scotch, ParMETIS or another parallel partitioning library to allow dynamic repartitioning as the mesh changes through a running simulation.
- Write reusable routines to handle the redistribution of mesh parts as part of repartitioning
- Test these new tools on a calving test case

## Progress Report

We have largely achieved our stated objectives, thanks to considerable help from Elmer researchers at CSC. For now, the new code developed by this project is in the elmerice_meshadapt branch of the Elmer FEM github (https://github.com/elmercsc/elmerfem/) but this will be merged into the elmerice branch once development is completed. It is now possible to adapt

Terminus of Kangerlussuaq Glacier with idealised calving event removed during remeshing.

3D elmer meshes, either adapting mesh resolution (e.g. local mesh refinement) or modifying geometry of the domain (e.g. calving events). The remeshing occurs in serial, but the associated bottleneck can be minimized by constraining the remeshing to the part of the domain where adaptation is required. Load balance can also be improved during simulations using Zoltan, which is a parallel repartitioning library with similar functionality to PT-Scotch and ParMETIS. Importantly, Zoltan is capable of *repartitioning,* which differs from standard partitioning in that it adapts an *existing* partitioning, so that it minimizes the required MPI communication while performing the load balancing.



We have finished the development of the core functionality, and remaining work is restricted to implementation. For example, it remains to develop a 3D calving predictor which will feed a level set function into the MMG3D interface. Some improvements will also be made to Elmer's interpolation routines to improve efficiency in cases where almost all nodes are fixed in space.

## Acknowledgements

The work has been performed under the Project HPC-EUROPA3 (INFRAIA-2016-1-730897), with the support of the EC Research Innovation Action under the H2020 Programme; in particular, the author gratefully acknowledges the support of Thomas Zwinger & Peter Raback of CSC, Finland and the computer resources and technical support provided by CSC.

# NUMERICAL SIMULATIONS OF RADAR LOGGING AND ITS PROPAGATION CHARACTERISTICS

*F. Zhou*

*Delft University of Technology, Delft, The Netherlands*

## Introduction

In the process of oil drilling, mud filtrate penetrates rock pores and alters the fluid properties in the formation, hereby affecting formation evaluation conducted with conventional well logging tools. We think the logging responses on the mud-contaminated formation also carry some useful information. For example, mud invasion depth has a close correlation with the permeability and porosity. We propose to use a borehole radar tool to detect the mud invasion depth. Once the invasion depth is decided, a quantitative relation is possible to be established between radar responses and reservoir permeability.

In this numerical work, we established a numerical model of radar logging in an open hole and invaded reservoir environment, and investigated the propagation characteristics of EM waves.

## Methods

The investigation was conducted in a numerical simulation way. An electromagnetic (EM) model was established to simulate the responses of borehole radar on the mud disturbed formation. The fluid model has been developed by the guest researcher where the multiphase and multicomponent flow formulas are numerically solved with an implicit pressure and explicit saturation (IMPES) method. The EM model was built up using gprMax. gprMax is a free ground penetrating radar data simulator based on finite difference time domain (FDTD) method, and the code has been developed by the host researcher.

## Results

Figures 1a and 1b present the built associated numerical model, and the corresponding EM response by time-lapse radar measurements.

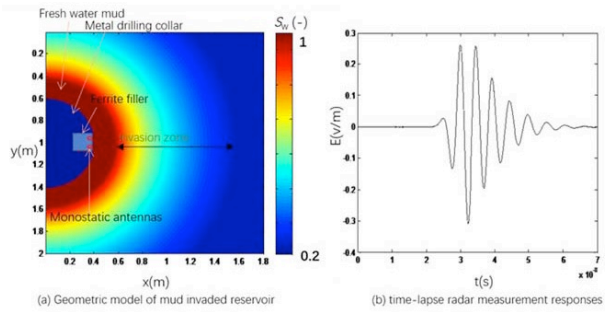



Figure 1 – The established mud invaded model with borehole radar (a) and the corresponding EM responses (b).

Figures 2a–d present the simulated EM snapshots in different times for a water-based mud invasion process. Note that the mud invasion model is simplified to displace the gradual fluid transition zone with a step model.

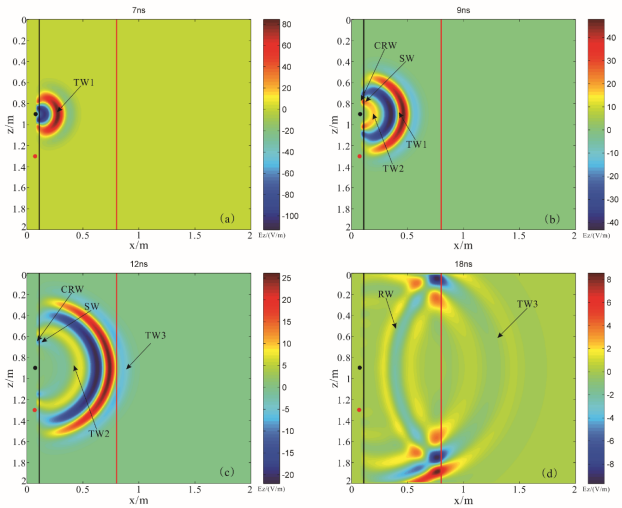



Figure 2 – Snapshots of the vertical component of the electric field in the simulation of water-base mud （The black dot represents the position of the transmitting antenna. The red dot represents the position of the receiving antenna, and the black solid line indicates the location of the borehole. The red solid line represents the interface between the invaded zone and the virgin formation).

## Conclusions

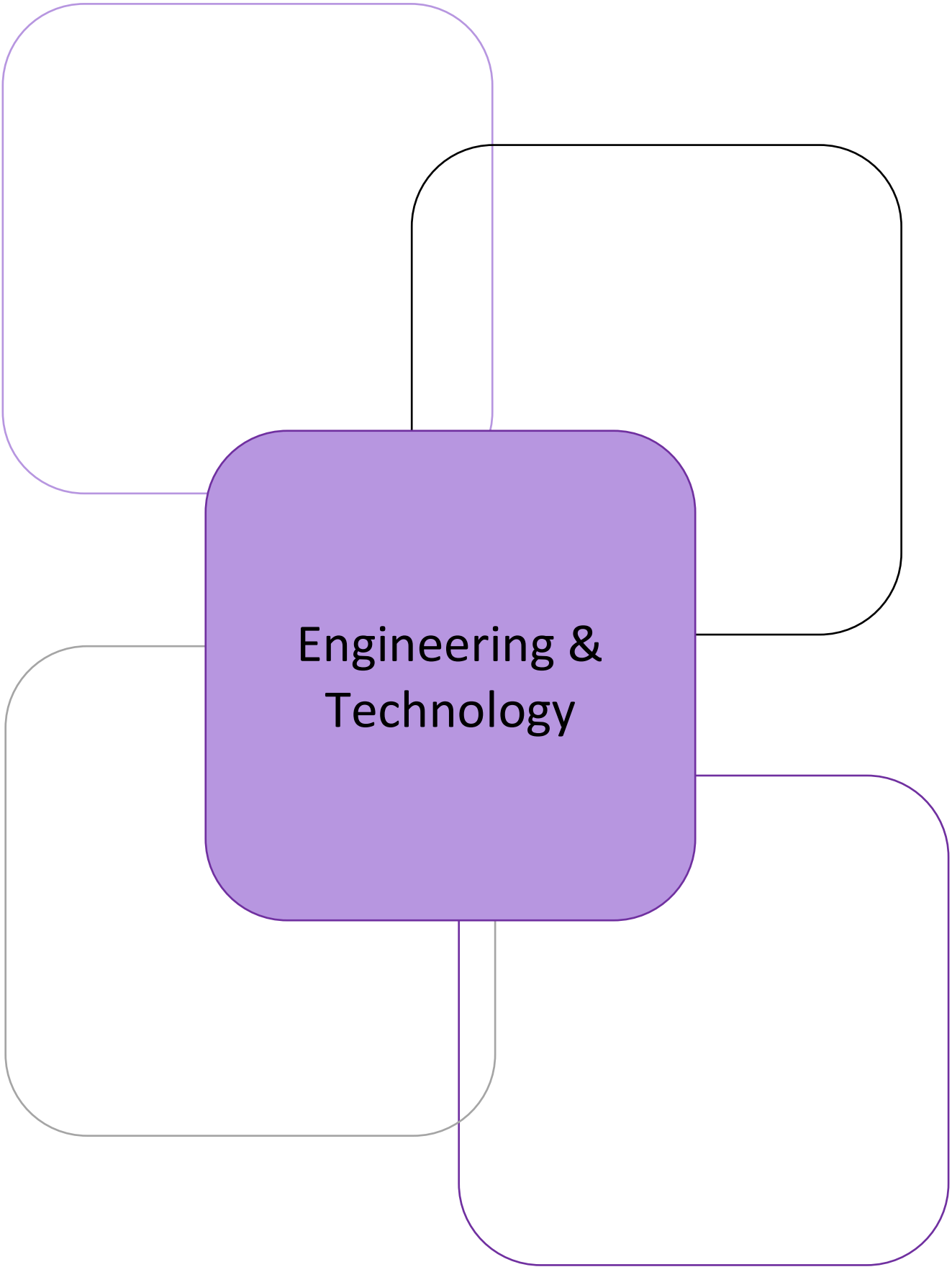
The borehole radar has a visible EM response on the mud-invaded reservoir, and the received EM signals have a correspondence with the gradual transition zone of mud invasion. In a water-based mud invasion environment, the EM waves propagating in the reservoir present a characteristic of alternating light and dark, complicating the useful signal extraction.

## References

[1] Heigl WM et al, Can we obtain invasion depth with directional borehole radar[J]? Petrophysics, 46(1): 52-61, 2005; [2] Slob E et al, Surface and borehole ground-penetrating-radar developments[J], Geophysics, 75(5): A103-A120, 2010; [3] Warren C et al, gprMax: Open source software to simulate electromagnetic wave propagation for Ground Penetrating Radar, Computer Physics Communications, 209, 163-170, 2016.

## Acknowledgements

The work has been performed under the Project HPC-EUROPA3 (INFRAIA-2016-1-730897), with the support of the EC Research Innovation Action under the H2020 Programme; in particular, the author gratefully acknowledges the support of Dr. Antonis Giannopoulos at The University of Edinburgh and the computer resources and technical support provided by EPCC.


Engineering &
Technology

# FRAMEWORK FOR CPACS

*M. Drougard*

*EPFL, Lausanne, Switzerland*

## Introduction

In the conceptual and primarily stages of an aircraft design, the designers need to take into account all the major parameters and requirements of the aircraft. They try to find out the best possible configurations for the aircraft. Such parameters as the wing size, fuel weight, take-off run will be determined. During these phases the aircraft and requirements will be often redefine and new analyses need to be performed. Multiples disciplines and teams as aeronautics, engine engineering, structures analyses are involved in this process. All these disciplines need to exchange data and to modify the aircraft in an efficient way. CPACS was developed to be a common format for these disciplines, but do not have high level parameters for the geometry of the aircraft as the wing span or the wing sweep. CPACS is also not fully integrated in the CFD simulation work-flow. Some manual steps are required to run a CFD simulation on a CPACS file. This paper will present a new way to extract and modify the sweep and dihedral parameters of a wing and present a work-flow to automatically run CFD simulations on the Cray XC40 system called Beskow [1] at KTH-PDC center which is the fastest super computing system in Scandinavia.

## Methods

CPACS geometric format is quite nested and has hierarchical dependences between its elements. CPACS format, also, make a heavy use of rotations, translations and scalings trans-formations. So, we first develop a library to represent the positions and the transformations using homogeneous coordinates and augmented matrices. With this library we can easily extract the sweep angle, define as the angle between the Y axis and the leading edge of the tip. To modify the sweep angle, we first compute the needed tip position to obtain the desired sweep angle. Then we computed the positions for the other sections of the wing. Finally, we encode these desired positions back in CPACS format using decomposition of matrices and linear algebra. We use the same kind of mechanisms to modify and extract the high-level dihedral parameter.

The second part of the project was to enable an aircraft CFD simulation based on CPACS file. To do this, we first need a mesh that represent the aircraft and then perform the simulation on it. Fortunately, there are already some software that perform the meshing and the simulation. The meshing can be created by SUMO and the simulation performed by SU2. But SUMO does not accept CPACS files as input. However, a software, called CPACS2SUMO [2] converts CPACS file to SUMO SMX file. So, we create scripts to connect these pieces of software together and to easily run multiple simulation using different simulation parameters. We also find out the best core number for SU2 simulation, see Figure 1.

## Results

A new software with a GUI to modify the aircraft geometry is on the way and will be release in August 2018. The software is based on TIGLViewer and use high level parameters. The scripts to prepare and launch CFD simulations on the supercomputer Beskow can be found on my Github [3]. For example, we compute the effect of the sweep angle on the pitching moment coefficient (CMy), see Figure 2.

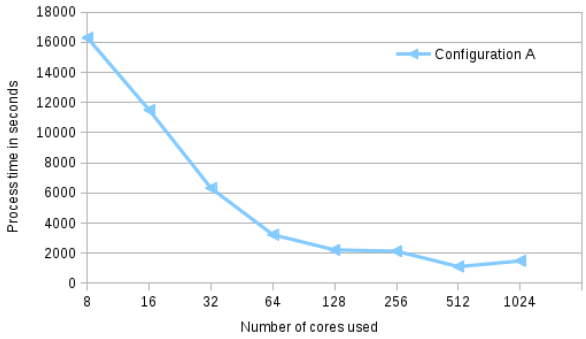



Figure 1 - Core numbers scalability for SU2 simulation

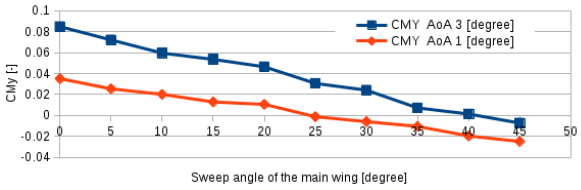



Figure 2 - Effect of variation of the sweep angle on the CMy

## Conclusions

This new framework brings really new perspectives of the utilization of CPACS format. The aircraft designer can now easily modify some high-level parameters and observe the impact using CFD simulation in the early change of the design process. This framework also opens the door to self-optimization loop using CFD simulations and high-level parameters.

## References

[1] PDC KTH. Web site PDC KTH. https://www.pdc.kth.se/hpc-services/computing-systems; [2] Aidan Jungo. CPACS2SUMO. https:https://github.com/aidan-cfse/CPACS2SUMO; [3] Drougard M, cpacsOnBeskow. https://github.com/ MaloDrougard/cpacsOnBeskow; [4] Airinnova. Airinnova web site. http://airinnova.se/.

## Acknowledgements

The work has been performed under the Project HPC-EUROPA3 (INFRAIA-2016-1-730897), with the support of the EC Research Innovation Action under the H2020 Programme; in particular, the author gratefully acknowledges the support of Airinnova [4] and the computer resources and technical support provided by PDC center of the KTH.

# PHYSICAL ORIGIN OF FINE PARTICLE PROBLEM IN BRITTLE DYNAMIC FRAGMENTATION

*A. Iravani[1], J. Åström[2]*

*[1]Department of Mineral Resources Engineering, Montanuniversitaet Leoben, Leoben, Austria; [2]CSC IT Center for Science, Espoo, Finland.*

## Introduction

Blasting with explosives and crushing with mills are two major processes for extracting ore minerals. Long-standing problems with these processes are "fines" production in blasting and the related energy consumption of mills. The energy consumption problem is, at least partly, a consequence of fines produced in mills. The creation of fines requires a lot of energy as the total created fracture surface grows large. Consequently, the two problems can be condensed into a "fines" problem [1]. A better understanding of the origin of fines could help eliminate these problems at an initial stage and improve the blasting and crushing practices.

## Methods

We employ a discrete-element code (HiDEM) to simulate blasting experiments of magnetic mortar cylinders. The experiments were conducted by Johansson [2] as a set of laboratories blasting fragmentations of magnetic mortar cylinders. In the HiDEM code, we model the cylinder with material properties similar to magnetic mortar, and use a time-dependent pressure pulse that crudely replicates the blasting process [3].

The blasting experiments create a compressive shock wave traveling outward in the radial direction. As a cylinder expands outward, it will induce tension in the tangential direction. In crushing, fragments are broken by continual shear deformation [4]. Such a process has a power-law behaviour. The tensile cracks can easily become unstable, branch, and further merge, forming fragments. This process is inherently universal and leads to a characteristic fragment size distribution (FSD) [5]. If n(s) describes the number-density of fragments with s number of grains, the FSD, or the number of fragments in a size-interval ds, can then be written with crushing, branching-merging and boulder components as [1]:

$$n(s)ds = C_1 s^{-ß} ds + C_2 s^{-\alpha} \exp(-s/C_3) ds + C_4 \exp(-s/s_b) ds$$

where $C_1$, $C_2$, $C_3$ and $C_4$ are non-universal constants, ß defines degree of grinding, $\alpha = (2D-1)/D$ with D being the dimension and $s_b$ characteristic size of the boulder. With a proper transformation from s to r and integrating the n(s) the mass passing fraction MPF(r) can be approximated. See [1] for detailed description.

## Results

Figure 1 displays the mass fractions for the three categories as functions of $P_{peak}$ and an experiment. The numerical data shows a significantly larger mass fraction for crushing. The origin of this is that the mortar grains are typically of size $r_{grain} \sim 0.1$ mm, while in the simulations $r_{grain} \sim 3$ mm is used. It is obvious that the numerical and experimental data match each other very well, apart from the expected deviations due to differences in $r_{grain}$, which also explain the slight difference in the branching-merging mass fraction for the highest $P_{peak}$, at which the crushing mass fraction grows large enough to begin depleting the branching-merging fraction.

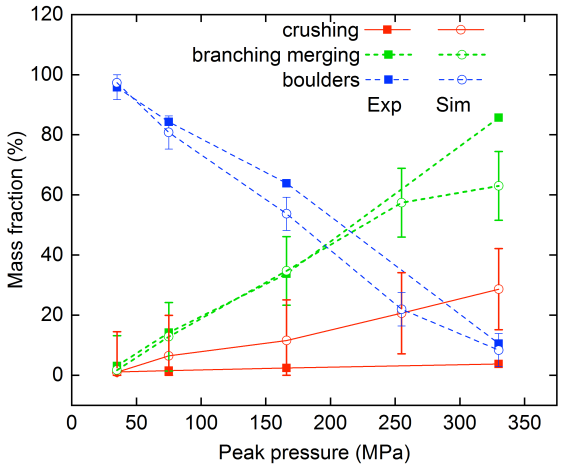



Figure 1 - The crushing, branching-merging, and boulder mass fractions as functions of $P_{peak}$. The estimated errors of the numerical results compared with the experimental data are shown by error bars.

## Conclusions

The blasting-fragmentation FSDs are composed of three parts, all of which can be described separately by universal fragmentation mechanisms. The fines arise mostly as a result of crushing. Fragments larger than fines are formed as a result of tensile-crack propagation. Tensile cracks branch and merge to form smaller fragments, and the tensile cracks set forth the largest fragments, called boulders. The results indicate two possible strategies for addressing the fines problem. (1) The blasting process should be designed with a minimum charge to reach cleavage threshold, and with a suitable geometrical proximity of applied charges to limit the resulting boulder size. (2) The blasting itself should be designed so that it minimizes compressive shear and maximizes tensile load.

## References

[1] Iravani A et al, Phys. Rev. Applied, 10(3):034001, 2018; [2] Johansson D, Lic. thesis, LUT, 2008; [3] Iravani A et al, in Proceedings of 12th international Symposium for rock fragmentation by blasting, p. 597, 2018; [4] Åström JA et al, Eur. Phys. J. E 35(5):40, 2012; [5] Åström JA, Adv. Phys. 55(3-4):247, 2006.

## Acknowledgements

The work has been performed under the Project HPC-EUROPA3 (INFRAIA-2016-1-730897), with the support of the EC Research Innovation Action under the H2020 Programme; in particular, the author gratefully acknowledges the support and the computer resources and technical support provided by CSC-IT Center for Science.

# STRUCTURE AND DYNAMICS OF NAFION/GRAPHENE OXIDE NANOCOMPOSITES

*G. Kritikos[1], K. Karatasos[1], A. Lyulin[2]*

*[1]Department of Chemical Engineering, Aristotle University of Thessaloniki, Thessaloniki, Greece; [2]Department of Applied Physics, Technische Universiteit Eindhoven, Eindhoven, The Netherlands*

## Introduction

The main challenge in Nafion membranes is to improve conductivity while on the same time to enhance mechanical stability. Cation conductivity increases by increasing water uptake [1]. Insertion of hydrophilic graphene oxide (GO) in a nafion membrane can serve this goal. Graphene sheets can affect the mechanical properties while, depending on other characteristics such as size and side groups, can also affect solvent diffusion [2]. The formation of a structured Nafion layer on the nanosheet can modify the formed nanochannels and even promote semi-crystallization.

Atomistic Molecular Dynamic simulations can enlighten us about the interplay between the two hydrophilic centres Nafion and GO [1,2]. Long simulations may allow a clarification in the atomistic level of the thermodynamic and dynamic properties observed in such materials.

## Methods

Each nanocomposite system contained one GO of 8x8 $nm^2$ lateral dimensions and 100 Nafion chains of 10 monomers each (referred to as NAFGO). We have focused on three temperatures (350, 300 and 250 K) and three hydration levels (λ) equal to 10, 15 and 20. At this hydration levels hydronium cations are present and they were taken into account one per side chain. For the nanocomposite systems the number of atoms were 105804, 120804 and 135804 for λ equal to 10, 15 and 20 respectively. As reference, we have taken into account bulk systems (in the absence of GO) of 15 Nafion chains containing the respective hydronium and water molecules (referred to as NAF).

The initial NAFGO and NAF systems were first equilibrated for 50 ns at 1000 K. Then we have cooled each system to 600 K in 50ns where Nafion is still mobile. Below that temperature we have followed a cooling procedure of 50 K/30 ns to the target temperature of 350 K. At 350 K we have hydrated the systems in order to reach the desired hydration levels. We have used the Gromacs MD software package.

## Results

Comparison of the mean square displacement of the bulk NAF system of 15 Nafion chains with another bulk of 100 Nafion chains showed insignificant size effects. In figure 1 we present the distribution of the centres of mass as a function of the distance from the GO surface. We observe that Nafion assumes zero density close to GO, thus forming channels which can promote ionic diffusion at the examined temperature. It is also found that GO affects the structure of the Nafion chains, promoting an ordering parallel to the sheet. The observed channel regions have a width that is comparable to the size of the Nafion side chain. Moreover, the GO exhibits more favourable interactions with the hydronium ions than with the more mobile water molecules.

Investigation of the dynamics close to the GO surface reveals a slower diffusion compared to the bulk for both hydronium and water due the interaction with the nanosheet. Far from the GO plane bulk dynamics are restored.

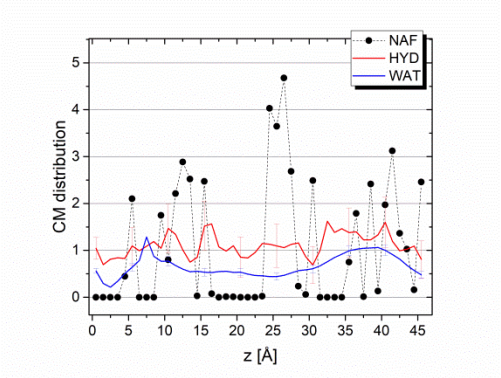



Figure 1 - Distribution of the centres of mass at T=350K (Nafion:NAF, Water:Wat and Hydronium:HYD).

## Conclusions

Graphene Oxide affects the structure of Nafion imposing the formation of parallel channels. Although GO is hydrophilic the concertation of hydronium on the sheet is higher than that of water. Future plans include the study of the characteristics of the diffusion in the formed is channels.

## References

[1] Sengupta S et al, Int. J. Hydrogen Energy 42, 27254 (2017); [2] Kritikos G et al, Mater. Today Commun. 13, 359 (2017).

## Acknowledgements

The work has been performed under the Project HPC-EUROPA3 (INFRAIA-2016-1-730897), with the support of the EC Research Innovation Action under the H2020 Programme; in particular, the author gratefully acknowledges the support of Prof. Lyulin A. host lab and the computer resources and technical support provided by the HPC Access Centre SARA.

# USING MESH MULTIPLICATION TECHNIQUES WITH CODE_SATURNE FOR LARGE SYNTHETIC JET SIMULATIONS

*A. Miró[1], C. Moulinec[2], Y. Fournier[3]*

*[1]Universitat Politècnica de Catalunya, Barcelona, Spain; [2]STFC Daresbury Laboratory, Daresbury, United Kingdom; [3]EDF R&D, Paris, France.*

## Introduction

Generating extremely large meshes for CFD is the key to make the best use of petascale/exascale machines. The best approach to generate such large meshes is to use an on-the-fly mesh multiplication (MM) algorithm. This results in a high-quality mesh since skewness and mesh quality can be controlled for the original mesh.

A mesh multiplication algorithm has been implemented in the open-source CFD Code_Saturne code [1]. The current algorithm is an improvement over an already existing version [2] and is able to handle different kinds of elements, e.g. hexahedra, tetrahedra, prisms, pyramids and generic polyhedra and also meshes made of different types of elements. Moreover, the grounds of a feature detection and curvature interpolation algorithm have been derived.

Finally, the MM algorithm has been successfully used in a relevant engineering study: the impingement of slotted and axisymmetric synthetic jets [3].

## Methods

The fully parallel mesh multiplication algorithm is divided into three steps: edges, faces and cells generation. A generic edge division is implemented by computing the edge mid-point. New vertices for the refined elements are built using these edge mid-points. Some elements, such as hexahedra, need a mid-face point and a mid-cell point. They are added by computing the face and cell center of gravity respectively. Then boundary and interior faces are built based on the element type and using the mid-face point if needed. Finally, the cells are sub-divided, also focusing on the element type, and the refined mesh connectivity is built. At this point, a usable mesh is obtained.

The quality of the refined mesh is further improved by detecting any curved edges and by providing a better approximation of the said curve in the refined mesh. This algorithm is divided into two steps: first, features in the unrefined mesh are detected and the faces/vertices belonging to a curve are flagged.

The aforementioned algorithm is applied in order to generate large meshes for time-accurate simulations of circular synthetic jet actuators. The coarse computational meshes are built using a serial meshing tool and are up to 1 million cells. Two refinements have been considered using the MM algorithm without the curvature interpolation.

## Results

The MM algorithm is also successfully applied to refine several types of meshes. Single tetrahedral and hexahedral meshes are used for validation. The MM is validated with a number of test cases such as paved meshes and prismatic triangular meshes. The new MM algorithm is shown to successfully handle these meshes, with respect to the previous implementation. Then, MM is applied to a series of community relevant cases, such as an airfoil (see Figure 1).

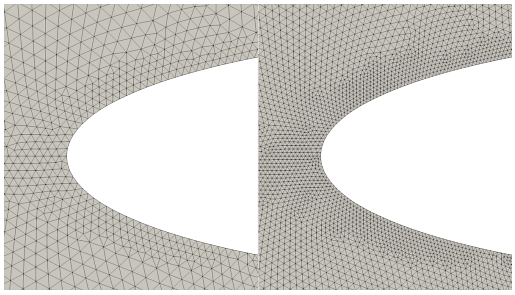


Figure 1 – Mesh multiplication algorithm applied to an airfoil (leading edge). Unrefined case, left; refined case, right.

Time-accurate large eddy simulations of circular SJA have been performed with good agreement with the existing literature and the results are published in [3].

## Conclusions

The division of element faces and generation of the correct number of nodes and cells have been successfully implemented into the MM algorithm for Code_Saturne. The algorithm is able to generate large meshes in parallel by splitting the original elements using Code_Saturne's mesh processing capabilities. Moreover, it can be the ground to build an Adaptive Mesh Refinement.

The MM algorithm is successfully applied to a relevant engineering study, having proved to be able to generate large meshes to study the impingement of circular SJA. An 8 million cell and a 16 million cell meshes have been produced using this technique.

## References

[1] Archambeau F et al, Int. J. of Fin. Vol. 1(1), 2004; [2] Kabelikova P et al, PRACE project (FP7/2007-2013):1–8, 2011. [3] Miró A et al, ICCFD10:1-18, 2018.

## Acknowledgements

The work has been performed under the Project HPC-EUROPA3 (INFRAIA-2016-1-730897), with the support of the EC Research Innovation Action under the H2020 Programme; in particular, the author gratefully acknowledges the support of STFC Daresbury Laboratory Scientific Computing Department and the computer resources and technical support provided by EPCC.

# FINE-GRAINED APPLICATION TUNING ON OPENPOWER HPC SYSTEMS

*L. Riha[1], O. Vysocky[1], A. Bartolini[2]*

*[1]IT4Innovations national supercomputing center VSB – Technical University of Ostrava, Ostrava, Czech Republic; [2]University of Bologna, DEIS, Bologna, Italy.*

## Introduction

This project is building on our experience achieved during the investigation of the Horizon 2020 project READEX. The READEX (Runtime Exploitation of Application Dynamism for Energy-efficient eXascale computing) [1,2] project analyzes HPC applications to find the best settings of the tuning parameters to minimize energy consumption. In this project our goal was to evaluate the READEX methodology on new hardware platform, in this case POWER8 based system D.A.V.I.D.E installed in CINECA and its energy and power monitoring system.

## Methods

The READEX project expects that HPC applications have different needs in separate regions of the code. These needs translate into energy savings because different regions require different optimal settings of the hardware parameters. In this study we tune:

- CPU Core (CF) and Uncore (UCF) frequencies
- number of OpenMP threads, thread placement

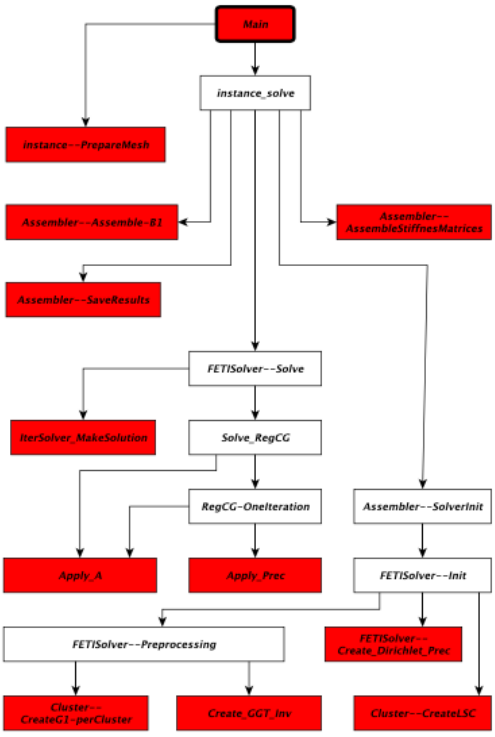



Figure 1 - Diagram of significant regions of the ESPRESO FEM library. The red regions are used for dynamic tuning.

User applications can be instrumented using manual instrumentation to analyze each region of the code separately. Figure 1 shows the instrumentation of the ESPRESO FEM library developed at IT4Innovations that is used for this work and has been ported to D.A.V.I.D.E. The red regions are the ones that are dynamically tuned because we do not tune the nested regions.

## Results

The final results are shown in the Table 1. We can see that up to 27.2% of energy can be saved against the default settings at a penalty of 15.3% in increase of the optimal, not default, runtime. In addition, one can tune for optimal runtime and save 23.6% of the runtime at penalty of 10.1% of extra energy consumption.

| | Total runtime and energy | Savings compared to the default settings |
|---|---|---|
| Default config. | 110.3 s; 76.7 kJ | - |
| the best time configuration | 84.2 s; 62.1 kJ | **23.6%** time savings extra **10.1%** of energy |
| the best energy configuration | 99.4 s; 55.8 kJ | **27.2%** energy savings extra **15.3 %** of runtime |

Table 1 – Final results of the dynamic tuning of the ESPRESO library on D.A.V.I.D.E POWER8 system

## Conclusions

POWER8 is an architecture where correct hyper-threading settings is a key to obtain optimal runtime and energy consumption. In addition, it allows for very fine-grained tuning of the CPU core frequency. Combination of these two aspects generates very significant potential for dynamic tuning that can be translated into good time and energy savings as shown in Table 1.

## References

[1] Schuchart J et al, The READEX formalism for automatic tuning for energy efficiency, 2017, Computing, 99 (8), pp. 727-745., DOI: 10.1007/s00607-016-0532-7; [2] Vysocký O et al, Evaluation of the HPC Applications Dynamic Behavior in Terms of Energy Consumption, Proceedings of the Fifth International Conference on Parallel, Distributed, Grid and Cloud Computing for Engineering (PARENG), 2017. doi:10.4203/ccp.111.3.

## Acknowledgements

The work has been performed under the Project HPC-EUROPA3 (INFRAIA-2016-1-730897), with the support of the EC Research Innovation Action under the H2020 Programme; in particular, the author gratefully acknowledges the support of University of Bologna, DEIS and the computer resources and technical support provided by CINECA.

# MULTI-SCALE MODELLING OF FIBRE REINFORCED CONCRETE MATERIALS

*J. Sliseris*

*Riga Technical University, Department of Structural Mechanics, Riga, Latvia.*

## Introduction

High performance fibre reinforced concrete (HPFRC) is becoming more popular in structural engineering. However, damage mechanics of HPFRC is not well understood, especially when fibres are oriented and material is additionally strengthened with traditional reinforcement bars (re-bars). Additional factors that affect properties of fibre concrete are fibre shape, aspect ratio, volume fraction and properties of concrete itself [1–2].

## Methods

Numerical modelling of damage and fracture of high-performance fibre concrete with re-bars is computationally intensive task. Pre-processing part includes a generation of discrete element model that is based on tetrahedron mesh (see Figure 1). The discrete element model consists of elements that represent concrete, re-bars and joint elements between re-bars and concrete. Re-bar elements do not coincident with concrete elements. Therefore a special Python routine is developed that calculates the location of joint elements that is consistent with pattern of re-bars. The Python code was managed to run in parallel on the KTH cluster. The most time consuming part is generation of joint elements between re-bars and concrete. Meanwhile, the cross section area and constitutive model parameters are obtained using homogenization of tetrahedron for 12 independent load cases that includes time consuming matrix operations. In order to model complex geometries with fine resolution, this procedure should be updated using parallel programming approach.

## Results

A significant progress in the development of numerical modelling framework of modelling fracture and damage of high-performance fibre concrete with re-bars was done. A fully automated procedure for the generation of buildings coarse and fine scale geometry and finite element model was created (see Figure 2). The generation of geometry was done within Salome-Meca framework. With assistance of KTH PDC personal it was possible to run this code on the KTH cluster. By using the cluster it was possible to simulate the whole building with relatively fine mesh that is not possible on personal computer.

## Conclusions

A new numerical modelling framework for high performance fibre reinforced concrete buildings are created. The modelling technology is fully scalable on simulation clusters. This method provide a possibility to model full building by taking into account material non-linearity and high finite element resolution.

The model will developed in further research to make it possible for practical use in the field of structural engineering. Additional routines for better parallelization and automation for two scale simulation will be developed in further research.

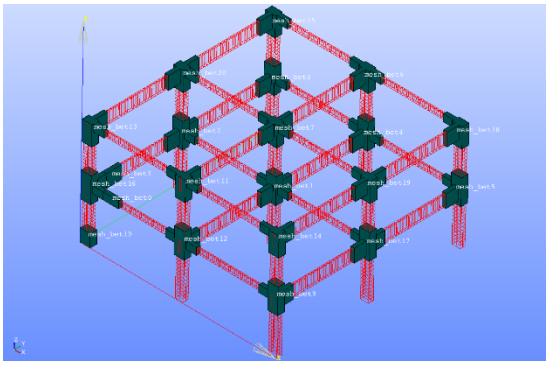


Figure 1 - Two story fibre reinforced concrete building.

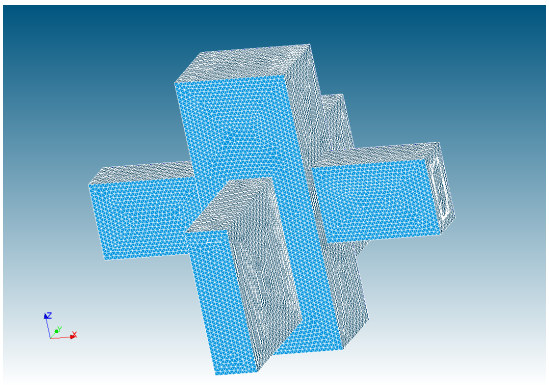


Figure 2 - Fine scale model for the main structural joint.

## References

[1] Sliseris J, Numerical analysis of reinforced concrete structures with oriented steel fibers and re-bars, Engineering Fracture Mechanics, Volume 194, 1 May, Pages 337-349, 2018; [2] Sliseris J, et al., Numerical modelling of flax short fibre reinforced and flax fibre fabric reinforced polymer composites, Composites Part B: Engineering, Volume 89, Pages 143-154, 2016.

## Acknowledgements

The work has been performed under the Project HPC-EUROPA3 (INFRAIA-2016-1-730897), with the support of the EC Research Innovation Action under the H2020 Programme; in particular, the author gratefully acknowledges the support of Department of Solid Mechanics and the computer resources and technical support provided by KTH HPC centre.

# DYNAMIC MODE DECOMPOSITION OF THE COMMON RESEARCH MODEL WAKE AT SUBSONIC STALL

*A. Waldmann*[1], *O. Lehmkuhl*[2]

[1]*Institute of Aerodynamics and Gas Dynamics, University of Stuttgart, Stuttgart, Germany;* [2]*Barcelona Supercomputing Center (BSC-CNS), Barcelona, Spain*

## Introduction

The aerodynamics of civil transport aircraft in flight are characterized by high Reynolds numbers due to the combination of high speed and large physical size. At the edges of the flight envelope, i.e. at significantly increased angles of attack, massive flow separation may occur. This is associated with large-scale turbulent fluctuations in the wake flow and possibly significant unsteady forces acting on the tailplane. Figure 1 shows an instantaneous snapshot of the Common Research Model (CRM) aircraft configuration at subsonic stall conditions, highlighting the separated wake using isosurfaces of $\lambda_2$.

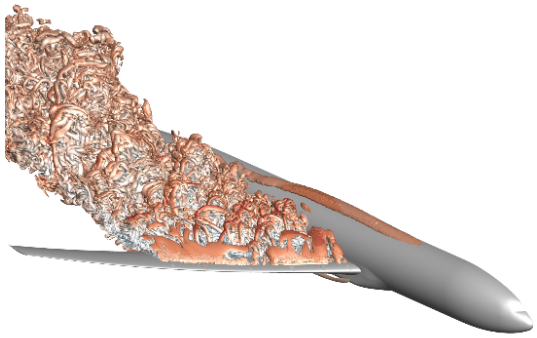


Figure 1 – Common Research Model wake at $M_\infty = 0.25$, $Re = 11.6 \cdot 10^6$, $\alpha = 18°$

It is therefore imperative for the prediction of such unsteady phenomena to understand the dynamics of such wake flows. Hybrid methods such as Detached Eddy Simulation (DES) are capable of resolving flows of this nature. However, the size of the resulting data set with a significant temporal and spatial resolution can be cumbersome to handle. While temporal statistics can provide significant insight into the behaviour of such wakes, research into modal decomposition methods has recently made possible the effective use of more advanced spatiotemporal analysis methods.

## Methods

Modal decomposition methods such as Proper Orthogonal Decomposition (POD) or Dynamic Mode Decomposition (DMD) have risen in popularity in the recent decades. While POD is useful for extracting the most energetic features, there is some interest in more reliable information on the frequency distribution of the fluctuation modes. This is provided by DMD, which extracts spatial modes from the flow field and assigns frequencies at which they occur.

Thorough derivations of DMD can be found in various publications, such as [1]. The work carried out in the present context focused on applying DMD on large-scale data sets from simulations using the same aircraft and flight conditions and two different solution types: Large Eddy Simulation results obtained using BSC's ALYA flow solver and DES results from the DLR TAU solver used at IAG. The unsteady results are fed into the DMD method implemented in Python. Most of the computational effort is spent in linear algebra routines on matrices whose size depends on the input data set.

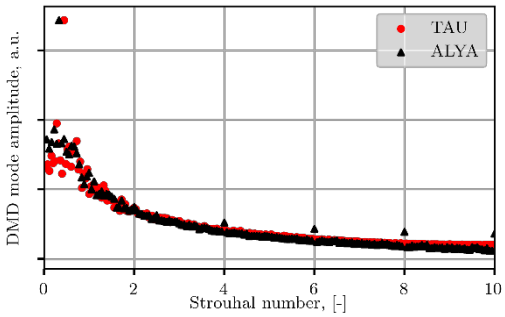



Figure 2 – CRM normalized dynamic mode spectrum at in the wake plane at 28.3% of semispan

Figure 2 shows the spectra obtained via DMD using time-resolved series of planar slices through wake flow of the CRM. Remarkably, while the codes employ different computational approaches, the dynamics of the CRM's wake flow field at this position show a significant degree of agreement between the two solutions. Most of the fluctuation energy is concentrated at in the lower frequency range, represented by Strouhal numbers between 0 and 1. In both cases there is a distinct dominant peak, which can be observed at $Sr = 0.33$ in the ALYA solution and at $Sr = 0.4$ for TAU. This mode corresponds to a meandering motion similar to a von Kàrmàn vortex street.

## Conclusions

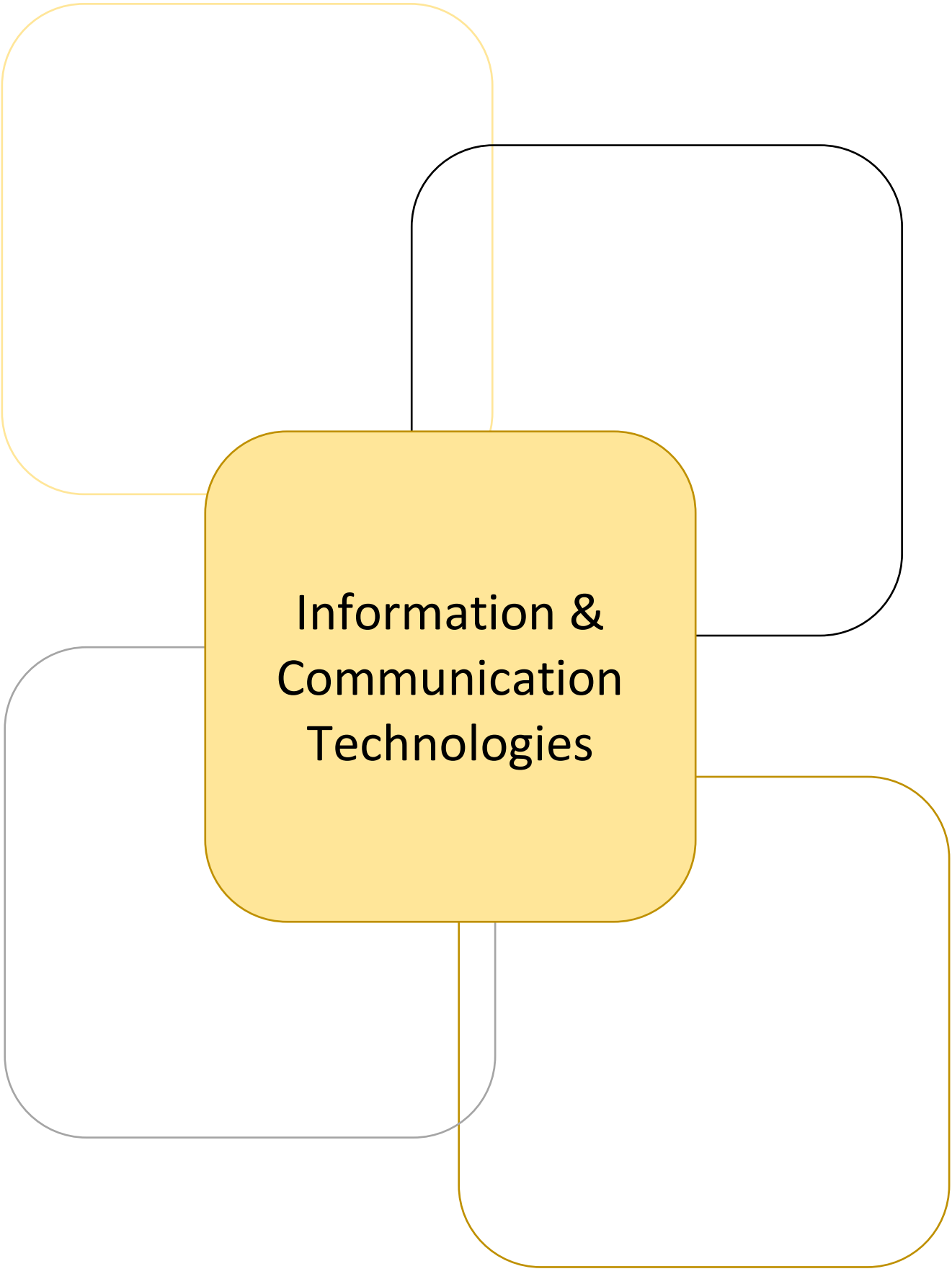
The DMD method was shown to be a viable method for the analysis of massively separated flows. Its straightforward implementation and scalability, combined with the availability of HPC resources, should enable future studies of the entire 3D flow field.

## References

[1] Schmid P, Dynamic mode decomposition of numerical and experimental data, J. Fluid Mech., 656:5-28, 2010.

## Acknowledgements

The work has been performed under the Project HPC-EUROPA3 (INFRAIA-2016-1-730897), with the support of the EC Research Innovation Action under the H2020 Programme; in particular, the author gratefully acknowledges the support of Oriol Lehmkuhl and the computer resources and technical support provided by BSC.


Information &
Communication
Technologies

# DETECTION OF OIL SPILLS IN THE SEA USING AERIAL SENSOR DATA AND DEEP LEARNING TECHNIQUES

*A.-J. Gallego[1], P. Gil[1], A. Pertusa[1], and R.B. Fisher[2]*

*[1]Computer Science Research Institute, University of Alicante, Alicante, Spain; [2]School of Informatics, University of Edinburgh, Edinburgh, United Kingdom.*

## Introduction

A quick response from governments is required in situations of marine pollution due to oil spills. When an oil slick is detected, the authorities activate the emergency protocols in order to control the environmental impact and the ecological damage in the sea.

In this work, we use deep neural autoencoders to segment oil spills from Side-Looking Airborne Radar (SLAR) imagery. Synthetic Aperture Radar (SAR) has been much exploited for ocean surface monitoring, especially for oil pollution detection, but few approaches in the literature use SLAR. Our sensor consists of two SAR antennas mounted on an aircraft, enabling a quicker response than satellite sensors for emergency services when an oil spill occurs.

## Methods

Based on the idea of denoising autoencoders, we use a type of segmentation autoencoder as proposed in [1] but specifically designed for oil spill detection. In this case, we aim to learn a codification that maintains only those input pixels that we select as relevant. This is achieved by modifying the training function so that the input is not mapped identically at the output. Instead, we train it with a ground truth of the input image pixels that we want to select. We will refer to this model as Selectional AutoEncoder (SelAE).

The topology of the network consists of a series of convolutional plus Max Pooling layers until reaching an intermediate layer in which the encoded representation of the input is attained. It then follows a series of convolutional plus upsampling layers that generates the output image with the same input size. In addition, we incorporate a series of residual connections as proposed in [2], from each encoding layer to its analogous decoding layer, which facilitates convergence and leads to better results.

In order to validate the effectiveness of the proposed method, we used a dataset containing 38 fight sequences supplied by the Spanish Maritime Safety and Rescue Agency (SASEMAR). For the ground truth, we used a binary mask delimiting the pixels corresponding to oil spills.

## Results

Experiments on TERMA radar were carried out to detect oil spills on Spanish coasts using deep selectional autoencoders and RED-nets. Different configurations of these networks were extensively evaluated using the HPC resources.

The computational resources from this machine were mainly exploited to parallelize the gridsearch process in order to explore several network configurations. Once the best configuration and parameter settings for each network were selected, we evaluated the results using different metrics. Moreover, we compared these results with three state-of-the-art methods for oil slick segmentation in SLAR images.

| Model | F1 | IoU |
|---|---|---|
| Graph-based | 48.28 ± 1.87 | 32.55 ± 0.16 |
| JSEG | 28.73 ± 0.46 | 16.50 ± 0.35 |
| SegSM | 55.78 ± 1.18 | 87.33 ± 0.51 |
| SelAE | 89.31 ± 0.93 | 92.14 ± 7.21 |
| SelAE (with residual connections) | **93.01 ± 0.85** | **100.00 ± 0.00** |

Table 1 - Evaluation results for the two architectures (with and without residual connections) using the chosen parameters after grid-search.

## Conclusions

The proposed network topology significantly outperformed previous approaches, correctly detecting 100% of the spills and obtaining an F1 score of 93.01% at the pixel level. The proposed network model performs accurately in SLAR imagery that has artefacts and noise caused by the aircraft manoeuvres, in different weather conditions and with the presence of look-alikes due to natural phenomena such as shoals of fish and seaweed.

## References

[1] Gallego A et al, Staf-line removal with selectional auto-encoders. Expert Systems with Applications, 89, 138–148, 2017; [2] Mao X et al, Image Restoration Using Very Deep Convolutional Encoder-Decoder Networks with Symmetric Skip Connections. NIPS, 29, 2802–2810, 2016.

## Acknowledgements

The work has been performed under the Project HPC-EUROPA3 (INFRAIA-2016-1-730897), with the support of the EC Research Innovation Action under the H2020 Programme; in particular, the author gratefully acknowledges the support of the School of Informatics at the University of Edinburgh and the computer resources and technical support provided by Edinburgh Parallel Computing Centre (EPCC).

# PORTING OPEN SOURCE CODE INCOMPACT3D ON GPU USING OPENACC

*W. Guo*

*Paul Scherrer Institut, Villigen, Switzerland.*

## Introduction

Incompact3d is a powerful high-order flow solver written in Fortran90 for academic research. In order to make the best use of supercomputers, Incompact3D is using the 2DECOMP&FFT library. The library is optimized for large-scale computations on supercomputers and scales well to tens of thousands of processors on both Cray and non-Cray systems. The library is built on top of MPI. Currently this code only works on CPU machines. The aim of this study is to add some OpenACC directives to Incompact3D and port the code on GPU machines. The second goal is to improve the performance of Incompact3D on GPU machines compared with its performance on CPU machines.

## Methods

The study is carried out on the cluster named GALILEO in CINECA. It has four NVIDIA K80 accelerators per node on 40 nodes. Before porting the code on GPU, PGPROF is used to profile Incompact3D. The profiling study shows that the subroutine named "convdiff" in Incompact3D takes up most of the time. This subroutine is accelerated by adding OpenACC directives and the loops it contains are ported on GPU.

Intermediate variables are created on the devices using the "!$acc enter data create ( )" directive. Input variables are copied to GPU devices before loops are executed. Kernels construct is used in subroutine "convdiff" to parallelize loops and accelerate them on GPU. In "convdiff", some loops are not able to be parallelized due to data dependency. They are executed serially on GPU. When the kernel ends, the data is brought back to CPU. Generally speaking, the more loops are parallelized on kernels, the more acceleration will be reached. However, frequent data movement between CPU and GPU will slow down the computation speed. So it is important to reduce unnecessary data transportation.

## Results

In the subroutine "convdiff", two loops and subroutine "derx" which is called inside "convdiff" are parallelized on GPU. A local timer is added before and after the kernel is executed. The time it takes to run the same part of the code on GPU is compared with the time it spends on CPU. The mesh used for the simulation is 64*65*42.

The code runs serially on CPU using one core. The results are shown below:

```
Time step =      2, Time unit =    0.005
 It takes time    1.6403209010604769E-002
 DIV U* Max=   1.8134400630548338E-002
 DIV U* Moyy=   7.5148911141671119E-004
          0 UT   0.6666052531493888       -6.1433385492803971E-005
 DIV U final Max=   2.0854393548366885E-014
 DIV U final Moy=   1.2215096621870103E-014
 U,V,W max=   0.7546199482718602        1.0112895763858543E-016
   1.2896631772907716E-017
 U,V,W min=  -3.9076054280946039E-018  -1.1779856604055006E-016
  -1.2192988530522931E-017
```

Figure 1. Results from pure CPU

```
Time step =      2, Time unit =    0.005
 It takes time    9.7253959975205362E-003
 DIV U* Max=   1.8134400630548157E-002
 DIV U* Moyy=   7.5148911141672495E-004
          0 UT   0.6666052531493888       -6.1433385492803971E-005
 DIV U final Max=   2.0276064686300839E-014
 DIV U final Moy=   1.2235624694019209E-014
 U,V,W max=   0.7546199482718602        1.0169816377914032E-016
   1.1903188583588990E-017
 U,V,W min=  -4.0428065090429927E-018  -1.1847619239835350E-016
  -1.2268033617098446E-017
```

Figure 2. Results from CPU and GPU acceleration

## Conclusions

From Figure 1 and Figure 2, it can be seen that the loops which are ported on GPU only takes 0.00973s, which is 60.8% of the time needed using pure CPU (0.016s). So, the local acceleration is achieved using OpenACC. The acceleration effect will be more significant if more loops are ported on GPU. If the code runs in parallel, CUDA aware OpenMPI is needed to pass device data between ranks and eliminate the need to bring data back to the host. Due to each node has four accelerators, data allocation on GPU is also needed to be specified. These can be done in the future study.

## References

[1] Laizet et al, Journal of Computational Physics 228.16 (2009): 5989-6015; [2] Laizet et al, International Journal for Numerical Methods in Fluids67.11 (2011): 1735-1757.

## Acknowledgements

The work has been performed under the Project HPC-EUROPA3 (INFRAIA-2016-1-730897), with the support of the EC Research Innovation Action under the H2020 Programme; in particular, the author gratefully acknowledges the support of Eric Pascolo, Pietro Bonfà, Riccardo Zanella, Debora Testi and the computer resources and technical support provided by CINECA.

# PERFORMANCE PORTABLE PDE-BASED SIMULATIONS WITH LIFT

*B. Hagedorn*[1], *M. Steuwer*[2]

[1]*University of Münster, Münster, Germany;* [2]*University of Glasgow, Glasgow, United Kingdom.*

## Introduction

Efficient solving of partial differential equations (PDE) is one of the most important problems of high-performance computing applications in academia and industry. PDEs are used to describe a wide range of physics phenomena like sound, fluid dynamics or quantum mechanics and have important applications ranging from simulating room acoustics to climate change or evaluating the risk of natural catastrophes. Based on the results of large-scale PDE-based simulations, major economic and political decisions are made with sometimes far reaching consequences for the society.

Efficient implementations of PDE solvers are challenging to write especially given the fast-changing landscape of accelerator HPC hardware. Once an efficient implementation in a low-level language has been developed for a particular hardware device the next hardware generation makes these efforts obsolete. This lack of performance portability will only become a more crucial issue with the increasingly diversification of hardware accelerators.

## Methods

Geometric Multigrid (GMG) is a method to solve PDEs which consists of four iteratively applied operations (smooth, residual, restrict and interpolate). The solution on a fine grid is approximated by restricting it to a coarser grid, on which the equation can be solved faster. Iterative application of the restrict operation creates a hierarchy of grids with increasing coarseness. On each level, smooth is applied to reduce high frequency errors. The steps are repeated until a level is reached where to cost of direct solution on the coarsest grid is negligible compared to the cost of the solution of the finest level. This solution is then interpolated up to the finer grids until the solution for the finest grid is computed. All four operations are stencil computations and hence benefit from being executed on many-core architectures like GPUs [1].

Programs executed on GPUs are usually written using low-level programming approaches like OpenCL or CUDA. In order to achieve high performance, expert knowledge is required to optimally manage hardware details such as the memory and thread hierarchy which are explicitly exposed to the programmer by these low-level approaches. However, the optimizations that achieve high performance vary significantly depending on hardware architecture and program, thus, tuning for one device generally leads to poor performance on other devices.

In order for a program to achieve performance portability, that is high performance across multiple architectures, it is necessary to manually optimize the program for every architecture. Lift [2] tackles the challenge of performance portability by automatically exploring the space of possible implementations by applying different combinations of optimizations. A program written in Lifts language is automatically transformed by applying semantics preserving rewrite rules. These rules automatically lower the abstract program by introducing low-level OpenCL implementation details such as how to manage the thread and memory hierarchies efficiently.

## Results

During this collaboration, we implemented the core computation building blocks of multigrid solvers in Lift, allowing developers to compute PDE-based application using automatically Lift-generated compute kernels. Besides the technical progress we disseminated our prior foundational results by presenting our paper “High Performance Stencil Code Generation with Lift” [3] at the prestigious International Symposium on Code Generation and Optimization (CGO'18) where we won the prestigious Best Paper Award. In addition, these results have been presented at the Compiler and Architecture Design group (CArD) at the University of Edinburgh, at the Scottish Programming Languages Seminar (SPLS), at the 20th Workshop on Compilers for Parallel Computing (CPC) and at the Dependable Systems Group at Heriot-Watt University Edinburgh. We presented a tutorial about the Lift Project at the 2018 IEEE International Symposium on Performance Analysis of Systems and Software (ISPASS).

## Conclusions

We are currently evaluating our multigrid PDE-solver implementation focussing especially on two HPC applications: room acoustics simulations and ground penetrating radar. We hope to publish our study with first results soon. This visit has continued an already long-lasting collaboration between the Universities of Glasgow and Edinburgh in Scotland and the University of Münster. We currently continue our collaborative work on the Lift project with a focus on its application in HPC computing. Visitor and host of this HPC-Europa 3 collaboration are keen to continue collaborating in the future.

## References

[1] Basu P et al, Compiler-based code generation and autotuning for geometric multigrid on GPU-accelerated supercomputers, Parallel Computing 64 (2017): 50-64; [2] Steuwer M et al, Generating performance portable code using rewrite rules: from high-level functional expressions to high-performance OpenCL code, ACM SIGPLAN Notices 50.9 (2015): 205-217; [3] Hagedorn B et al, High Performance Stencil Code Generation with Lift, International Symposium on Code Generation and Optimization. 2018.

## Acknowledgements

The work has been performed under the Project HPC-EUROPA3 (INFRAIA-2016-1-730897), with the support of the EC Research Innovation Action under the H2020 Programme; in particular, the author gratefully acknowledges the support of Michel Steuwer (University of Glasgow), the entire Lift team and the computer resources and technical support provided by ICHEC.

# DISTRIBUTED ALGORITHM FOR THE ANALYSIS OF PROPERTIES OF COMPLEX NETWORKS

*F. Magalhães*

*Instituto Superior Técnico, Universidade de Lisboa, Lisboa, Portugal.*

## Introduction

In recent years, the study of networks has received a renewed interest from the scientific community. It blends with a wide range of areas, with networks emerging for example from social interactions, biological phenomena, the world wide web, financial networks, and the power grid.

This paper proposes a Monte Carlo method to compute metrics of complex networks, such as the centrality of nodes in a network. The proposed method uses random walks over a matrix that represents the network, to calculate functions based on the powers of the matrix. In particular, several interesting metrics can be derived, such as the inverse of the matrix multiplied by a vector, corresponding to Katz centrality, and the trace of the inverse of the matrix.

The proposed method aims to be parallel and scalable, in order to deal with large networks, that may exceed the memory limits of individual machines.

## Methods

The proposed solution is based on the method described in [1]. Considering a matrix B, let A = I - B, where I is the identity matrix., then

$$B^{-1} = (I - A)^{-1} = \sum_{k=0}^{\infty} A^k$$

as long as, considering $\lambda _{r}(A)$ as the $r$-th eigenvalue of $A$, the following is held:

$$\max_r \lambda_r < 1$$

In order to approximate $B^{-1}$, the method under study approximates the sum the of first n powers. This is a good approximation for a large enough $n$, since it converges as n tends to infinity. The method calculates independently each row of each power of the matrix A, using random walks, similarly to Markov Chain Monte Carlo.

In order to calculate the approximation of the l-th row of $A^k$, the k-th power of matrix A, the method builds a vector r with the same length as the row, initially with 0 in all entries.

A play starts with the value of 1 and, at each step, multiplies it by the sum of the values of the entries in the current row. When a play reaches the end of its k steps, its value is added to the vector r, at the index of the row it is currently at. In the end, the values in r must be divided by the number of plays, yielding an approximation of the desired i-th row of Ak.

To speed up computation, the sum of all entries in each row is initially calculated and stored in a vector. The matrix representation in memory is similar to CSR, with this added information.

When the objective is simply to calculate the product of a vector by the inverse of the matrix, a vector, which is the case both in systems of linear equations and when calculating the Katz centrality of a network, the algorithm can be modified to store only vectors rather than the full matrix.

Based on the powers of matrices, several metrics can be computed.

## Results

The implementation of the method was tested over different types of networks, with small world properties, simulating real complex networks, and executed in supercomputers featuring high speed networks interconnecting nodes with several processors each. It is shown to scale when the problem size increases, even beyond the memory limits of a single machine. It also scales as desired when increasing the number of processes.

When computing the trace of the inverse, Monte Carlo can be much faster than alternative methods, for a comparable error.

## Conclusions

The proposed method can deal with very large matrices, and surpass individual machine's memory limitations. It is adaptable to compute several metrics of complex networks, at the same time if desired, by computing different functions of the adjacency matrix.

The method can be faster than the existing alternatives at computing an approximation of metrics of very large matrices. Therefore, the existing implementation of this method is suitable for very large problems, where an estimation of the result is desired.

## References

[1] Newman MEJ, Networks: an introduction. Oxford New York: Oxford University Press, 2010.

## Acknowledgements

The work has been performed under the Project HPCEUROPA3 (INFRAIA-2016-1-730897), with the support of the EC Research Innovation Action under the H2020 Programme; in particular, the author gratefully acknowledges the support of José Ramón Herrero from the Department of Computer Architecture and the computer resources and technical support provided by BSC.

# A HIGHLY PARALLEL ALGORITHM FOR COMPUTING THE ACTION OF A MATRIX EXPONENTIAL ON A VECTOR BASED ON A MULTILEVEL MONTE CARLO METHOD

*J.A. Acebrón[1,2], J.R. Herrero[3], J. Monteiro[2]*

*[1]Dept. Information Science and Technology, ISCTE-University Institute of Lisbon, Lisbon, Portugal; [2]INESC-ID, Instituto Superior Técnico, Universidade de Lisboa, Lisbon, Portugal; [3]Dept. d'Arquitectura de Computadors, Universitat Politècnica de Catalunya, Barcelona, Spain.*

## Introduction

A novel algorithm for computing the action of a matrix exponential over a vector is proposed. The algorithm is based on a multilevel Monte Carlo method, and the vector solution is computed probabilistically generating suitable random paths which evolve through the indices of the matrix according to a suitable probability law. We have proved that the complexity is significantly better than the classical Monte Carlo method, allowing the computation of much more accurate solutions. We have developed a highly scalable implementation capable of solving some test problems very efficiently using high performance supercomputers equipped with a large number of cores.

## Methods

A major disadvantage of any Monte Carlo method is the slow convergence rate to the solution of the numerical method, in general of the order O(N-1/2). There already exist a few statistical techniques to mitigate such a poor performance, where the multilevel method clearly stands out. An excellent review has been recently published in [1] describing in detail the method as well as a variety of applications where it was successfully applied.

One of the main contributions of this paper is to develop a multilevel method for the problem of computing the action of a matrix exponential over a vector. This is done by conveniently adapting the probabilistic representation of the solution derived in [2] to the multilevel framework. Figure 1 compares the relative complexity of the methods.

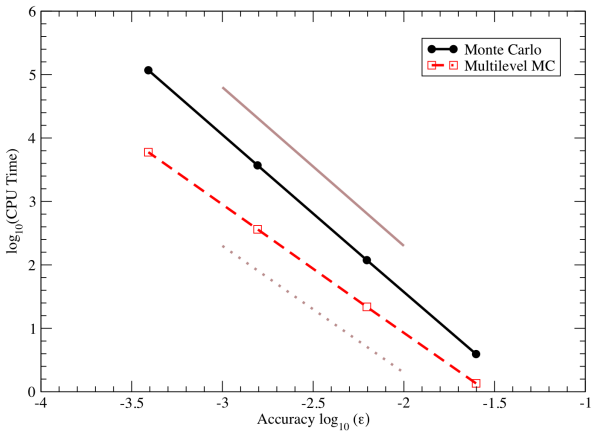



Figure 1 - Computational time as function of accuracy ε, both in log10 scale. The brown solid line corresponds to an ancillary function of slope -5/2, while the dotted line to a function of a slope -2. The results correspond to the communicability for a single node of a smallworld network of size $n=10^6$.

## Results

The following table presents the time required to compute the total communicability of a small-world network of size $n=10^8$ as function of different number of cores for the multilevel method. The accuracy was kept fixed at $\varepsilon=10^{-7}$. The speedup column indicates how much faster the execution is relative to half the number of cores (previous row in the table).

| Cores | Time(s) | Speedup |
|---|---|---|
| 1200 | 315 | |
| 2400 | 175 | 1.8 |
| 4800 | 87 | 2.0 |
| 9600 | 50 | 1.7 |

In all cases, the speedup is very close to the ideal, even for such a large number of cores. This is because most of the calculations are totally independent, corresponding to the Monte Carlo simulations performed at each level of the method. For the defined level of accuracy ε, a very large number of samples is required, exceeding the number of $10^9$ for the coarsest level. Communication is required between levels, but the overhead is negligible.

## Conclusions

The multilevel Monte Carlo method was conveniently recast to be able to compute the action of a matrix exponential over a vector. The underlying algorithm after parallelization has been shown to be highly scalable, which in practice enables simulation of largescale problems for extremely large number of cores. Finally, whenever available, simulations based on a standard Krylov-based method have been conducted, and the performance compared with the multilevel method. The multilevel method clearly outperforms the deterministic method for solving problems consisting in large matrices, not only in terms of computational time, but also as memory requirements.

## References

[1] Giles MB, Multilevel Monte Carlo methods, Acta Numerica, 24 (2015) 259-328. [2] Acebrón JA, A Monte Carlo method for computing the action of a matrix exponential on a vector, submitted (2018).

## Acknowledgements

The work has been performed under the Project HPCEUROPA3 (INFRAIA-2016-1-730897), with the support of the EC Research Innovation Action under the H2020 Programme; in particular, the author gratefully acknowledges the support of Computer Architecture Department at Universitat Politècnica de Catalunya and the computer resources and technical support provided by Barcelona Supercomputing Center (BSC).

# COMMUNICATION REDUCTION IN THE CONJUGATE GRADIENT METHOD

*E.S. Quintana-Orti*

*Universidad Jaume I, Castellón, Spain.*

## Introduction

The Conjugate Gradient method (hereafter, CG) is the Swiss army knife of numerical methods for the iterative solution of sparse linear systems with symmetric positive definite (s.p.d.) co-efficient matrix. The efficient implementation of a preconditioned variant of CG (PCG) for the solution of very large-scale sparse linear systems requires the combination of parallel programming models such as MPI and OpenMP.

## Objective

The general goal for this visit was the development of an MPI+OmpSs (or OpenMP) version of PCG that replaces the non-scalable ILU-type preconditioner present in the current code with a highly- parallel one, based on Jacobi, that allows to assess the impact of communication-reduction techniques in large-scale configurations.

## Results

After the research stay, the following specific goals were achieved:

- Initial port and evaluation. The work to be performed at CINECA commenced with the migration of the current version of PCG to the platforms in the centre. As expected, this stage was smooth as the code only relies on the BLAS libraries, MPI as the message passing tool and OpenMP. The porting was followed by an initial evaluation of the performance and scalability of the code.
- Integration and evaluation of a Jacobi preconditioner. The initial plan to integrate a block-Jacobi preconditioner was replaced by the integration of a plain Jacobi scheme, with a simpler parallelization using MPI+OpenMP. The solution was encoded into the initial PCG method (replacing the ILU-type preconditioner), and the result was evaluated from the points of view of performance and scalability.
- Integration and evaluation of communication-avoiding techniques. The PCG solver as used as a basis to develop new variants of the method that reduce the number of synchronization points. Eijkhouts variant was partially encoded and its evaluation was still an ongoing task at the end of the research stay.
- Integration and evaluation of adaptive precision techniques. The PCG solver was be integrated into an outer-inner solver with mixed precision and iterative refinement.

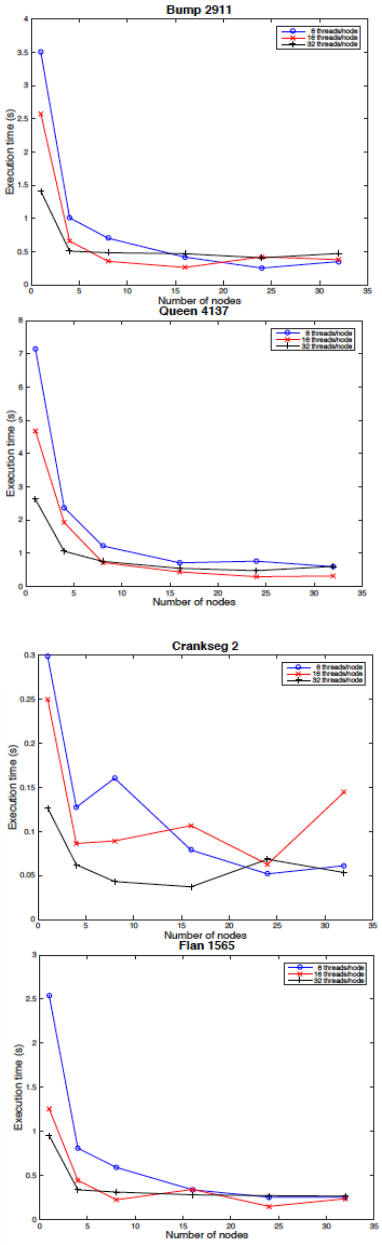



Figure 1: Performance of the MPI+OpenMP version of PCG (with Jacobi preconditioner) executed on Marconi, for four large-scale matrices from the SuiteSparse matrix collection.

## Acknowledgements

This work was performed during Enrique's visit to CINECA, funded by Project HPC-EUROPA3 (INFRAIA- 2016-1-730897), with the support of the EC Research Innovation Action under the H2020 Programme.

# A PARALLEL MULTI-OBJECTIVE ALGORITHM FOR OPTIMISING HIGH-PRESSURE/TEMPERATURE TREATMENTS IN FOOD INDUSTRY

*M. Ruiz Ferrández*

*University of Almería, Almería, Spain.*

## Introduction

In our previous work, we proposed a decision tool to support food engineers in designing high-pressure and temperature (HPT) treatments for food processing. More precisely, it focuses on finding the initial and refrigeration temperatures and the pressure profile to be provided to the HPT equipment such that the final enzymatic activity in the food and the maximum temperature reached during the whole process are minimal and the final vitamin activity is maximal. To this aim, we designed a multi-objective methodology which consists of using a preference-based evolutionary algorithm coupled with a decision tool. Now, in this research stay funded by the HPC-Europa3 program, the main goals are: (i) to migrate the HPT numerical simulation to an open-source solver, (ii) to explore some parallelism strategies applied in both, the HPT numerical simulation and the multi-objective algorithm, and (iii) to implement some mechanisms to improve the convergence of the optimization algorithm and the number of optimal points in the final solution.

## Methods

The HPT numerical simulation consists of numerically solving the heat transfer partial differential equation system coupled with a kinetic equation for each of the involved enzymes or vitamins over an axisymmetric domain [1]. In this work, it has been implemented using the open-source OpenFOAM package and their results were compared with the commercial software COMSOL Multiphysics.

Regarding the optimisation, the multi-objective algorithm used in this work is called Weighting Achievement Scalarizing Function-Genetic Algorithm (WASF-GA) [2]. It is a meta-heuristic method based on a population of points that evolve toward the set of optimal solutions focusing on a region of interest preferred by the person solving the problem. In our industrial problem, the most computationally-expensive task in time is the evaluation of the objective functions as it implies the HPT numerical simulation. Since the evaluation of each point is independent of the rest of the population and there is no need to communicate during this phase, it has been parallelized by assigning a fraction of the population to each available processing element.

Furthermore, we have implemented some mechanisms to improve WASF-GA: the introduction of an external list for storing the most promising points and the generation of an advanced population that accelerates the convergence using local search techniques.

## Results

For the HPT numerical simulation, the temperatures obtained with OpenFOAM and COMSOL are quite similar. However, the computational time consumed by OpenFOAM is much larger than the COMSOL one, so even using some parallelism strategies OpenFOAM cannot compete against the commercial software COMSOL.

The parallel version of WASF-GA exhibits an almost linear speedup when there is enough computational load (see Figure 1). Remember that the speedup is the acceleration experimented by a program using several processing units instead of a single one. Additionally, it shows scalability, which means that the larger the size of the problem to solve, the better its performance.

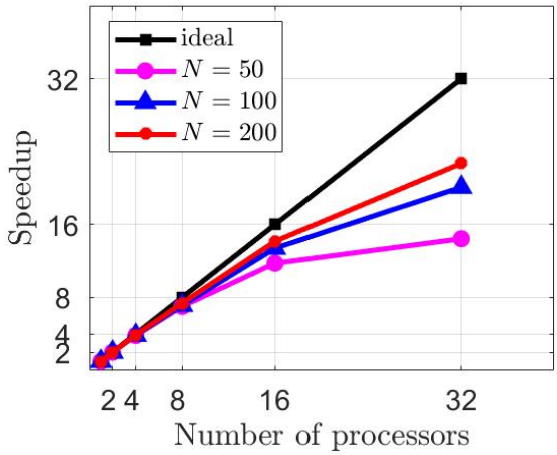



Figure 1 – Speedup of the WASF-GA parallel version varying the number *N* of points in the population

## Conclusions

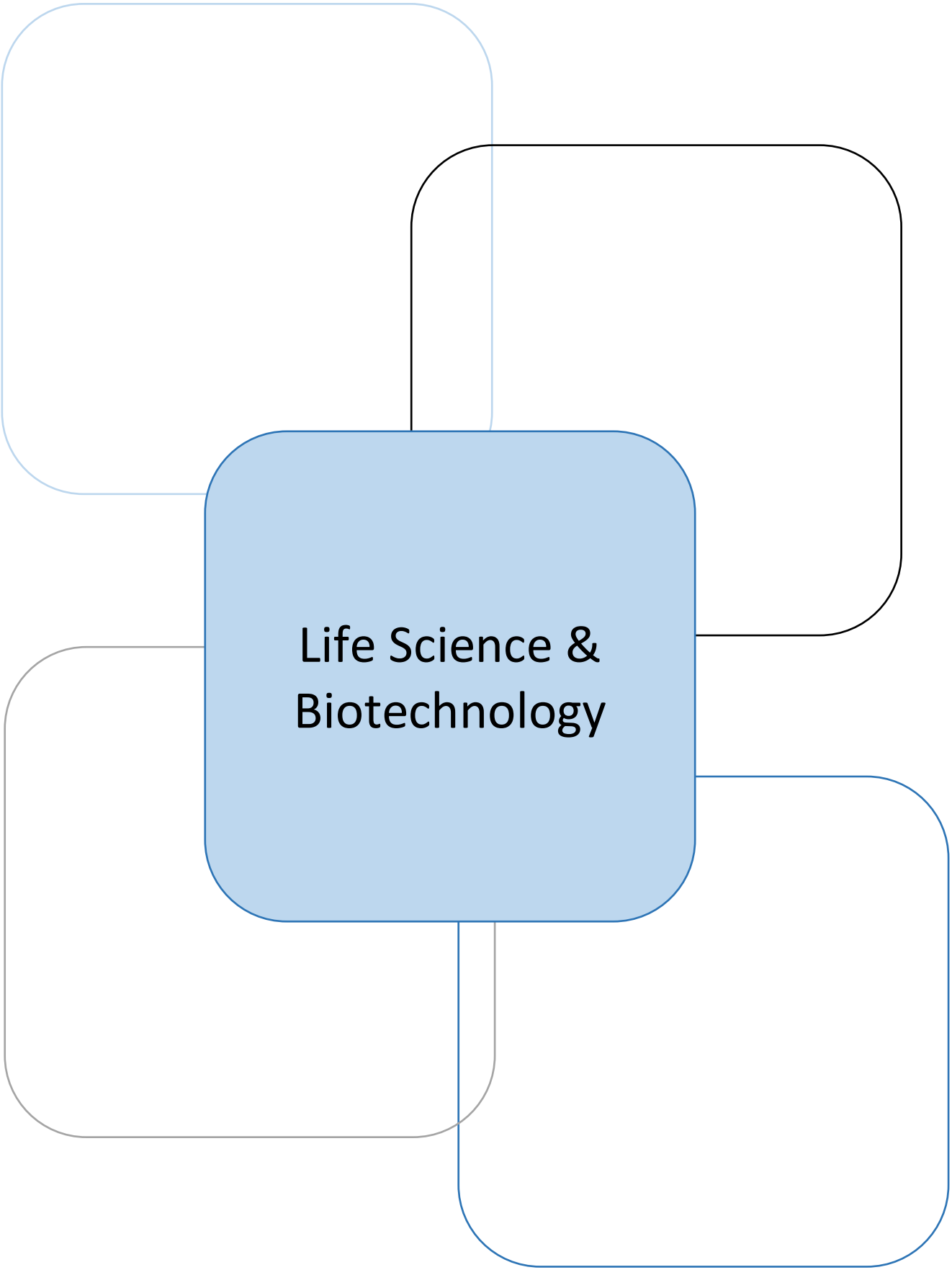
The parallel version of WASF-GA allows a significant reduction in the computational time, so that a larger number of points in the solution set and a larger number of iterations could be considered for the optimisation procedure, which translates into an increase in the number of optimal points provided to food engineers and an increment in the method accuracy, respectively.

## References

[1] Infante JA et al, Math Models Methods Appl Sci, 19(12): 2203-2229, 2009; [2] Ruiz AB et al, JOGO, 62(1):101-129, 2015.

## Acknowledgements

The work has been performed under the Project HPC-EUROPA3 (INFRAIA-2016-1-730897), with the support of the EC Research Innovation Action under the H2020 Programme; in particular, the author gratefully acknowledges the support of Ben Paechter from the School of Computing at Edinburgh Napier University and the computer resources and technical support provided by EPCC.


Life Science &
Biotechnology

# ANTIOXIDANT ACTIVITY VS. PHOTOCHEMICAL PROPERTIES OF PLANAR, BIDENTAL CHELATING, DYES

*N.P. Benetis[1], M. Paloncýová[2], S. Knippenberg[2]*

*[1]Dept. of Environmental Engineering and Antipollution Control, TEI of Western Macedonia, Kila Kozanis, Greece; [2]Division of Theoretical Chemistry and Biology, School of Biotechnology, Royal Institute of Technology, Stockholm, Sweden.*

## Introduction

The planar bidental chelators with conjugated double bond system quercetin and luteolin, among other polyphenols, as well as o-phenanthroline and neocuproin, are known as antioxidants, protecting eukaryotic cells from DNA single strand breakage by hydrogen peroxide [1]. Their photochemical properties render them also environmental cellular probes at the molecular level. Both these characteristics depend on their affinity and a possible passage through cell membrane. Molecular dynamics simulations were performed with the quest of possible incorporation of the above substances in the bilayer, seeking also their position and the orientation of the ground to the excited states' transition electric dipole moments with respect to the membrane surface.

## Methods

The properties of the molecular probes of interest were studied by means of MD simulation and QM/MM hybrid Quantum Mechanics Molecular Mechanics computations. The position and the orientation of these probes in the cell membrane is of utmost importance for confocal spectroscopy investigations, governed by photo-selection and the relative alignment of the transition dipole moment vector compared to the electric field polarization [2]. The photochemistry studies require also knowledge of the excited stated of the probes. Two parallel methods were utilized for this purpose. ADC, Algebraic Diagrammatic Construction scheme for the polarization propagator provides a series of methods for the calculation of excited states. The optical properties of ortho-phenanthroline were studied in particular because the literature data for this molecule are relatively scarce. Complementary optical data for the probes of interest were obtained by Time Dependent Density Functional Theory TD-DFT.

The study of their antioxidant properties, also in focus of this investigation, can be enriched utilizing the optical properties of the above molecular probes as Ligands chelating Fe(II) and Cu(I) ions. In particular this ability is both related to their photo excitation and the antioxidant activity. Aiming to investigate the Fenton reaction mechanism of Me-$(H_2O)_x$-Ligandy adducts we sought for the anticipated strong Me-O or Me-N bonds, using NBO, Natural Bond Order computations.

## Results

The MD simulation and the QM/MM study of the above-mentioned molecular probes along with the caffeic acid methyl ester have shown significant affinity to cell membranes, judging from their interaction with DPPC based model membrane. Running the MD trajectories of the probes alone the stationarity states were derived, preparing for the MD/MM investigation. The obtained stationarity times are: 196, 176, 113, and 142 ns respectively for Caffeic acid methyl ester, luteolin, o-phenanthroline and quercetin.

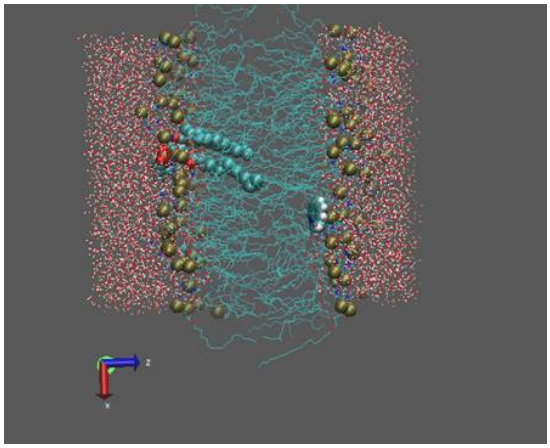


Figure 1 - Snapshot from Molecular Dynamics (MD) simulation trajectory of DPPC- based model membrane in presence of the antioxidant drug, the bidental planar chelator ortho-phenanthroline, at 50 °C. A VDW representation of a whole molecule of the DPPC lipid (light blue beads) is part of the left halve leaflet of the membrane. The VDW representation of the drug, white beads in the right edge of the hydrophobic zone of the bilayer, has entered in the bilayer starting from the right outer water zone of the water/ bilayer system. The phosphorus atoms in the lipids head groups are represented as gold spheres.

## Conclusions

The MD simulatioas showed that all the five sudied probes entered the cell mebrane from the water zone in the first decades of nsec and remained in the lipophilic part of the membrane close to the head groups of the lipids. However, a passive penetration of the model membrane was not possible.

## References

[1] Melidou M et al, Free Radical Biology & Medicine 39 1591 – 1600, 2005; [2] Osella S et al, J. Chem. Theor. Comput. 12, 6169–6181, 2016.

## Acknowledgements

The work has been performed under the Project HPC-EUROPA3 (INFRAIA-2016-1-730897), with the support of the EC Research Innovation Action under the H2020 Programme; in particular, the author gratefully acknowledges the support of Department of Materials and Environmental Chemistry (Arrhenius Laboratory, Stockholm University, Sweden) and the computer resources and technical support provided by PDC in KTH in Stockholm, Sweden.

# INVESTIGATION OF NEW GLYCOSIDE DERIVATIVES AS LIGANDS FOR BIOMEDICALLY RELEVANT LECTIN MOLECULAR MODELING OF GALECTIN BINDING COMPOUNDS

*D. Capasso[1], I. Galdadas[2], A. Iadonisi[3], F.L. Gervasio[2,4]*

*[1]Department of Pharmacy, University of Naples Federico II, Naples, Italy; [2]Department of Chemistry, University College London, London, United Kingdom; [3]Department of Chemical Sciences, University of Naples Federico II, Naples, Italy; [4]Institute of Structural and Molecular Biology, University College London, London, UK.*

## Introduction

Lectins are carbohydrate-binding proteins that recognise various carbohydrates attached to proteins and lipids, known as glycoconjugates, on cell surfaces and extracellular matrices. Galectins, one of the lectin families, are able to recognise β-galactoside sugars with high specificity. Some galectins, unlike most of the lectins, are not membrane bound but soluble proteins with both intra- and extracellular functions. Galectins have important functions in several aspects of cancer biology, such as in regulation of apoptosis, tumour metastasis, migration, and angiogenesis[1]. Galectins are often overexpressed in cancerous cells, modulating tumour progression and influencing the disease outcome[2]. Therefore, development of potent and selective small inhibitors capable of reducing the activity of galectins is of great importance. Many natural or synthetic carbohydrate ligands of galectins can inhibit their biological activity, even to nanomolar range of activity[3]. Simple molecules containing two saccharide residues with a bridging sulfur atom can display high affinity towards galectins and potential anti-cancer activity[4]. Moreover, the replacement of the oxygen atom in the glycosidic linkage with sulfur ensures an increased biostability of the molecule by precluding its hydrolytic cleavage when used as pharmaceuticals. The goal of the current project is the rational design of new glycomimetic inhibitors, featuring high affinity, stability, and selectivity towards Gal-1, Gal-3, and Gal-9. In particular, new diglycosylated compounds equipped with alternative bridging groups between the sugars such as a disulfide moiety. In preliminary studies, glycosyl disulfides were, indeed found to exhibit cytotoxic activity against some tumour cell lines[5].

## Methods

Prior to docking calculations, the crystal structures of Gal-1 (PDB ID 4Y24), Gal-3 (PDB ID 5H9P), and Gal-9N isoform (PDB ID 3WLU) were prepared using the Protein Preparation Wizard of Schrödinger suite[8] in order to add any missing atom to the starting protein structures. The protonation state of each residue was determined using the PROPKA algorithm at pH 7.0 as implemented in Maestro. Crystallographic water molecules that formed less than two hydrogen bonds with non-water atoms were removed. The co-crystalised ligands were then removed and the resulting apo structures were simulated for 1μs each in the NVT ensemble. The MD simulations were performed using GROMACS 5.1.4. Each system was described by the Amber99SB*-ILDN[9] force field with the dihedral corrections of Best and Hummer[10], solvated with TIP3P water molecules[11] and enclosed in a cubic box with periodic boundary conditions. The conformations obtained from the MD simulations were then clustered using an RMSD-based algorithm using a 1.2Å criterion.

To dock the proposed compounds into the conformations of Gal-1, Gal-3, and Gal-9N that resulted from the MD simulations, we used the extra precision (XP) scoring function and docking protocol of Glide 7.6[12] as implemented in the Schrödinger suite. Glide searches for favourable interactions between a ligand and a receptor. In our study, Glide was run with flexible ligand conformations, where multiple conformations of the ligand were generated before docking.

## Results

Among the thiogalactosides described in the literature, we selected the simplest β-galactoside sugar, thiodigalactoside, and its selenium counterpart, as lead compounds and we tried to optimize their structures further. Derivatization at position C3 has been proposed to improve the affinity towards galectins through favourable interactions with charged residues[6]. Although derivatization at positions C4 and C6 is feasible, the hydrogen bond formed between the β-galactoside scaffold and galectins indicates that these interactions not only stabilise the β-galactoside in the binding site, but are also necessary for substrate recognition. Therefore, we have focused on potential modifications only at position C3 (Figure). After the comparison of several crystal structures of Gal-1, Gal-3, and Gal- 9N bound with different ligands, we have come up with substitutions that may enhance both the affinity and selectivity towards specific galectins. To test the potential use of 3,3'-derivatised thiodigalactosides as inhibitors of galectin-1, -3, and -9N, we have created a virtual library of candidate compounds and docked these compounds to all three structures, assessing their binding poses and binding scores. Moreover, to identify alternative conformations of each receptor that might be druggable and that could be exploited to design inhibitors with higher specificity towards each galectin, we run molecular dynamics (MD) simulations of hundred nanoseconds for each galectin, clustered the resulting conformations and then docked the virtual library to each of the newly identified conformations.

The presence of many arginines in the binding groove of all three galectin structures led us to test the affinity of ligands when different arene modifications are introduced in position C3. The exploitation of arene-guanidinium interactions in drug design has attracted considerable interest in recent years since such interactions can be as strong as cation-anion interactions[7]. Therefore, the proposed structures were based not only on the Glide scores but also to the binding pose of each compound and our chemical intuition. On the basis of the *in-silico* results, the most promising compounds will be synthesized, and their biological activity will be assessed on tumour and human endothelial cells.

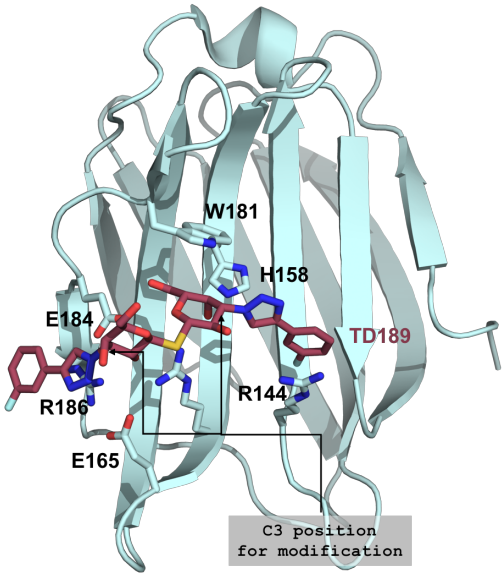



Figure **1** - Crystal structure of Gal-3 (cyan) in complex with the thiodigalactoside TD139 (red). Important residues for the binding are depicted in (cyan) sticks.

## References

[1] Nat. Rev. Cancer 2005, 5(1):29-41; [2] Chembiochem 2017, 18(8):782-789; [3] Bioorganic Med. Chem. 2008, 16:7811-7823; [4] Angiogenesis 2011, 14:293-307; [5] Org. Biomol. Chem. 2011, 9:6278-83; [6] Angew. Chem. Int. Ed. 2005, 44:5110-5112; [7] Proc. Natl. Acad. Sci. USA 1999, 96:9459-9464; [8] Schrödinger LLC, New York, 2018; [9] Proteins 2010, 78(8):1950–1958; [10] J Phys Chem B 2009 113(26):9004–9015; [11] J. Chem. Phys. 2004.121, 10096–10103; [12} J. Med. Chem. 2006, 49(21):6177-6196.

## Acknowledgements

The work has been performed under the Project HPC-EUROPA3 (INFRAIA-2016-1-730897), with the support of the EC Research Innovation Action under the H2020 Programme; in particular, D.C. gratefully acknowledges the support of Department of Chemistry, University College London and the computer resources and technical support provided by EPCC.

# COMPARATIVE ANALYSIS OF MOLECULAR MOTIONS IN SIRTUIN2 PROTEINS

*S. Dotolo*[1,2,3], *A. Facchiano*[1], *A. Pandini*[3]

[1]*Institute of Food Science, National Research Council, Avellino, Italy;* [2]*Dept. of Biochemistry, Biophysics and General Pathology, University of Study of Campania, Naples, Italy;* [3]*Dept. of Computer Science, Brunel University of London, London, United Kingdom.*

## Introduction

Sirtuin2 is an NAD+-dependent protein deacetylase, evolutionarily conserved from bacteria to humans. CobB is a bacterial Sirt2 homologue characterized by a large Rossmann-fold domain and a smaller Zinc-binding domain (critical point for substrate recognition in bacteria). It was previously suggested that selective substrate-binding in CobB is mediated by distal molecular interactions between the zinc-binding domain and the pocket residues around the acetyl-lysine modified residue on the substrate. This support the hypothesis of an allosteric regulation of SIRTUIN2 proteins. However, the molecular mechanism involved in this process, it is still unknown. We here present a computational study of the molecular dynamics of CobB to unveil SIRTUIN2 mechanism of substrate recognition and to understand if there are mechanisms of allosteric regulation that could explain the role of distal molecular interactions.

## Methods

The workflow is based on the combined application of residue coevolution and molecular simulation analysis: [a] The dynamics of the two systems was simulated using tCONCOORD. Functional local correlated motions will be detected using GSATools and GROMACS 2016. Coevolved positions were detected and recorded using MetaPSICOV and DCA. Network and pathways of signal transmission between the zinc-binding region and the substrate binding site was extracted and annotated with the location of co-evolved positions. [b] Local correlated motions and communication pathways will be extracted as in [a] for SIRT2, but the critical residues with functional role will be selected by comparison with the list of functional and co-evolved residues in CobB. Mapping of literature annotated pathogenic mutations will further inform on the relevance of critical residues.

## Results

A Multiple Sequences Alignment (MSA) of all sequences with SIR2 domain was generated from CobB in E.coli K12 (UNIPROT P75960; PDB code 1S5P). The MSA has been analyzed and filtered by means of R (Bio3d) and Python (BioPython), to retain the human sequences, choosing for human sirtuin2 the structure PDB 5D7P (UNIPROT Q8IXJ6) [Figure 1]. Through combined use of analysis of loal correlation motions (in the form of Mutual Information matrices derived with GSATools) and residue coevolution (in the form of MetaPSICOV matrices), it was possible to detect the evolutionary properties of domain organization, interacting residues and functional structural transitions. The co-evolution profile of DCA (Direct Coupling Analysis) matrix has been generated a domain decomposition to identify the dynamics and evolutionary-related domains, using SPECTRUS-evo web server, based on the analysis of co-evolutionary couplings between pairs of amino acid positions in a multiple sequence alignment. The resulting clusters are consistent in relation domains. In more details, the coevolution coupling by analysis MetaPSICOV matrices demonstrated consistency with the domain organization. It has evident a coevolution coupling between the zinc-binding domain and active site.

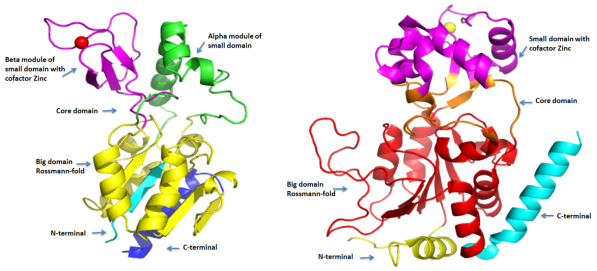



Figure 1 – Structure of CobB (1S5P) and human sirtuin 2 (5D7P).

The conformational changes of the two systems were simulated using tCONCOORD. It has been possible, to predict protein conformational flexibility of ensembles of structures, based on geometrical considerations and representing a good sampling of the conformational space. After that, we have analyzed different properties, applying some important GROMACS 2016.4 tools: gmx and GSAtools. The first tool has been used to calculate the trajectories and to investigate some important dynamics properties based on principal collective motions. It has been seen that some dynamics properties are preserved. This information is important, because this result is not expected. The analysis of local correlated motions has been executed using GSAtools, highlighted similarity in the most important local conformational changes of the two proteins. This is reported as centrality measure derived from the network of local correlation in the MI matrix [Figure 2].

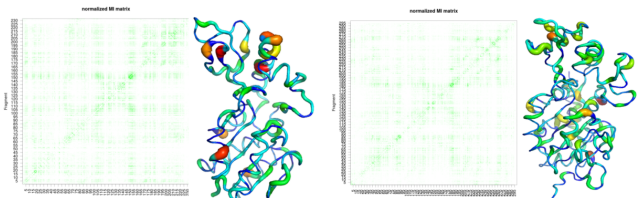


Figure21 – MI and network of CobB and human sirtuin 2.

## Conclusions

Residues position identified on CobB have been mapped on the human Sirt2 in order to evaluate if the distal molecular interactions to recognize a specific substrate have been conserved. It is necessary to execute the molecular dynamics simulations, using GROMACS to validate the tCONCOORD results and to calculate the new dynamics properties of two systems. From the profile of the wild-type protein we planned to investigate the dynamics of pathological mutants and the existence of compensatory mutations. This approach could lead to the identification of putative allosteric sites for novel drug-design. Finally, it could be useful apply this protocol also on human Sirt2 in order to evaluate if the distal molecular interactions to recognize a specific substrate have been conserved or not.

## References

[1] Rumpf T et al, Seeding for sirtuins: microseed matrix seeding to obtain crystals of human Sirt3 and Sirt2 suitable for soaking. Acta. Crystallogr. F. Struct. Biol. Commun. 2015;71(Pt 12):1498-510; [2] Zhao K et al, Structure and Substrate Binding Properties of cobB, a Sir2 Homolog Protein Decetylase from Escherichia coli. J. Mol. Biol. 2004; 337(3):731-41; [3] de Juan D et al, Emerging method in protein co-evolution. Nat. Rev. Genet. 2013; 14(4):249-61; [4] Pandini A et al, The Gearbox of the Bacterial Flagellar Motor Switch. Structure. 2016;24(7):1209-20; [5] Pandini A et al, GSATools: analysis of allosteric communication and functional local motions using a structural alphabet. Bioinformatics 2013; 29(16): 2053–2055.

## Acknowledgements

The work has been partially supported by Traineeship Erasmus plus grant and HPC-Europe3 project.

# THE INTEGRATION OF MOLECULAR NETWORKS UNCOVERS THE MECHANISMS OF DRUG SENSITIVITY IN CANCER THERAPY

*A. Federico[1,2], G. Scala[3,4], V. Marwah[3,4], A. Serra[3], V. Costa[2], A. Ciccodicola[1,2], V. Fortino[5], D. Greco[3,4]*

*[1]Department of Science and Technology, University of Naples "Parthenope", Naples, Italy; [2]Institute of Genetics and Biophysics "Adriano Buzzati Traverso", CNR, Naples, Italy; [3]Faculty of Medicine and Life Sciences, University of Tampere, Tampere, Finland; [4]Institute of Biotechnology, University of Helsinki, Helsink , Finland; [5]Institute of Biomedicine, University of Eastern Finland, Kuopio, Finland.*

## Introduction

In the last decade, somatic alterations were identified in certain genes driving cancer development [1, 2, 3]. Moreover, several of these alterations are implicated as determinants of treatment response in the clinic [4]. Recently, longitudinal studies from The Cancer Genome Atlas and the International Cancer Genome Consortium have generated comprehensive catalogues of the cancer driver genes involved in tumorigenesis across a broad range of cancer types. However, the so far developed pharmacological strategies based on the somatic background suffer from multiple limitations. In this work, we developed a computational framework aimed at overcoming the limitations of the classical pharmacogenomic approaches, uncovering the therapeutic potential of deregulated cellular processes in cancer cells. Here, we present a case study on the invasive breast cancer cohort from The Cancer Genome Atlas.

## Methods

The breast cancer gene expression data have been retrieved from The Cancer Genome Atlas (TCGA) [TCGA. Available online: https://tcga-data.nci.nih.gov/tcga/]. The differentially expressed genes were identified performing the Wilcoxon signed-rank test and the F-test on the variance and the union of the genes within the top 10% of the distributions were selected for further analysis. The co-expression network was inferred using the INfORM tool [5]. The topological functional annotation was based on the Reactome database. The druggability has been based on the latest annotation of Drugbank database. The network-coverage analysis has been performed through the use of a customized genetic algorithm.

## Results

For the differential gene expression analysis, 112 patients were considered, for a total of 224 RNA-Seq datasets. Through the differential expression analysis, we identified 4347 most differentially expressed genes [see Methods section] to be considered for further analyses. Subsequently, we inferred the highly responsive modules from the top deregulated genes through the use of the INfORM software. We analysed the patterns of connectivity within the graph by the walktrap algorithm and identified 165 gene communities. To select the communities with a biological relevance, we filtered out all the communities constituted by less than 10 genes. In this step, we selected 35 biologically relevant communities of highly connected genes with an average dimension of 132 genes. In order to assess the role of each of the detected communities of co-expressed genes, we carried out a topological functional analysis. 29 out of 35 detected communities were significantly enriched in at least one pathway (pval < 0.01). In particular, we identified "extracellular matrix organization and degradation" as the most deregulated process. Thereafter, we built a drugability map integrating the results obtained in the previous sections of this work. We identified 5 drugs targeting 10 communities, 148 pathways, with a median network rank of 5, indicating that the drug-targeted genes are the most connected hubs. Finally, the drugs selected by the genetic algorithm are quazepam, copper, halothane, doxorubicin and hymenialdisine (Drugbank entries: DB01589, DB09130, DB01159, DB00997, DB02950, respectively) [Figure 1].

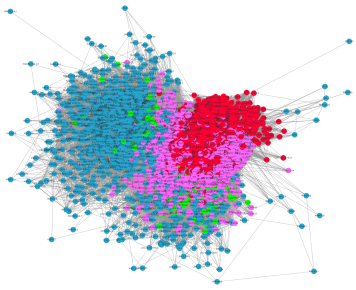


Figure 1 - Drugs selected by the genetic algorithm

## Conclusions

In this case study, we identified cellular processes that can be targeted from certain drugs, both known to be effective against cancer cells and new compounds candidate to repositioning. In particular, we described modules of coexpressed – functionally related – drugable genes. Although the finding of suitable drug molecular targets is still challenging, this pilot study paves the way to the definition of a drugability map based on the functional deregulation of cancer cells. Further work is needed to make this computational framework applicable at patient-scale level.

## References

[1] Sonawane AR et al, Understanding Tissue-Specific Gene Regulation. Cell Rep. 2017 Oct 24;21(4):1077-1088; [2] Bailey MH et al, Comprehensive Characterization of Cancer Driver Genes and Mutations. Cell. 2018;173(2):371-385; [3] Sanchez-Vega F et al, Oncogenic Signaling Pathways in The Cancer Genome Atlas. Cell. 2018;173(2):321-337; [4] Mok T et al, A small step towards personalized medicine for non-small cell lung cancer. Discov Med. 2009;8(43):227-31; [5] Marwah VS et al, INfORM: Inference of NetwOrk Response Modules. Bioinformatics. 2018;34(12):2136-2138.

## Acknowledgements

The work has been performed under the Project HPC-EUROPA3 (INFRAIA-2016-1-730897), with the support of the EC Research Innovation Action under the H2020 Programme; in particular, the author gratefully acknowledges the support of the computer resources and technical support provided by CSC - Center for Science.

# STRUCTURAL AND DYNAMIC STUDIES ON DOPAMINE-2 RECEPTOR BINDING SELECTIVITY TOWARDS G-PROTEINS AND ARRESTINS: A METADYNAMICS APPROACH

*C. A.V. Barreto*

*Data-driven Molecular Design, CNC - Center for Neuroscience and Cell Biology, University of Coimbra, Coimbra, Portugal.*

## Introduction

Since its discovery, dopamine has attracted an incredible amount of attention due to its critical role in cell-signalling and regulation of dopaminergic pathways [1]. Dopamine is released into the synaptic cleft and activates the Dopamine Receptor (DR), a member of the super-family G-Protein Coupled Receptors (GPCRs) [1]. DRs' mode of action is regulated by G-protein activation [2]. However, growing evidence suggest that DRs do not signal exclusively G-proteins and can engage in G-protein independent pathways [3]. A predominant independent pathway involves recruitment of β-arrestin, which is classically associated with GPCRs' desensitization [4]. As such, drugs can induce or stabilize a receptor in different conformations, leading to the activation of different downstream pathways, phenomenon known as "functional selectivity" [5]. Here, to address how drug dependent GPCR signalling is modulated by "functional selectivity", we aimed to investigate and characterize, at the atomic-level, the modes of interaction of D2R with its two predominant signalling partners, G-Proteins and Arrestins, and explore the dynamic properties of the complex systems in their natural environment.

## Methods

Preparation of Models: Models of D2R and internal partners Gi and Arrestin-3 were obtained through homology model protocol. Active models were refined using refinement interface of HADDOCK 2.2 [6]. To prepare the PDB files for the refinement stage each model complexed with Gi1 and Arrestin-3 structures were aligned with 3SN6 [7] and 4ZWJ [8] structures, respectively, using the PyMOL software.

Molecular Docking of Dopamine: AutoDockTools [9], a package of MGLTools was used to perform docking parameterization. Docking itself was performed using Autodock4.2 (version 4.2.6, released in 2009) [9]. Using the top clusters from each run, all distances between the center of mass of the dopamine and the alpha-C-atom (Cα) of the flexible residues used in the docking were measured to attain an initial evaluation of these models. BINANA software [10] was also used to perform further analysis of the dockings.

Metadynamics of the Receptor-Ligand-Protein: a) System setup: Before setting up the two systems, the Dopamine-2 Receptor selected models were subjected to PPM server [11] to calculate spatial orientations respecting to the Membrane Normal defined by the Z-axis. The system was set up in Membrane Builder function of CHARMM-GUI server [12,13]. In the initial steps Dopamine-2 Receptor models were protonated in the residue ASP52, two SS bonds were defined in CHARMM36 and a palmitoyl was added to the last residue. Further, two lipid tails were added to Gi first two residues, a myristoyl and a palmitoyl, respectively. Parameters for dopamine were generated with CHARMM General Force Field tool. All the terminals were patched except for those where a lipid tail was added. The complexes were then embedded in a POPC:CHL membrane with a ratio of 9:1. NaCl ions were added to until it reached a total concentration of 0.15M. Final systems included approximately 370 1-Palmitoyl-2-oleoylphosphatidylcholine (POPC), 40 Cholesterol (CHL), 300 ions and 50 000 water molecules. b) System Initiation: Energy minimization and equilibration of the system were performed with GROMACS 5.1.1. [14]. Systems were relaxed to remove any possible steric clashes by a set of 50000 step of Steepest Descent energy minimization. For equilibration, an adapted version of the protocol from Schneider et al. was applied [15]. The system was heated using Nosé-Hoover thermostat from 0 to 310.15 K in the NVT ensemble over 1 ns with harmonic restraints of 1000 kJ/mol/nm2 to membrane, proteins and dopamine atoms. Then a 1 ns equilibration NPT run was performed using the Parrinello-Rahman pressure coupling at a reference pressure of 1 bar. To equilibrate the membrane two 1 ns NPT runs were performed, with gradually reduced positional restraints to lipids heavy atoms (500 kJ/mol/nm2, 250 kJ/mol/nm2), followed by a 1 ns unrestrained run. The restraints of receptor were also reduced gradually in three additional NPT runs of 1 ns each, first with 500 kJ/mol/nm2 for all heavy atoms and then 500 kJ/mol/nm2 and 250 kJ/mol/nm2 for C-alpha atoms. The restraints of dopamine were maintained for all steps. Finally, a 2ns simulation without any restrains is performed to complete the equilibration of the system under study. c) Collective variables for Metadynamics: The choice of collective variables (CVs) is one of the most important steps for metadynamics simulation. A CV is a function of the microscopic coordinates of the system and it should distinguish between initial and final state and describe all the intermediates. The number of CVs used should be kept small because the use of many CVs implies that high-dimensional space must be explored [16]. In this project the aim is to analyze the interaction between the receptor and the ligand and between the receptor and the intracellular partner and how the one affects the other. For that at least one CV must be defined to characterize each interaction. To characterize the ligand-receptor interaction a protocol based on Schneider et al [15] was applied. The distance between the center of mass of Dopamine and the center of mass of receptor helical bundle (Cα atoms). Characterizing the interaction between receptor and G-protein/Arrestin is more challenging since it is a large interface with multiple interaction points. However, Masone and Grosdidier describe a good method for protein-protein interactions [17]. The idea is to use the number of intramolecular hydrogens bonds as a collective variable to characterize the docking of two proteins. Their approach was to use every residue on the monomers, even the ones that were not on the surface. We had knowledge of the potential interacting residues in each case from literature revision and in silico classification by SPOTON [18]. We will restrain the covalent variable only to intracellular regions of each receptor to save computational time. Receptor regions such as ICL2, ICL3 and junction of TM7 and helix-8 seem to be the most important regions for G-Protein coupling. These regions interact with αNβ1-loop, area around β2-β3, the area around α4-β6 and c terminal of α5 helix seem to be interactions points for Gi. Unfortunately, the same level of knowledge is not available for Arrestin-3, it is only known that For Arrestin-3, recent structural studies propose that the finger loop on N-

domain and the C-loop and the short helix on C-domain are the regions that interact with receptor [8,19,20]. A study conducted by Peterson et al. [21] showed selective mutants for G-Protein and Arrestin coupling. A mutant with only G-Protein binding was D2R:L125N-Y133L and with Arrestin binding was D2R:A135D-M140D. All these residues are located around the DRY motif in TM3, a critical alpha helix for transmission of conformational changes from ligand binding to signaling molecules [21]. d) Metadynamics Simulations: For ligand binding sampling used an initial height of 0.8 kJ/mol for the bias, a width of 0.0125 nm for the Gaussian applied to the ligand-receptor distance and a deposition interval of 2500 steps with a bias factor of 15 [15]. For the second CV we applied a HBOND_MATRIX function of PLUMED with a r0 2.5 [17].

## Results

Docking analysis with Pymol and BINANA: We selected the best three dockings from Autodock poses and clustering. Docking poses from G-Protein model have overall better binding energy than the Arrestin models (Autodock scores). The top three cluster from 3SN6 model have a lowest binding energy of -10.11 kcal/mol, -9.09 kcal/mol and -7.92 kcal/mol with a population of 28, 4 and 5, respectively. For 4ZWJ [8] model the top three clusters have a lowest binding energy of -9.15 kcal/mol, -8.87 kcal/mol and -8.47 kcal/mol with a respective population of 1, 1 and 34. The top 3 clusters from each model were submitted to measurement analysis by python scripting using PyMOL. This simple approach was applied to have a general idea of what residues were closer to the ligand. Both models have close contacts with ASP86, although the pose from G-Protein active model seems to be closer to the serine cluster of TM3 than the pose from Arrestin active model. This difference is interesting since the serine cluster has been reported to be crucial for effective D2R-G-Protein coupling [22]. BINANA results further emphasize the differences in contacts observed in the initial analysis. Hydrogens bonds and salt bridges between the receptor and dopamine were found on both models. However, only the G-Protein active model establishes contacts with the serine cluster.

Metadynamics: The initial snapshot of each system is presented in Figure 1. We run 150 ns of dynamics for each system within this project duration, which is not sufficient to conclude any meaningful movement or conformation. However, in the nanoseconds produced the system was stable with no steric clashes.

## Conclusions

Since both systems were stable within these first nanoseconds of simulation, we can conclude that the system assembles and initiation protocol is valid for GPCR complexes. From the simulation possible to produce within this project it's not possible yet to conclude any meaningful movements, since it takes at least 500 ns to have valuable data in metadynamics simulation. Further extensions of the simulations are necessary and will be performed.

## Acknowledgements

The work has been performed under the Project HPC-EUROPA3 (INFRAIA-2016-1-730897), with the support of the EC Research Innovation Action under the H2020 Programme; in particular, the author gratefully acknowledges the support of Prof. Alexandre Bonvin and the Computational Structural Biology group at University of Utrecht and the computer resources and technical support provided by SURFSARA.

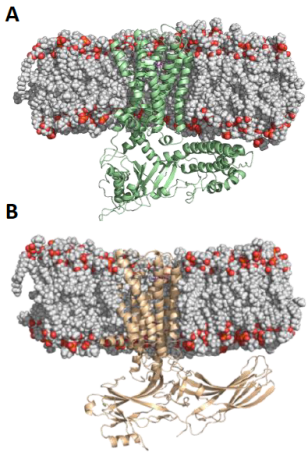
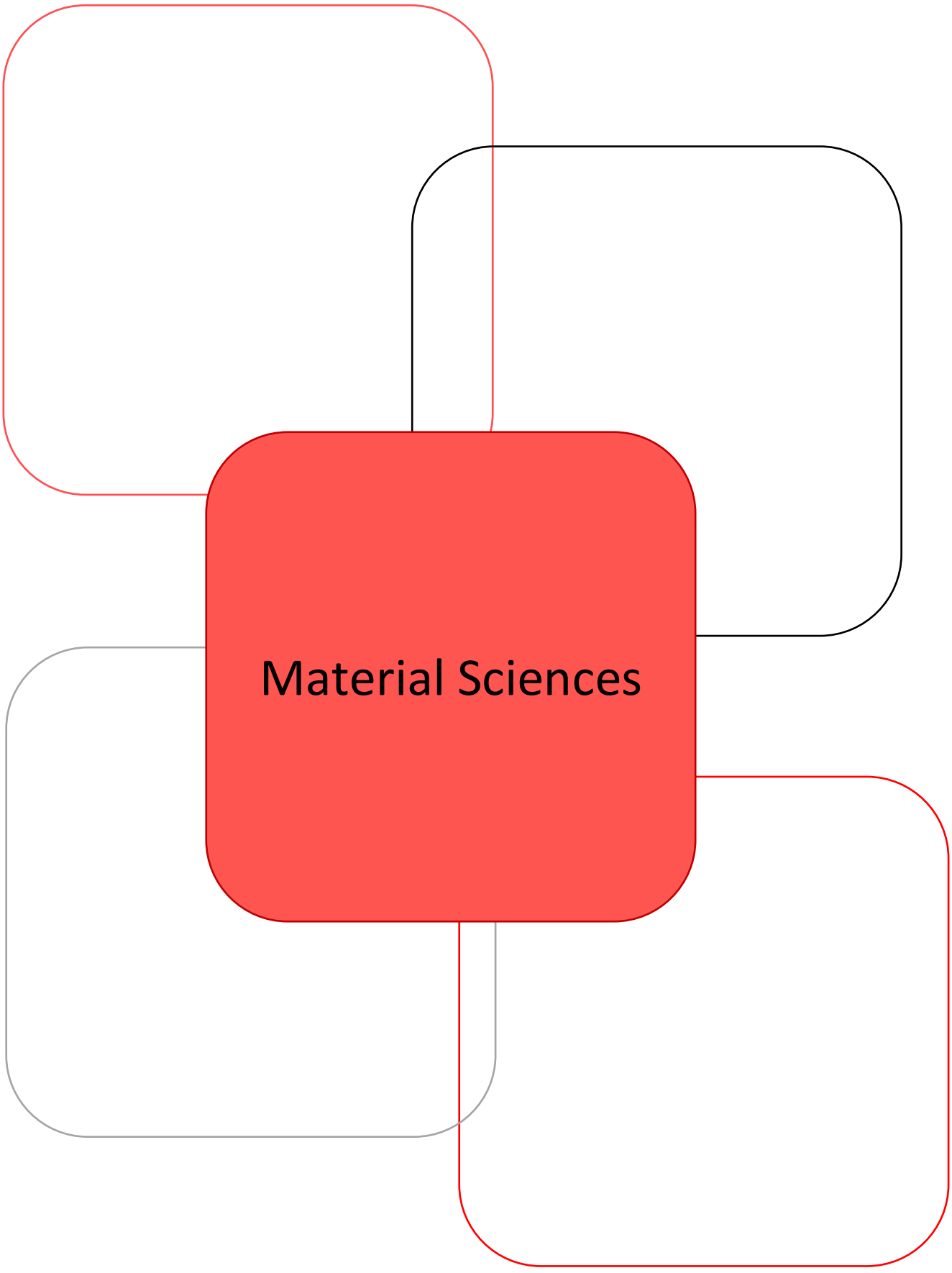



Figure 1 - A) D2R-Gi complex with Dopamine coupled embedded in a bilayer lipid. B) D2R-Arr3 complex with Dopamine inside a bilayer lipid membrane.

## References

[1] Beaulieu J-M et el, The Physiology, Signaling, and Pharmacology of Dopamine Receptors. Pharmacol. Rev. 2011, 63(1), 182–217; [2] Beaulieu J-M et el, Dopamine receptors - IUPHAR review 13. Br. J. Pharmacol. 2015, 172(1), 1–23; [3] Luttrell, LM et el, The role of beta-arrestins in the termination and transduction of G-protein-coupled receptor signals. J. Cell Sci. 2002, 115(Pt 3), 455–465; [4] Ranjan R et el, Novel Structural Insights into GPCR-β-Arrestin Interaction and Signaling. Trends Cell Biol. 2017, 27(11), 851–862; [5] Lane JR et el, A kinetic view of GPCR allostery and biased agonism. Nat. Chem. Biol. 2017, 13(9), 929–937; [6] Van Zundert GCP et el, The HADDOCK2.2 Web Server: User-Friendly Integrative Modeling of Biomolecular Complexes. J. Mol. Biol. 2016, 428(4), 720–725; [7] Rasmussen SGF et el, Crystal structure of the β 2 adrenergic receptor-Gs protein complex. Nature 2011, 477(7366), 549–557; [8] Kang Y et el, Crystal structure of rhodopsin bound to arrestin determined by femtosecond X-ray laser. Crystal structure of rhodopsin bound to arrestin determined by femtosecond X-ray laser, doi: 10.1038/nature14656; [9] Morris GM et el, AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. J. Comput. Chem. 2009, 30(16), 2785–2791; [10] Durrant JD et el, A novel algorithm for ligand-binding characterization. J. Mol. Graph. Model. 2011, 29 (6), 888–893; [11] Lomize MA et el, OPM database and PPM web server: Resources for positioning of proteins in membranes. Nucleic Acids Res. 2012, 40(D1), D370–D376; [12] Jo S et el, CHARMM-GUI: A web-based graphical user interface for CHARMM. J. Comput. Chem. 2008, 29(11), 1859–1865; [13] Wu EL et el, CHARMM-GUI membrane builder toward realistic biological membrane simulations. Journal of Computational Chemistry. October 2014, 1997–2004; [14] Abraham MJ et el, Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. SoftwareX 2015, 1–2, 19–25; [15] Schneider S et el, The Dynamic Process of Drug–GPCR Binding at Either Orthosteric or Allosteric Sites Evaluated by Metadynamics BT - G Protein-Coupled Receptors in Drug Discovery: Methods and Protocols. Filizola M, Ed. Springer New York, 2015; 277–294; [16] Barducci A et el, Metadynamics. Wiley Interdiscip. Rev. Comput. Mol. Sci. 2011, 1(5), 826–843; [17] Masone D et el, H-bond network optimization in protein–protein complexes: Are all-atom force field scores enough? Proteins Struct. Funct. Bioinforma. 2011, 80 (3), 818–824; [18] Moreira IS et el, SpotOn: High Accuracy Identification of Protein-Protein Interface Hot-Spots. Sci. Rep. 2017, 7(1), 8007; [19] Shukla AK et el, Visualization of arrestin recruitment by a G-protein-coupled receptor. Nature 2014, 512, 218; 20] Sensoy O et el, Understanding the Differential Selectivity of Arrestins toward the Phosphorylation State of the Receptor. ACS Chem. Neurosci. 2016, 7(9), 1212–1224; [21] Peterson SM et el, Elucidation of G-protein and β-arrestin functional selectivity at the dopamine D2 receptor. Proc. Natl. Acad. Sci. 2015, 112(22), 7097 LP-7102; [22] Coley C et el, Effect of Multiple Serine/Alanine Mutations in the Transmembrane Spanning Region V of the D2 Dopamine Receptor on Ligand Binding. J. Neurochem. 2001, 74(1), 358–366.


Material Sciences

# EVIDENCE OF POLARON FORMATION IN $Nb_{12}O_{29}$

*D.B. Migas[1], N.V. Skorodumova[2]*

*[1]Belarusian State University of Informatics and Radioelectronics, Minsk, Belarus; [2]Department of Materials and Engineering, Royal Institute of Technology, Stockholm, Sweden.*

## Introduction

There are some substoichiometric phases of $NbO_{2.5-\delta}$ and one of them, namely $Nb_{12}O_{29}$, has "incongruous" properties [1]. It is revealed to be an itinerant antiferromagnet with Neel temperature of 12 K and, at the same time, it displays metallic temperature dependence of resistivity from 300 down to 0.3 K. Moreover, it is expected to be an insulator assuming that the formula $Nb_{12}O_{29}$ can be viewed as $Nb_2^{4+}Nb_{10}^{5+}O_{29}^{2-}$ where two and ten Nb cations are present in the $d^1$ and $d^0$ configurations, respectively, in order to accommodate charge needs of twenty nine O anions. Such an ionic model can promote polaron driven conductivity behavior in $Nb_{12}O_{29}$. However, no polaron formation was detected in experimental works [1]. One possible explanation to this issue could be the coexistence of delocalized and localized electrons originated from Nb cations in the $d^1$ configuration. In this work by means of *ab initio* calculations within the GGA and GGA+U approximations we make an attempt to shed some light on "incongruous" properties of $Nb_{12}O_{29}$.

## Methods

The structural optimization and density of states (DOS) calculations for all polimorphs of $Nb_{12}O_{29}$ have been performed by the first principles total energy projector-augmented wave method (code VASP) [2] with plane wave basis set. Exchange and correlation potentials were included using the generalized gradient approximation (GGA) of Perdew-Burke-Ernzerhof. We set the energy cutoff at 500 eV. The convergence in total energy (less than 0.001 eV per formula unit) with respect to a number of k-points was checked to be sufficient if we used the 7×7×7 mesh of Monkhorst-Park points. DOSes were calculated by the tetrahedron method with Blöchl corrections. We also used the GGA+U approach to take into account the on-site d–d Coulomb interaction.

## Results

The calculated DOSs are shown in Figure 1. If the NM state is treated by the GGA approach, $Nb_{12}O_{29}$ has turned out to be a metal or degenerate semiconductor (see Fig.1 *a*). Spin polarized calculations provided a slight shift in energy for peaks in DOSs for different spin polarizations, which is especially pronounced near the Fermi level (Fig.1 *b*). The shape of DOSs in the case of the AFM state is identical for different spin polarization (Fig.1 *c*). In the case of the GGA+U approach the shift in energy of DOS peaks for different spin polarizations near the Fermi level progressively increases with the U value for Nb-d electrons (Fig.1 *d*, *e*, *f*) manifesting the appearance of semiconducting properties. At U=5.5 in the gap region a separate band is formed by splitting off from the conduction band (Fig.1 *f*). The separate band is mainly composed of d states of some Nb atoms with a slight admixture of O-p states, as can be clearly seen in Fig.1 *g*. For the AFM state the GGA+U approach does not provide any shift of DOS peaks for different spin polarizations relatively to each other, and the evolution of the separate band formation for both spin polarizations can be also easily traced (Figure 1 *h*, *i*, *j*). The formation of a separate band in the gap region in $Nb_{12}O_{29}$ is found to be accompanied by polaronic lattice distortion.

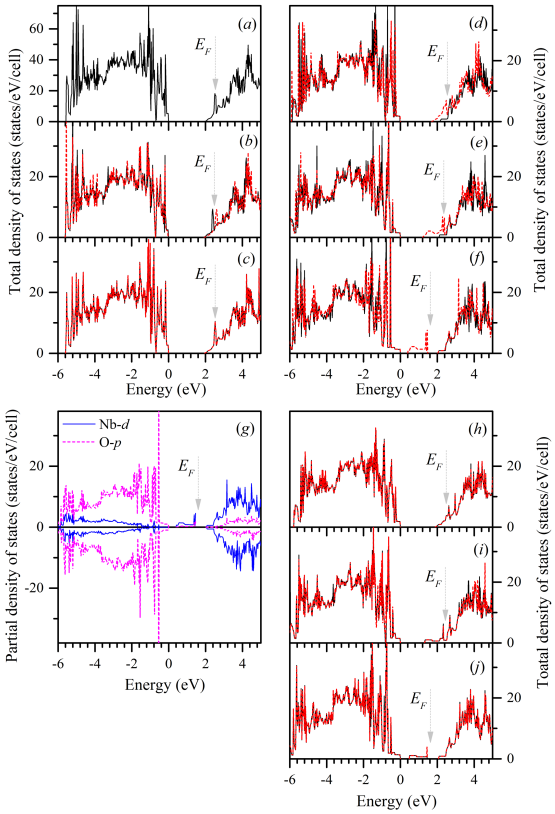



Figure 1 - Total (*a* – *f*, *h* – *j*) and projected (*g*) density of states (ev/states/cell) of $Nb_{12}O_{29}$ for different spin polarizations

## Conclusions

By using the GGA+U approach it is possible to describe localized states in the gap region of Nb12O29 due to the polaron formation.

## References

[1] Nico C et al, Prog. Mater. Sci. 80,1, 2016; [2] Kresse G et al, Phys. Rev. B 49, 14251, 1994.

## Acknowledgements

The work has been performed under the Project HPC-EUROPA3 (#HPC17ZYU8G), with the support of the EC Research Innovation Action under the H2020 Programme; in particular, the author gratefully acknowledges the support of Professor N. Skorodumova from KTH and the computer resources and technical support provided by the PDC Center for High Performance Computing at KTH.

# OPTICAL PROPERTIES OF ULTRA-THIN MONOELEMENTAL SEMICONDUCTORS

*M. Muževi[1], I. Lukačevi[1], N. de Leeuw[2]*

*[1]Department of Physics, Josip Juraj Strossmayer University of Osijek, Osijek, Croatia; [2]School of Chemistry, Cardiff University, Cardiff, United Kingdom.*

## Introduction

Since the discovery of graphene, there has been an explosion in 2D material research [1]. Other monoelemental 2D materials were predicted and synthesized – such as silicene, germanene, phosphorene, but there still exists a search for 2D forms of other elements. We study the potential 2D structures of elemental antimony, indium and aluminium.

One possible application of 2D materials is in optoelectronic devices. If such use is to be explored, their optical properties and interaction with light are of prime interest and importance [2]. Once known, natural modulation of optical properties of ultra-thin materials can be exploited by the variation of the number of layers.

## Methods

First-principles calculations were performed based on the density functional theory (DFT) [3, 4]. For the theoretical simulations we used the QuantumEspresso code which solves Kohn-Sham equations in a planewave basis set [5]. To calculate the optical properties, we use the random phase approximation [6]. Plane-wave energy cut-off was set between 60 and 80 Ry, depending on the structure and the type of element. Monkhorst- Pack k-point meshes between 8´8´1 and 14´14´1, depending on the structure and type of element, were used to approximate the integration over the first Brillouin zone during the geometry optimizations of monolayers. All atoms and crystal unit cells have been fully relaxed until residual forces and stresses were converged to 0.05 eV/Å and 0.5 kbar, respectively. Total energy convergence criterion was 10-3 eV/atom.

## Results

Among the studied elements, antimony proved to have three energetically stable forms – planar, buckled and puckered. Both indium and aluminium proved to have two stable forms – planar and triangular. Although stable in regard to forces acting on the atoms, further stability of these structures is to be explored using different methods.

For all structures, we calculated the electronic band structure, dielectric function, DOS and charge density. From the dielectric function we calculated absorption, reflection and energy loss spectrum, as well as extinction coefficient and refractive index. Example of calculated properties can be seen in Figures 1 and 2.

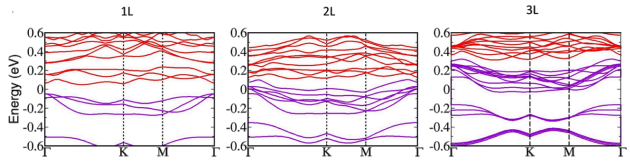



Figure 1 – Electron band structure of 1, 2 and 3 layers of buckled antimonene. Purple colour are valence bands and red colour are conduction bands. Transition from semiconductor to semimetal can clearly be seen.

We also studied the possible modulation of the optical properties with respect to the number of layers. An example can be seen in Figure 2. What we found out, in all cases, is clear and distinct changes with the increase of the number of layers – e.g. shifts of the absorption edge towards the visible spectrum.

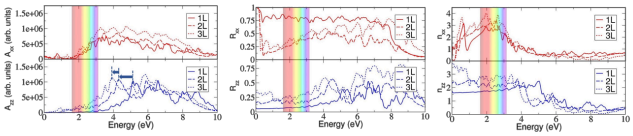



Figure 2 – Absorption spectrum, reflectivity and refractive index of 1, 2 and 3 layers of buckled antimonene. Coloured insert is the energy region of the visible spectrum.

## Conclusions

Using ab-initio simulations we modelled the possible stable 2D allotropes of elemental antimony, indium and aluminium. We simulated the characteristics of their electronic and optical properties. Also, we investigated how their electronic and optical properties depend on the number of layers of the proposed 2D materials and found they can be effectively modulated toward improved functionability.

## References

[1] Novoselov KS et al, Nature, 306:666-669, 2004; [2] Zhou X et al, Adv. Funct. Mater., 28(170658), 2018; [3] Hohenberg P et al, Phys. Rev., 136(B864), 1964; [4] Kohn W et al, Phys. Rev., 140(A1133), 1965; [5] Gianozzi P et al, J. Phys. Condens. Matter., 21 (395592), 2009; [6] Gajdoš M et al, Phys. Rev. B, 73(045112), 2006.

## Acknowledgements

The work has been performed under the Project HPCEUROPA3 (INFRAIA-2016-1-730897), with the support of the EC Research Innovation Action under the H2020 Programme; in particular, the author gratefully acknowledges the support of prof. Nora de Leeuw of Cardiff University and the computer resources and technical support provided by EPCC and HPC Wales.

# CHEMICAL ORDERING IN AGPT NANOALLOYS: STRUCTURE

*S. Olobardi*

*Dipartimento di Scienze Chimiche e Farmaceutiche, Università di Trieste, Trieste, Italy; Departament de Quimica Fisica and Institut de Quımica Teorica i Computacional (IQTCUB), Universitat de Barcelona, Barcelona, Spain.*

## Introduction

Nanoparticles (NPs) composed of atoms of more than one metal are often referred to as nanoalloys. They represent a lively research subject, thanks to their usage in catalysis, magnetism, optics, nanomedicine, etc. Properties of nanoalloys strongly depend not only on their geometric structure and size, but also on their particular composition. The latter makes nanoalloys much more tunable for tailored applications compared to monometallic particles. The project is focused on the study of three nanoparticles (NPs) of a PtAg nanoalloy with different size and local shape, although they all belong to the truncated octahedral structural shape family. These NPs contain 116, 140, and 201 atoms. For each one an Ag-rich situation has been studied, including both 1:3 and 1:6 Pt:Ag concentration ratios.

## Methods

The optimization procedure for obtain the chemical ordering of these bimetallic nanoalloys is the method for developing lattice models that aims at transferring ab initio or DFT level of accuracy to NP sizes: this is called the topological energy method (denoted TOP)[4,5]. This method is based on the analysis of the energy related to topological degrees of freedom: these take into account the formation of heteroatomic bonds and the different coordination of atoms in different positions of a certain NP. According to the TOP method, the energy of a given nanoparticle is approximated by the topological energy. In the case of a bimetallic NP of a chosen shape, composed of x atoms of A and y atoms of B, it depends only on the mutual positions of A and B atoms in a predetermined crystalline lattice. The parameters in the energy expression for each NP size and composition are derived by a rigorous fitting procedure based on energies from a limited set of density functional (electronic structure) calculations for NPs of the same size and composition. In the equation of E(TOP), we call $\varepsilon_i$ the energetic parameters associated with each degree of freedom considered in the topological energy. They will be referred to as descriptors. In contrast to parameters in many empirical methods, each descriptor $\varepsilon_i$ has a clear physical meaning. From a starting NP structure (the first guess for the chemical ordering) and an initial guess set of $\varepsilon_i$ descriptors, the program generates many NP structures with different chemical ordering via Monte-Carlo simulation. The simulation is carried out in steps, with simultaneous exchanges of n random atoms of A with n random B atoms. The code that performs Monte-Carlo simulation is written in a way that whenever it finds a structure with lower energy than the previous calculated ones, the geometry of this new structure is recorded in a file. As a result, a series of new XYZ files are obtained after the MC simulation, and some of these structures are selected for carry out DFT optimization calculations. 14 structures for each NP have been selected, both low-energy and high-energy. Then the descriptors are determine by fitting the E(TOP) structural parameters values to the total energies E(DFT) of these optimized NP structures via multiple linear regression using an adequate statistic software. In this study the free statistics software Wessa has been used. As a result of the fitting procedure, a new set of descriptor values is obtained, and the new descriptor values are used to re-optimize the chemical ordering in the nanoparticle. The Monte-Carlo simulation has to be repeated with the new descriptor set and the 10 low-energy structures are selected and again optimized at the DFT level for test E(TOP)'s precision (i.e. the precision of the complete topological expression themselves), δ. This can be estimated by calculating twice the residual standard deviation (RSD) between the calculated E(TOP) and E(DFT) energies. We have to determine also the trueness or accuracy of the topological energies (ΔE). If the precision is not satisfactory, these 10 structures have to be included in the fitting set, and we restart repeating the process. If the previous requirement are filled, we proceed calculating the accuracies of the individual descriptors. The electronic structure calculations were performed using the periodic plane-wave code VASP [6]. The code employs the PBE exchange-correlation functional [7] which was found to be one of the most appropriate among common functionals to describe transition metals [8]. The interaction between valence and core electrons was treated within the projector augmented wave (PEW) approach. In order to moderate the computational cost, proper energy cutoff of plane-wave basis sets was used. The one-electron levels were smeared by 0.1 eV using the first-order method of Methfessel and Paxton [9], and the converged energies were extrapolated to the zero smearing. All calculations were performed only at the gamma-point in the reciprocal space, and the full set of atoms was allowed to relax during the geometry optimization. To reduce computational costs, the minimal separation between NPs has to be larger than 0.7 nm, a typical safe value which guarantees negligible interaction between adjacent NPs [10].

## Results

For the PtAg nanoparticle with 116 atoms, $Pt_{29}Ag_{87}$ and $Pt_{17}Ag_{99}$ structures were considered respectively with proportions 1:3 and 1:6. After the descriptors $\varepsilon_i$ of the E(TOP) expressions were obtained also the 1-number of structures used for the fitting of the descriptors, $N_{FITTING}$, 2- precision of the corresponding topological energy expression δ, and 3-its accuracy ΔE were recorded. Regarding the $Pt_{29}Ag_{87}$ nanoparticle, although no exact criteria are defined to be met for the precision δ (but as a rule of thumb, $\delta \leq 200$ meV), even with a very poor $N_{FITTING}$ we are always below this general pre-established threshold of δ = 200 meV. In the case of $Pt_{17}Ag_{99}$, the situation is quite different: even including more and more low energy structures in the fitting procedure and repeating it several times in order to further optimize the descriptor values, the precision δ is always above the threshold and never goes below δ = 200 meV. Furthermore, the precision does not even approach linearly to the threshold, but shows an inconsistent trend. What we can surely notice is that an Ag-shell/Pt-core structure is present for both 1:3 and 1:6 proportions. All the corner, edge and terrace positions are occupied by silver atoms and the stability of Pt atoms in the inside of the nanoparticle is very high. The greater loss of energy that we can have is for Pt atoms in edge positions,

followed by the corner and the terrace ones. The situation is the same for the 140-atoms PtAg nanoparticle with again both the proportions 1:3 and 1:6 (the structures $Pt_{35}Ag_{105}$ and $Pt_{20}Ag_{120}$) and the nanoparticle with 201 atoms (the two structures 1:3 and 1:6 $Pt_{50}Ag_{151}$ and $Pt_{29}Ag_{172}$). The optimization of AgPt nanoparticles yields Ag atoms preferentially occupying positions with lower coordination numbers. This effect leads to the CoreShell-like structure of these lowest-energy AgPt homotops, which have the surface shell enriched with Ag and the core composed mostly of Pt.

## Conclusions

The energy gain due to the formation of heteroatomic bonds is rather small for these materials and plays a secondary role in the determination of the NP ordering. This reflect the experimentally known bulk immiscibility of Pt and Ag and all descriptors do not change drastically with decreasing Pt content. Furthermore, the PtAg "bonding" gets further disfavored for less Pt. This is an important finding related to the possibility of forming single surface atoms of one metals fully surrounded by atoms of another metal, nowadays called "single-atom catalysts".

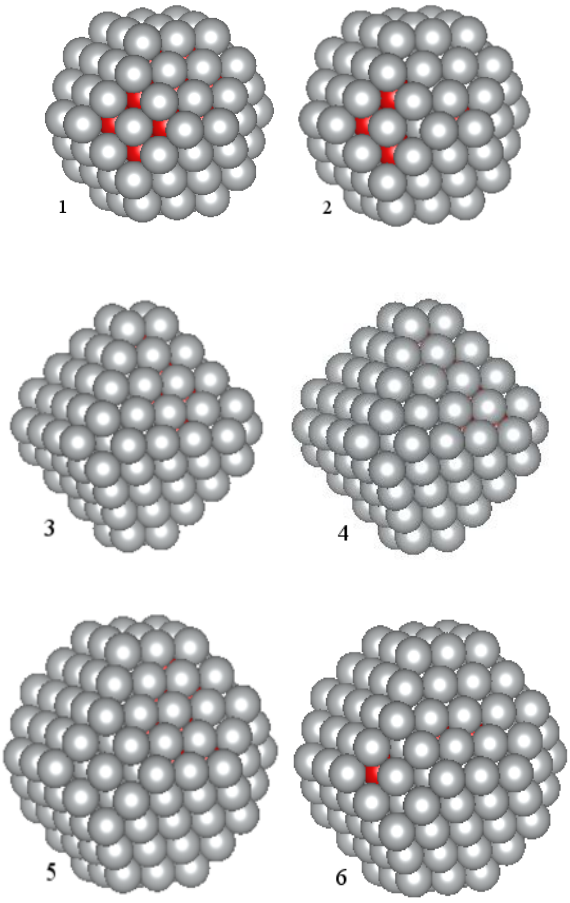



Figure 1 - Structures of different AgPt nanoparticles with optimized chemical ordering (1 and 2 are the 116-atom NPs $Pt_{29}Ag_{87}$ and $Pt_{17}Ag_{99}$; 3 and 4 are the 140-atom NPs $Pt_{35}Ag_{105}$ and $Pt_{20}Ag_{120}$; 5and 6 are the 201-atom NPs $Pt_{50}Ag_{151}$ and $Pt_{29}Ag_{172}$. Pt atoms are displayed as red spheres; Ag atoms are displayed as grey spheres.

## References

[1] Schaal M et al, Surf. Sci. 603, 690 (2009); [2] Cottancin E et al, Eur. Phys. J. D 24, 111 (2003); [3] Barcaro G et al, J. Phys. Chem. C, 118 (2014) 28101−28108; [4] Kozlov SM et al, Chem. Sci. 2015, 6, 3868-3880; [5] Kovács G et al, J. Phys. Chem. C 2017, 121, 10803−10808; [6] Kresse G et al, Phys. Rev. B: Condens. Matter Mater. Phys., 1996, 54, 11169–11186; [7] Perdew JP et al, Phys. Rev. Lett., 77 (1996), 3865; [8] Janthon P et al, J. Chem. Theory Comput., 9 (2013) 1631; [9] Methfessel M et al, Phys. Rev. B: Condens. Matter Mater. Phys., 1989,40, 3616–3621; [10] Viñes F et al, Chem. Int. Ed. 46 (2007) 7094.

## Acknowledgements

The work has been performed under the Project HPC-EUROPA3 (INFRAIA-2016-1-730897), with the support of the EC Research Innovation Action under the H2020 Programme; in particular, the author gratefully acknowledges the support of Departament de Quimica Fisica and Institut de Quımica Teorica i Computacional (IQTCUB), Universitat de Barcelona (Spain) and the computer resources and technical support provided by BSC (Barcelona Supercomputing Center).

# DFT Study of Stability of Gd - doped $BiFeO_3$

*V. Ribić[1], A. Dapčević[2], N. Skorodumova[3], D. Luković Golić[1], G. Branković[1]*

*[1]Institute for Multidisciplinary Research, University of Belgrade, Belgrade, Serbia; [2]Faculty of Technology and Metallurgy, University of Belgrade, Belgrade, Serbia; [3]KTH - Royal Institute of Technology, Stockholm, Sweden.*

## Introduction

Bismuth ferrite, $BiFeO_3$, is a promising single phase multiferroic material, with both ferroelectric and magnetic order parameters at room temperature. It has a rhombohedrally distorted perovskite structure, with lattice parameters $a$ = 5.5775(5) Å, $c$ = 13.8616(8) Å, $\alpha$ = 90°, $\gamma$ = 120°.[1] Despite of the room-temperature multiferroicity, the use of $BiFeO_3$ in practical devices has been limited because of a large leakage current induced by defects, low magnetic moment and a very weak magneto-electric coupling. The addition of even small amounts of trivalent rare-earth elements is known to stabilize the $BiFeO_3$ phase, and improve its physical properties. The addition of trivalent rare-earth element ions (La, Sm, Pr, Eu, Dy, and Gd) resulted in abrupt enhancement of magnetization. Based on experimental results, we have noticed that there is no continual linear dependence between a dopant concentration and certain functional and/or structural properties. We have found the sharp changes of properties for some dopant concentrations and thus the discontinuity from general trends. This study has proved that, in the case of Gd as a dopant, there is the periodical occupancy of specific site and, as a consequence, formation of a supercell structure.

## Models

The supercells for Gd-doped $BiFeO_3$ having 8.33 % of Bi replaced by Gd are constructed based on the Rietveld refinement of experimental X-ray powder pattern. These results reveal that the Gd-doped $BiFeO_3$ most likely crystallizes in 2x2x1 super-cell (s. g. #158, i.e. *P*3*c*1) with composition $Bi_{22}Gd_2Fe_{24}O_{72}$ and Gd occupying 2*a* position. The lowering in symmetry comparing to single cell characteristic for $BiFeO_3$ (s. g. #161, i.e. *R*3*c*) was necessary in order to find a particular site for Gd.

Within *P*3*c*1 space group Gd can occupy the following Wyckoff positions:

| | | |
|---|---|---|
| **2 c 3..** | **⅓, ⅓, z** | **⅔, ⅔, z + ½** |
| **2 b 3..** | **⅓, ⅔, z** | **⅓, ⅔, z + ½** |
| **2 a 3..** | **0, 0, z** | **0, 0, z + ½** |

## Methods

DFT calculations were employed to study the stability range of $BiFeO_3$ superlattices with respect to different dopants positions. Structural optimizations at the PBE level were carried out using Quantum Espresso. Core electrons were treated by pseudopotentials compiled in the optimized and tested Standard Solid State Pseudopotential database. All pseudopotentials were found to converge within a plane-wave cutoff energy of 65 Ry. Gamma-centered sampling was used. We used 0.02 Ry value of the Gaussian spreading for Brillouin-zone integration in metals and Marzari-Vanderbilt cold smearing. Relaxations were performed for nine distinct models that are representing different distributions of dopant.

## Results

DFT screening of Gd reveals the uniform distribution of Gd dopant in the BiFeO3 structure. Namely, the calculations showed that Gd has high preference to occupy the 2a Wyckoff position. Structural model with both Gd atoms in this position appears to has the lowest energy.

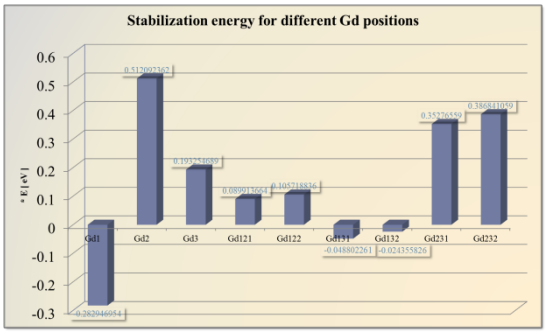



Figure 1 – Energies for 9 different models in eV relative to average value of calculated energies

The same trend was confirmed by calculations with varying volume of the cell.

## Conclusions

In this study we show that, in Gd-doped BiFeO3, dopant atoms are located at particular Wyckoff position in the BiFeO3 structure. This is a good start at proving that dopants are not randomly distributed over crystal structure but periodically occupy specific sites and, as a consequence, the modulated structures are formed in order to decrease the system energy.

## References

[1] Lukovic Golic D et al, Journal of the European Ceramic Society, 36: 1623–1631, 2016; [2] Khomchenko VA et al, Acta Materialia, 57: 5137–5145, 2009; [3] Gebhardt J et al, Journal of Physics: Conf. Series 921 012009, 2017.

## Acknowledgements

The work has been performed under the Project HPC-EUROPA3 (SNIC 2018/3-191), with the support of the EC Research Innovation Action under the H2020 Programme; the author gratefully acknowledges the support of KTH - Royal Institute of Technology, Department of Materials Science and Engineering, and the computer resources and technical support provided by PDC.

Physics

# UNDERSTANDING THE SURFACE CHEMISTRY OF AMORPHOUS CARBON USING MACHINE LEARNING AND DFT

*V.L. Deringer*

*Department of Engineering, University of Cambridge, Cambridge, United Kingdom.*

## Introduction

Carbon is a chemical element with fascinating structural complexity, and this is especially true for its amorphous (non-crystalline) forms. Amorphous carbon materials are also highly interesting for applications: dense, "diamond-like" or "tetrahedral amorphous" (ta) forms have long been used for coatings due to their attractive mechanical properties, and more recently ta-C has been suggested as an electrode material for detecting bio-molecules. Still, open questions remain with regard to its atomic-scale structure. Accurate and predictive computer simulations are now needed to better understand the surface structure of ta-C as well as its chemical functionalisation (most prominently, with hydrogen and various oxygen-based functional groups), which in turn determines the (electro-) chemical behaviour of ta-C films.

## Methods

We have recently introduced a machine-learning based interatomic potential for amorphous carbon [1] that makes use of the Gaussian approximation potential (GAP) framework [2]. GAP provides a method for accurately fitting the high-dimensional potential energy sur-face by "learning" quantum-mechanical (here, density-functional theory, DFT) reference data. Once developed and validated, a GAP can hence be used to carry out atomic-scale simulations at close-to-DFT quality but orders of magnitude faster. In an ongoing research collaboration with colleagues at Cambridge and Aalto Universities, we have been studying ta-C using atomistic GAP simulations. Most notably, we recently described the growth mechanism of ta-C films in large-scale, explicit simulations with hitherto inaccessible accuracy and in excellent agreement with experimental results [3].

However, the current version of the carbon GAP [1] has been "trained" for the pure element exclusively. The next challenge is therefore to extend the scope of simulations beyond that. We use a combination of different, complementary approaches: generating surface slab models in long GAP-driven molecular-dynamics (MD) simulations [4] and subsequently introducing functionalisation by density-functional methods. In the present case (Figure 1), we place oxygen atoms randomly on top of the slab models, and gradually heat the systems to observe preferred binding sites and reactivity.

## Results

Our results are in line with chemical intuition and previous experimental knowledge (see the original work [4] for a detailed discussion): at low temperature, epoxy groups 1 are abundant at the surfaces but diminish gradually due to the ring strain; keto groups 3 play the most important role at higher simulation temperatures, and are expected to be abundant at ta-C surfaces.

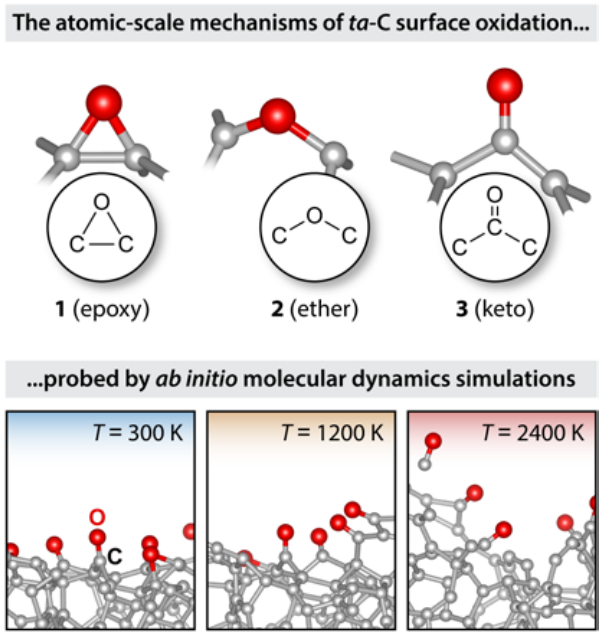



Figure 1 - Atomistic simulations of oxygen-based functional groups at the surface of a ta-C slab model. Top: Example fragments with chemical drawings; bottom: snapshots from ab initio MD simulations performed during the present visit. Adapted with permission from Ref. [4]. Copyright 2018 American Chemical Society.

## Conclusions

Machine learning methods combined with DFT are opening up exciting new research directions for the ac-curate modelling and understanding of complex functional materials. The work performed during this visit [4] as well as our overarching and ongoing collaboration [3–5] attest to this for the technologically important case of amorphous carbon materials.

## References

[1] Deringer V L & Csányi G, Phys. Rev. B, 95(09):094203, 2017; [2] Bartók AP et al, Phys. Rev. Lett. 104(13):136403, 2010; [3] Caro MA et al, Phys. Rev. Lett., 120(16):166101, 2018; [4] Deringer V L et al, Chem. Mater., in press; published online at DOI: 10.1021/acs.chemmater.8b02410 [5] Caro MA et al, Chem. Mater., in press; published online at DOI: 10.1021/acs.chemmater.8b03353.

## Acknowledgements

The work has been performed under the Project HPC-EUROPA3 (INFRAIA-2016-1-730897), with the support of the EC Research Innovation Action under the H2020 Programme; in particular, the author gratefully acknowledges the support of Dr Miguel Caro and Prof. Tomi Laurila (Department of Electrical Engineering and Automation at Aalto University, Finland) and the computer resources and technical support provided by CSC – IT Center for Science, Finland.

# DISCRETE MODEL FOR POPULATION DYNAMICS IN TWO-DIMENSIONAL COMPRESSIBLE TURBULENCE

*G. Guccione, R. Benzi, F. Toschi*

*Eindhoven University of Technology, Eindhoven, The Netherlands; Roma Tor Vergata University, Roma, Italy.*

## Introduction

The dynamics and genetics of populations of microorganisms in marine environments is a fascinating scientific topic at the crossroads of biology, statistical physics and fluid dynamics [1- 4]. The ocean covers the major part of the Earth's surface and it hosts millions of different living species. The field of population genetics can describe how different species outcompete one another over time.

The different species of micro-organisms can present biological differences affecting reproduction, feeding behaviors, resistance to disease and many others factors. Well-adapted individuals, which have inherited favourable characteristics, survive and grow faster than others, passing on the genes that make them successful. They are said to have a selective advantage. It is not possible to determinate a priori, which of two species taken into account one with selective advantage and one without, would be the dominant one in a competition. However, there is the possibility to calculate a probability.

For a system in absence of advection, Kimura in his work [7], derived a theoretical prediction on the fixation probability of one organism:

$$P_{fix} = \frac{1- e^{-sNf}}{1- e^{-sN}}$$

This formula describes the fixation probability for a species with selective advantage s in a population of size N that makes up an initial fraction f of all organisms. However, with regard to populations living in a fluid environment, it is largely unknown, both theoretically and experimentally, how population genetics change in the presence of turbulence.

In particular, effective compressibility can have a large effect. Fluid flows can be weakly compressible in many natural settings due to inertial effects. Also, effective compressibility can arise when passively traveling organisms living at a specific depth experience the convergence or divergence of water masses. These areas where two different water masses meet are called marine fronts, and they are characterized by high biological productivity. It is important to understand the variation in physical factors that a population can withstand and continue to thrive in these environments in order to support our ocean ecosystem.

It was recently found that in presence of weak flows compared to the no-flow case, on average, organisms born at sources acquire a significant advantage. It is shown that a turbulent flow can significantly diminish the effect of selective advantage on fixation probabilities. This project focuses on the enhancement or suppression of selective advantage in population genetics subject to advection of two dimensional compressible turbulence. To understand this question we started following a specific approach: We have used a discrete model, implementing first a one-dimensional code and then extended it to two-dimensional one.

The advantage of this approach is to allow us not to follow particle by particle but directly the set of particles that are present in a specific box composing the lattice. Because of this last statement, it is therefore possible to study a large number of microorganisms.

If the total number of particles per unit volume, N, is very large, a macroscopic description of the system in terms of continuous fields, e.g. density or concentration, is usually appropriate. A prototypical model for these reaction diffusion systems is the Fisher- Kolmogorov-Petrovskii-Piskunov (FKPP) population equation [5], [6] describing the space–time evolution of microbiological population densities due to the diffusion mechanism.

$$\partial_t c(x,t) = D\partial^2_{xx}\, c + \mu c(1-c)$$

Where c is the sum of the concentrations of the two species c = c1 + c2 ; $\mu$ is the rate of growth and D the diffusion.

In absence of noise the FKPP equation exhibits a traveling wave solution with a front speed given by $v$= = 2?$D\mu$ called Fisher velocity.

## Methods

In the first approach to solve the population dynamics, we were previously using a continuous space model [9-11]. In order to accommodate a greater number of organisms we adopt a discrete space model.

We implemented a one-dimensional code with a probabilistic diffusion: we insert a certain number of particles in a box, in which every individual has the possibility to move right, left or to remain in the same position. In this way we have discretized the diffusion.

Discreteness is accompanied by the noise so we had to change the logistic dynamics compared to the continue approach; To understand the impact of sources and sinks on fixation prior to tackling turbulence, we took a stationary velocity field with the sinusoidal form u(x)=A0sin(x-p/4) in a domain size 2p with periodic boundary conditions.

This kind of compressible flow leads to a reduction of the average population size when the flow strength is going to increase [9].

Particles are uniformly distributed, in particular we consider two species A and B; the first one is restricted to the source of the sine and the species B is homogeneous distributed throughout the field. The maximum of this field is small and it is comparable with the Fisher velocity of the system.

All single particles, for every timestep, are carried by the velocity field; if the particles shifted their position to another box the number of the neighbour cells will change.

We consider a model called "birthcoagulation process",

$$\begin{cases} A + A \xrightarrow{\chi} A \\ A \xrightarrow{\gamma} A + A \end{cases}$$

Where c indicates the death (coagulation) rate and g denotes the birth rate.

The abovementioned processes can be modelled box by box; we get the Fisher velocity value not with the expected velocity, but we obtain a lower value because of the

discretization. In 2003, Doring gave an estimation of how the Fisher velocity value has to be far from the ideal one [8]. The convergence of a discrete process is very slow, as the particle growing up in number the convergence line is approaching to unit value.

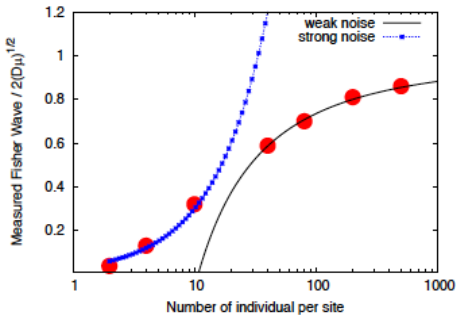



Figure 1 – Convergence trials algorithm: Fisher wave's behaviour vary with the number of individual per site.

A schematic drawing of the convergence is presented in Figure 1. There are shown two theoretical estimates; the blue and black lines get in the limit of respectively weak and strong noise. The red points are the results of our simulation to validate the convergence of the algorithm; They are perfectly in accordance with the theory. The convergence shows that already with 50 particles we are in weak noise regime; On the contrary with 10 particles we are in strong noise regime, where the Fisher velocity is about 0.3 times than the theoretical one.

We decided to simulate in strong noise regime.

## Results

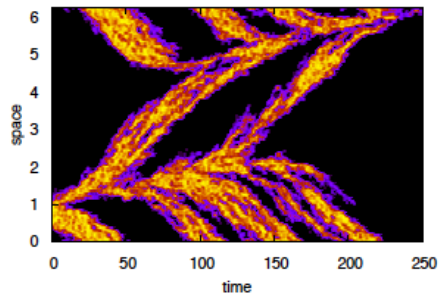



Figure 2 – Velocity alterations over the time; Coloured areas mean the part where competition between the two species takes place. Black regions represent the place in which just one species is present.

In figure 2 is shown how the velocity value of the function is changing over the time. Besides the Fisher wave it is also clear the effect of the displacement due to transport velocity.

The yellow lines represent the competition between the two species; initially are concentrate in the source, where the population A was placed. Subsequently this population expands itself till it wins on the other one, which dies.

The black region constitutes the part where there is just one species. To sum up, population A is initially concentrated in a small region of the entire space and subsequently is going to occupy the domain; population B, largely extended at the beginning, is going to die.

## Conclusions

We have presented a discrete model of both one and two dimensional system subject to compression with weak velocity. This new approach allowed us to study a large number of individuals implementing the process of diffusion and advection particle by particle and in addition the birth-coagulation process.

We could validate the convergence of the algorithm, the competition procedure and we could verify, taking a stationary flux, the correspondence of the simulations with the theoretical Kimura formula.

## References

[1] Wakita JI et al, (1994), J Phys. Soc. Japan, 63:1205 ; [2] Jacob E et al, (2004), Biofilms, 1:239 ; [3] Kimura M et al, (1964), Genetics, 49:561 ; [4] Korolev KS et al, (2009), Rev. Mod. Phys.82:1691-1718; [5] Fisher RA, The wave of advance of advantageous genes Annual Eugenics, 7 (1937), pp. 255-369 ; [6] Kolmogorov AN et al, A study of the diffusion equation with increase in the amount of substance and its application to a biology problem Bull. Univ. Moscow, Ser. Int. A, 1 (1937), pp. 1-16; [7] Kimura M., On the probability of fixation of mutant genes in a population, Genetics 47, 713 (1962) ; [8] Doering, CR et al, (2003). Interacting particles, the stochastic Fisher–Kolmogorov– Petrovsky–Piscounov equation, and duality. Physica A: Statistical Mechanics and its Applications, 325(1-2), 243-259 ; [9] Pigolotti S et al, (2012). Population genetics in compressible flows. Physical review letters, 108(12), 128102; [10] Benzi R et al, Physica (Amsterdam) 238D, 2003 (2009) ; [11] Perlekar P et al, Phys. Rev. Lett. 105, 144501 (2010).

## Acknowledgements

The work has been performed under the Project HPC-EUROPA3 (INFRAIA-2016-1- 730897), with the support of the EC Research Innovation Action under the H2020 Programme; in particular, the author gratefully acknowledges the support of Eindhoven University of Technology (TU/e), Department of Applied Physics and the computer resources and technical support provided by SURFsara.

# PARALLEL POWER DEPOSITION ON PLASMA FACING COMPONENTS

*L. Kos[1], T. Johnson[2]*

*[1] Faculty of Mechanical Engineering, University of Ljubljana, Ljubljana, Slovenia; [2]KTH, Department of Fusion Plasma Physics, Stockholm, Sweden.*

## Introduction

The SMITER code and its graphical user interface, based on SALOME platform, is a suite of object-oriented FORTRAN-95 modules designed to map profiles of heat scape-off layer (SOL) heat flux density flowing parallel to magnetic field lines onto the surfaces of Plasma Facing Components (PFCs). This requires that magnetic field lines on flux surfaces within the magnetic equilibrium be followed in 3D geometry until they intersect a solid surface, the geometry of which is generally obtained from a CAD model of the structure. In practice, field lines are actually followed backwards from the surface in question, with proper mapping of the heat flux profile specified in the free SOL taking into account magnetic flux expansion. This field line following must take into account the full neighbouring structures around the object of interest to ensure that field lines are not intersected locally by other solid surfaces.

## Methods

Parallelisation of the power calculation code "powcal" was analysed and divided into:

1) Running SMITER GUI interactively with VirtualGL acceleration on tegner.pdc.kth.se. Parallel execution gives user immediate access to CPU cores allocated interactively. Such use can run *powcal* in parallel on more than one node.

2) Interfacing SALOME graphics user environment with Job manager that submits jobs to PDC's HPC Tegner and Beskow under SLURM batch manager. Besides classical submit scripts there is also possibility to run remotely python batch jobs.

3) ParaView Catalyst is a solution that provides in-situ filedline traces for all or selected triangles on the target. Without Catalyst we could easily end up with thousands of VTK files. With Catalyst one can create graphics pipeline and save a “co-processor” script that lowers IO complexity and creates just field-lines of interests and the corresponding “connection lengths” to the user.

## Results

Field-line tracing parallelisation of the SMITER code in FORTRAN now achieved nearly linear speedup as shown in the table for a benchmark case from Ref. [1].

| #CPU | Time with blocking MPI comm. | Time with non-blocking MPI comm. | Time without MPI comm. |
|---|---|---|---|
| 1 | 170 s | 165 s | (168 s) |
| 2 | 86 s | 83 s | (84 s) |
| 4 | 43 s | 42 s | (42 s) |
| 8 | 21 s | 22 s | (21 s) |
| 16 | 12.2s | 12 s | (10 s) |
| 32 | 6.9 s | 6.5 s | (5.3 s) |
| 64 | 4.9 s | 4.2 s | (2.7 s) |

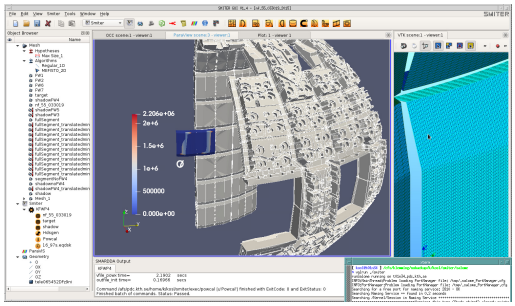


Figure 1 – SMITER GUI study running on compute node of tegner.pdc.hth.se showing in the centre window peak power deposition of 2.2 MW/m$^2$ on the target panel that is combined with the ITER blanket CAD geometry inside a ParaView window. Mesh detail of the target is show in the right window.

Timing in the table are given for one first-wall panel shown in the centre of Figure 1 consisting of high-density triangular mesh (256K triangles) and many millions of surrounding triangles of a lower density against which we calculate shadow on the target. For this, relatively small case, we compare evaluation times when using blocking and non-blocking strategy for MPI communication of the results into parent thread. Initial workload division is shown in the last column where no MPI communication were yet applied. When we compare times of the initial solution in the last column we see that the speedup achieved is nearly linear due to good workload balancing. Comparing the impact of the MPI communication against ideal (no communication times in parentheses) we see that the overall time is 1-2 seconds. The difference between non-blocking and blocking communication is small as we only collect results at the end whereas with non-blocking we don’t prescribe CPU priority.

## Conclusions

With parallelisation we resolved main obstacle for efficient power deposition of the PFCs and large cases are now possible. Impact of the communication for large scale cases could be further lowered with non-blocking send/collect in blocks while computing.

## References

[1] Kocan M et al, Nucl.Fusion. 55 033019 (2015), doi: 10.1088/00295515/55/3/033019

## Acknowledgements

The work has been performed under the Project HPC-EUROPA3 (INFRAIA-2016-1-730897), with the support of the EC Research Innovation Action under the H2020 Programme; in particular, the author gratefully acknowledges the support of Lilit Axner and the computer resources and technical support provided by KTH Royal Institute of Technology, PDC Center for High Performance Computing.

# MAGNETIC FIELD EVOLUTION IN DISK-CORONA SYSTEM

F. Nauman[1,2], D. Mitra[2]

[1]Niels Bohr Institute, Copenhagen, Denmark; [2]NORDITA, AlbaNova University Center, Stockholm, Sweden.

## Introduction

For many astrophysical systems with a gravitational source such as a young star or a compact object, excess matter from the ambient medium starts to fall on to the central object. This infalling material typically takes the shape of a disk like structure due to non-zero initial angular momentum. Many of these objects seem to be magnetized. What is the origin and morphology of these magnetic fields?

For this project, we used a layered structure where one layer was hydrodynamically forced while the other was unforced with low resistivity. This unforced highly conductive layer represents a highly magnetized region "corona". The goal of our project is to determine whether the magnetic fields that develop in the turbulent midplane due to the dynamo process can lead to the build up of magnetic fields in the corona and whether magnetic helicity from the forced turbulent region flows into the corona to feed the fields.

## Methods

We used the publicly available finite difference fluid and magnetohydrodynamics code, PENCIL, written in FORTRAN that has been tested to run efficiently over thousands of CPU cores. For our project, PENCIL code was run over 512 CPUs on Beskow for several hours (typically 8 hours).

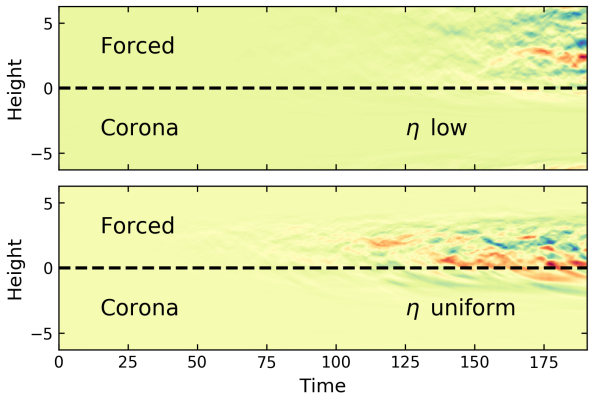



Figure 1 - The evolution of the horizontal magnetic field (Bx) of the two cases we studied. The forced region is where we force turbulence using a white noise in time and isotropic in space forcing and the corona region is unforced. In the top plot, we use a low resistivity in the halo while in the bottom case, the corona resistivity is the same as the forced turbulence region.

## Results

We show the evolution of the horizontal magnetic field in the disk and halo for two cases in the Figure 1. We use maximally helical forcing that is white noise in time and isotropic in space. The difference in the two cases displayed in Figure 1 is the resistivity in the corona: top: corona has a lower resistivity by a factor of 100 compared to the forced turbulence region while bottom: corona has the same resistivity as the forced region. In this simple experiment, we notice a difference in the diffusion across the disk-corona boundary. The diffusion seems to be more pronounced in the case where the resistivity in the corona is the same as the forced region. Eventually we want to study what happens when there is density stratification due to gravity [1], and whether helicity fluxes cause the growth of the field in the corona [2].

## Conclusions

We find that the magnetic field grows in the forced region in both cases but the diffusion of the field to the corona seems to be more pronounced in the case where we use uniform resistivity in the entire box. These results are encouraging as they suggest that even if the corona is not turbulent itself, it can still be fed from the turbulent regions. In this simple study, we did not include the effects of density stratification, shear, rotation that might lead to different results. Particularly, density stratification is likely to lead to stronger fields in the corona due to the buoyancy.

## References

[1] Mitra D et al, 2014, MNRAS, Volume 445, Issue 1, p.761-769; [2] Mitra D et al, 2010, Astronomische Nachrichten, Vol.331, Issue 1, p.130.

## Acknowledgements

The work has been performed under the Project HPC-EUROPA3 (INFRAIA-2016-1-730897), with the support of the EC Research Innovation Action under the H2020 Programme; in particular, the author gratefully acknowledges the support of NORDITA and the computer resources and technical support provided by PDC Center for High Performance Computing, KTH (Stockholm).

# IMPROVING PARALLEL EFFICIENCY OF EOF-LIBRARY – ELMER FEM AND OPENFOAM COUPLER

*J. Vencels*[1], *P. Råback*[2], *S. Ilvonen*[2]

[1]*University of Latvia, Riga, Latvia;* [2]*CSC – IT Center for Science Ltd, Espoo, Finland.*

## Introduction

EOF-Library is software that couples Elmer FEM [1] and OpenFOAM [2] simulation packages. It enables efficient internal field interpolation and communication between finite element and finite volume frameworks. Coupling is based on peer-to-peer strategy and uses Message Passing Interface which results in low latency, high data bandwidth and parallel scalability. There are known applications of EOF-Library in magnetohydrodynamics and plasma physics, but coupler also can be easily extended to other applications where volumetric fields are shared between Elmer and OpenFOAM.

One such illustrative example is electromagnetic levitation of liquid metal [3] shown in Figure 1.

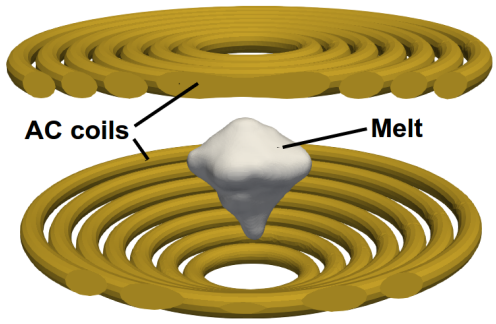



Figure 1 – Simulation of levitating liquid metal.

## Methods

Peer-to-peer (P2P) communication allows MPI processes from one software communicate directly with MPI processes from another one without any central entity (server). Communication pattern between processes is established during initialization phase.

In this work we implemented bounding box algorithm which allows faster search for P2P pairs and we did strong scaling test.

We took unit cube geometry and meshed it using tetrahedrons (13.1M elements / cells). To make Elmer and OpenFOAM meshes different while keeping the same number of tetrahedrons, we rotated one mesh around x-axis.

We ran our test on supercomputer Taito at CSC - IT Center for Science. Computing nodes had two 8 core Intel Xeon E5-2670 (Sandy Bridge) 2.6 GHz CPUs, 64 GB of memory and were connected with high-bandwidth, low-latency Infiniband FDR interconnect.

## Results

Strong scaling efficiency is shown in Figure 2. We are measuring and comparing wall time and duration it takes to transfer one scalar field in both directions. Field interpolation and communication (coupling time) takes 10.4s on 4 cores and 0.36s on 256 cores.

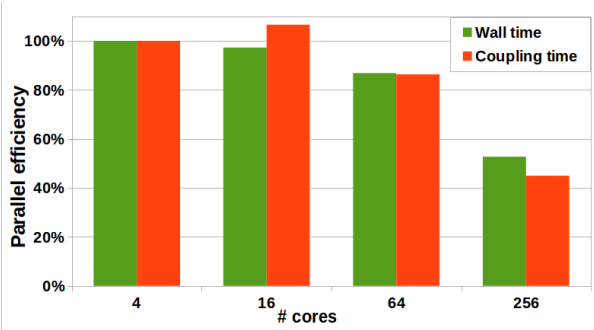



Figure 2 – Strong scaling efficiency compared as a wall time (green) and pure coupling time (red).

## Conclusions

EOF-Library uses peer-to-peer MPI-based coupling strategy to achieve high performance and scalability. It reuses native Elmer and OpenFOAM routines for search and interpolation tasks. Scaling test up to 256 cores (16 nodes) shows that EOF-Library makes negligible performance overhead for real simulations. In most realistic cases interpolation and field communication combined will take less than a second.

## References

[1] Keranen J et al, Efficient parallel 3-D computation of electrical machines with Elmer. (2015); [2] Weller H et al, A tensorial approach to computational continuum mechanics using object-oriented techniques. (1998); [3] Vencels J et al, Simulation of 3D MHD with free surface using open-source EOF-Library: levitating liquid metal in an alternating electromagnetic field. (2017).

## Acknowledgements

The work has been performed under the Project HPC-EUROPA3 (INFRAIA-2016-1-730897), with the support of the EC Research Innovation Action under the H2020 Programme; in particular, the author gratefully acknowledges the support of CSC – IT Center for Science and the computer resources and technical support provided by their HPC center.

# TOWARDS EFFICIENT SIMULATIONS OF SUSPENSIONS OF SOFT CAPSULES IN MULTI-COMPONENT LATTICE BOLTZMANN FLUIDS

*M. Wouters, J. Harting*

*Eindhoven University of Technology, Eindhoven, the Netherlands; Helmholtz Institute Erlangen-Nürnberg, Erlangen-Nürnberg, Germany.*

## Introduction

Compared to solid particles, soft particles are very efficient emulsifiers since the deformation at an interface allows for an additional reduction of the interfacial energy [1]. As a result, soft (micro-)particles are abundant in everyday life and play a key role in for example cosmetics, medical applications, paints, or the food industry. Unfortunately, the dynamics of a suspension of soft particles is difficult to describe analytically, while experimental methods can be invasive or struggle to extract the evolution over time of all relevant properties locally.

Therefore, an efficient model for the simulation of a suspension of soft capsules in a multi-component model can be very useful to increase the knowledge and improve the quality of many products which rely on soft particles. Hence we started to develop a model for the simulation of fluid-filled capsules in a multi-component fluid. In the current work the model only handles rigid capsules for simplicity, but deformability of the capsules has been kept in mind during development and can be easily added after the optimisation of the code for rigid capsules.

## Methods

For the underlying fluid-solver we start from an existing lattice Boltzmann method [2] with a pseudo-potential model [3] for simulating multiple fluids efficiently [4]. To model the capsules, we added infinitely thin boundaries which are discretised homogeneously via triangulation. The fluid-structure coupling is done via simple mid-grid bounce-back boundary conditions [5] with a leap-frog time integration scheme to dampen oscillations which occur due to the discrete movement of the boundary position when the particle moves.

The code is written in the Fortran programming language and is parallelised with the message parsing interface (MPI). The model is added as an extra module to the existing framework LB3D [2], which has been shown to have excellent scaling behaviour for large, complex simulations of suspensions [5,6,7].

## Results

The single-core performance of the original implementation was analysed with 'Intel VTune Amplifier' during the project. After mitigating the most significant bottlenecks, a speed-up of 89% was measured for the single-core performance of the code for a suspension with a volume fraction of 13%. Next to general restructuring and loop optimisations, a large part of the speed up resulted from favouring to reduce the number certain function calls by storing additional information calculated during the bounce-back stage for re-use in the next time step. The resulting increase in memory usage of ~3% is negligible since the code performance is not memory bound.

The self-imposed restrictions, which will allow us to extend the rigid structure into a deformable boundary, complicated the parallelisation since more information has to be communicated between CPU's, especially when a particle spans multiple CPU domains. After parallelisation, a good (weak-)scaling behaviour with less than 5% increase in computation time was achieved, as shown in figure 1. Each CPU was given a subdomain with $32^3$ fluid nodes and a volume fraction of ~13%.

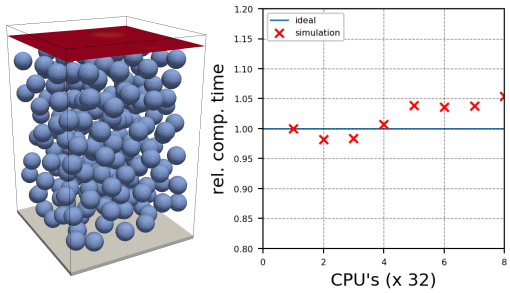



Figure 1 – (left) Early stage snapshot of a drying suspension of rigid capsules. The fluid-fluid interface is shown in red, and the substrate in grey. (right) Weak-scaling behaviour up to 256 cores, where each CPU handles $32^3$ fluid nodes.

## Conclusions

A new implementation to model capsules in a multi-component flow were both optimised and parallelised. The single-core performance nearly doubled, and the (weak-)scaling test showed increases below 5% for systems up to 256 CPU's. The extension of the model to handle deformable boundaries rather than rigid boundaries was kept in mind during this work, which mitigates the effort to extend the model to deformable boundaries in a multi-component fluid in future work.

## References

[1] Mehrabian H et al, *Soft Matter*, 12(4):1062-1073, 2016; [2] Hessling D et al, *J. of Chem. Phys.*, 145(5):054111, 2017; [3] Shan X et al, *J. Stat. Phys.*, 81:379, 1995; [4] Harting J et al, *NIC Symposium 2012: Proceedings; 25 Years HLRZ*, 2012; [5] Ladd A, *J. Fluid. Mech.*, 271:285-309,1994; [6] Harting, J et al, *Phil.Trans.Roy.Soc. A*, 363(1833):1895-1915,2005; [7] Aouane, O et al, *High Performance Computing in Sc. and Eng.'17,* Springer, 2018

## Acknowledgements

The work has been performed under the Project HPC-EUROPA3 (INFRAIA-2016-1-730897), with the support of the EC Research Innovation Action under the H2020 Programme; in particular, the author gratefully acknowledges the support of Manuel Zellhöfer and Othmane Aouane from the Helmholtz Institute Erlangen-Nürnberg and the computer resources and technical support provided by HLRS.