PRACE application form   **12th PRACE Project Access Call**

Detailed Project Document

Project name	**AMPEX-P**: SMALL MOLECULES AB INITIO ENERGY TRANSFER IN PLASMAS
Research field	Small molecules molecular dynamics

Project leader

Title	Dr
Last name	Esposito
First name	Fabrizio
Organisation name	CNR-NANOTEC
Department	Consiglio Nazionale delle Ricerche
Group	
Country	Italy

Preamble. The emphasis of PRACE on community proposals (for which in some recent calls a percentage of computing time was reserved) stimulated our Chemistry, Molecular and Material Sciences and Technologies (CMMST) Virtual Research Community (VRC) (see A. Laganà, G. Sipos, R. McLennan, "Towards a CMMST VRC, (VIRT&L-COMM.2.2012.2, http://www.hpc.unipg.it/ojs/index.php/virtlcomm/article/view/40/38)) to assemble a coordinated plan of submission under the name "AMPEX" (AB-INITIO MOLECULAR PROCESSES EFFICIENCY AS COLLABORATIVE SERVICE). The spirit of AMPEX leverages on a Virtual Community approach to complex computational simulations. In such approach highly demanding requests of both parallel computing runs and handling of large amounts of data are addressed to PRACE (from proper user interfaces built on interoperable and standardised components and services of distributed resources) for the part related to the use of highly parallelized packages. This enables effective collaboration between developers (f.e. application developers) and domain researchers based on the sharing of methodologies (models, algorithms and applications) and of data inside the same scientific application and across the different ones. In this way, in fact, the most appropriate machines for different types of software solutions will be used when creating applications and transforming them into production tools for innovative technologies. This implies that huge amounts of data will be transferred and stored during initial, intermediate and final stages of the elaboration allowing the creation of large shared repositories of efficiency parameters for Molecular processes. Three community proposals for the following fields of applications are considered:
AMPEX-A: SMALL MOLECULES AB INITIO REACTIVE EFFICIENCY IN ASTROCHEMISTRY (Toulouse, Perugia, Chieti, Vitoria) Biennial. Packages: NWCHEM, MOLPRO, VENUS, DL_POLY, BERTHA, RWAVEPR (in-house quantum reactive TD scattering code for atom-diatom reactions), MIXQC (in-house mixed quantum-classical code for reactive and inelastic scattering in diatom-diatom collisions); see related proposal.

AMPEX-P: SMALL MOLECULES AB INITIO ENERGY TRANSFER IN PLASMAS (Bari, Milano) Biennial. Packages: QCT4M (in-house QCT code), QUANTUM ESPRESSO. See this proposal.

AMPEX-C: LARGE MOLECULES AB INITIO STRUCTURAL PROPERTIES IN CATALYST DESIGN (Aachen, Tubingen). Biennial. NWCHEM, GAUSSIAN (see related proposal).

Research Project

Plasma phenomena in non-equilibrium conditions are currently actively studied due to their wide use in technological applications, such as plasma etching, plasma deposition, material plasma treatment, gas lasers, negative and positive ion-beam sources. Modelling is very helpful in the study of plasma [1], because of its ability to reveal details impossible or very difficult to access in the experimental approach. Low gas temperature plasma phenomena in non-equilibrium conditions are theoretically studied by means of coupled kinetic Boltzmann transport equations (BTE) for the different components of the plasma system: electrons, ions, atoms and molecules. To solve BTE, deterministic (state-to-state molecular dynamics) and stochastic methods such as DSMC (Direct Simulation Monte Carlo) and PIC (Particle-in-Cell) are the different approaches used. Both are based on the knowledge of particle-particle interaction potentials or cross sections for collisional events in bulk and on boundaries.

DSMC is the standard simulation method for the non-equilibrium dilute gas flows met in high altitude hypersonic flight or gas microflows. It is based on the dynamics of binary molecular encounters. Due to the high computational cost of the method, molecular collisions are traditionally dealt with by phenomenological models, or by using pre-calculated cross sections. The availability of accurate potential energy surfaces (PES) makes it possible to develop high-fidelity hybrid methods combining (Quantum-)Classical Trajectories with DSMC [2]. Typically pre-collision states are created by DSMC, while post-collision states are obtained by CT.

The PIC principles are applied in the Hall-effect thruster (HET) model [3]. Successful applications of HETs onboard numerous satellites on one side, and the high power electric propulsion interplanetary missions on the other side, are stimulating interest for this kind of aerospace propulsor. The challenge nowadays is to increase its power, but the fundamental physical mechanisms on which HET is based are not completely clear, and this issue prevents any further improvement. For this reason a numerical study of high power HET is required. Also in this case the elementary processes data are the inputs of the HET model, which is based on the particle representation of plasma electrons and ions, tracked since their production by ionization, by emission from lateral surfaces and from cathode, till they reach lateral walls or anode/cathode.

Another tool of value is the fluctuation theory of plasma [4], which is an answer to open issues deriving from the difficulty of formulating adequate theoretical approaches and solving the related equations in special regimes. Particularly challenging is to study the spectral density of plasma fluctuations in the presence of neutrals and/or dust grains, and in this context the elementary processes become important. The simulation of realistic plasma environments can be greatly improved with this kind of studies; in particular, modifications in the spectra due to the presence of solid particulates have been proposed as a diagnostic of dust in laboratory and space plasmas. A better understanding of plasma correlations will also find applications in plasma source design (material processing, telecommunications), material design and nanotechnology (plasmonics).

Vibrational state-to-state kinetic models are required for a quantitative treatment of heat problems due to entering planetary atmospheres [5]. As input of the models, a large amount of detailed (ro)-vibrational transitions among the appropriate atomic/molecular species

(containing N and O for Earth re-entry, C and O for Mars) and in interaction with spacecraft surface, is required. The state-to-state air species vibrational kinetics is also useful in the detailed study of atmospheric pollutants.

The models presented here need detailed data of vibration–vibration (V–V), vibration–translation (V–T) energy exchange processes and dissociation/recombination in molecular collisions, as well as the gas-surface interaction. The aim of the present proposal is to massively calculate this kind of elementary processes cross sections and probabilities to be used by mesoscale models in order to study fundamental phenomena and real scale device for observable quantities. Accurate plasma modelling requires at least the vibrational resolution of the molecular processes involved, but molecular rotation deeply modifies both vibrational energy exchange and dissociation, therefore it should be included at least in an approximate way. Among all the possible processes to consider, special attention is devoted to the atom-molecule collisions, because of their efficiency in modifying the internal molecular state and in dissociating molecules, generally to a greater extent with respect to molecule-molecule collisions. The calculation of detailed dynamical data for atom-molecule collisions involving air species will be performed in this project. In particular the $N+N_2$ collisional system will be treated in great detail, using some recent PESs [6,7]. The first one is highly accurate, but very time consuming, while the second one (derived from a 4-body PES) is computationally faster, but probably not as accurate as the first PES. Extensive detailed and accurate calculations and comparisons on both PESs have not been performed, while they would be obviously of great help for the numerous air species containing models. The processes to consider will be dissociation and vibrational energy transfer in the temperature range from 2000 to 20000K, considering virtually the whole rovibrational ladder (about 10000 states, treated with suited approximations). The dynamics will be studied by means of the quasiclassical trajectory method (QCT), that has shown both good accuracy and computational efficiency on massive calculations with this collisional system in the studies of the present proposal's project leader and collaborators [8,9,10]. Also the hybrid (OpenMP+MPI) QCT code "QCT4M" has been entirely written from scratch by the same author, with the specific goal of massive calculations for modelling (QCT for Modelling is the acronym). A complete, accurate and detailed database is expected from these calculations. This database can be used virtually in all the modelling research activities presented here. Comparisons with experimental results will be possible, with an accurate assessment of the PESs adopted.

Another goal of this proposal is the computational study of the catalytic properties of surfaces. This can be studied in a very detailed way by the Molecular Dynamics (MD) simulations that enable to describe at atomistic level the elementary processes due to the interaction of atomic and molecular gaseous species with solid surfaces. These processes include atomic and molecular adsorption, elastic and inelastic scattering, atom recombination and molecular dissociation. In detail, the study of graphene/substrate (like SiO_2/Si) will be performed, due to the interest in electronics, as well as the study of the H and H_2 interactions with a surface of CsMo, for modelling of divertor in nuclear fusion devices. The study will be carried out using the semiclassical collisional method for heterogeneous reactions (see for example Refs [11-15]). The key role of complete study is represented by the determination of PES where the reaction occurs, that will be performed in the framework of the present project.

Two years are necessary to complete the present project. The reason is that the project is not only of computational nature. The partial results obtained will be compared with data of both experimental and theoretical origin (in this last case also from the past works of the project leader and his collaborators), and subsequent work will depend on the results of these comparisons. In the $N+N_2$ case, the actual performances of two different PESs will be compared, stressing not simply the computational aspects but also the reliability of the final result. In fact, PESs are very often tested considering only a very low vibrational excitation, therefore it is not uncommon to find flaws when a complete study like the present one considers the whole rovibrational ladder. This is also the reason for using more than one PES in this work. Concerning the surface interactions, the first year will be devoted to obtaining the PES for O,O2 on graphene/substrate, while at the beginning of second year the obtained PES will be tested in MD code and possibly some modifications or re-parametrization will be necessary. At the same time the building up of the PES for the second system will be started.

Resources Management

Concerning QCT atom-molecule collisions, the project will start with some preliminary calculations to optimize the PES codes (not specifically written for OpenMP treatment) and to fix the fundamental parameters of the calculation (discretization of impact parameter and collision energy, numerical accuracy compatible with reasonable computing time, maximum interaction range, etc.). Then a thorough evaluation of the two PESs [6,7] will be necessary, in order to select the optimum compromise between accuracy and computational speed for the massive calculations. Finally, the complete set of dynamical data will be calculated using only one of the two PESs.

Concerning surface interaction, the first three months will be devoted to the optimization of cluster or slab surface models considered. The next nine months will be devoted to determining the points as a function of the distance from the surface and as a function of the interatomic distance (in the case of the molecule), for at least two sites of incidence on the surface. These points, by means of a fitting procedure, will allow defining the PES to be used in the study of reaction dynamics. The same procedure will be followed in the second year to determine the PES for the other considered system.

Months	1-3	4-6	7-9	10-12	13-15	16-18	19-21	22-24
$N+N_2$ parameters	█							
$N+N_2$ (PES assessment)		█	█	█				
$N+N_2$(complete calculations)					█	█	█	█
O,O2/graphene-substrate	█	█	█	█				
H,H2/CsMo					█	█	█	█

Numerical methods (1 page)

The numerical method for quasiclassical trajectories has been written from scratch by the project leader [16], and then improved over the years by adding some computational capabilities, such as distribution over a wide computational grid and MPI parallelism. The last recent improvement has been the use of OpenMP threads in the MPI code, in order to fully exploit the power of the Fermi machine [17]. Batches of trajectories are distributed on the Fermi nodes by MPI calls, while each trajectory in the batch is computed by a single thread on the physical 16-core chip in the node. Some work is generally necessary in order to make the PES code thread-safe. The computational result, reported in the speedup evaluation section, is quite appreciable. Each trajectory step is accurately checked with a double integration with two different time steps. This allows the use of an adaptive time step, which in turn allows to reach the best compromise between accuracy and computational speed. This is particularly important for reacting chemical species having a metastable three-body state in the strong coupling region, as the case of $N+N_2$ using all the recent PESs. The calculation of cross sections instead of the more common rate coefficients makes things a bit harder: the discretization of collision energy is an issue in particular for processes with a threshold, considering that in a massive calculation it is difficult and/or inefficient to specialize the parameters for each transition. Increasing the number of discretization points is not recommendable, because of the inefficient calculations in energy regions where the trend is smooth. A solution has been proposed in [18], with a continuous distribution of collision energy of trajectories and a special interpolation by a polygonal based on a feature extraction procedure. In that case all the relevant features of the cross section are correctly extracted, with a computational load comparable to a quite coarse-grained discretization. Some experiments in this sense will be done in the first part of the project, because of the practical importance of the topic for massive calculations.

Quantum Espresso is a suite of computer codes for electronic-structure calculations and materials modelling, based on density-functional theory, plane waves, and pseudopotentials. It contains the two different core packages PWSCF (Plane-Wave Self-Consistent Field) and CP (Car-Parrinello) for the calculation of electronic-structure properties within Density-Functional Theory (DFT), using a Plane-Wave (PW) basis set and pseudopotentials. In this proposal we will use PWSCF package.

Explain why this project needs to run on a Tier-0 system

Scalability tests have been performed during past ISCRA-C projects on the QCT4M (Quasi-Classical Trajectories for Modelling) in-house software. Results of recent tests are shown below. The system used for tests is very near to the exact object of the proposal. In fact, $N+N_2(v=40,j=80)$ collisions have been computed on the faster of the two possible PESs proposed here [7]. The amount of resources required by this project is estimated in the last section of the present document, and together with the high efficiency shown here fully justify the choice of a Tier-0 system.

# Fermi nodes	absolute timing (s)	speedup
2	6036.37	1.0000
3	3017.44	2.0004

4	2017.8	2.9915
8	865.68	6.9730
16	405.92	14.8708
32	200.27	30.1411
64	95.49	63.21
128	47.33	127.53
256	23.64	255.34
512	11.89	507.68
1024	6.06	996.10

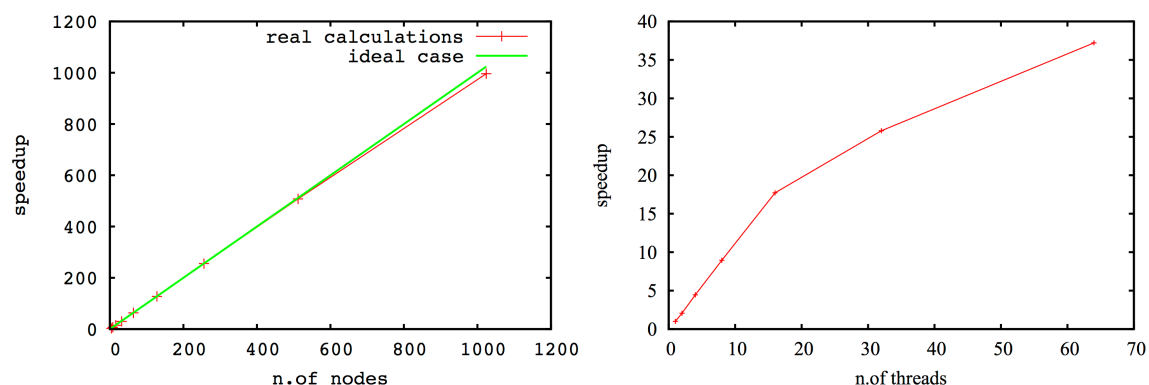Tab.1. QCT4M speedup as a function of the number of nodes (each node has 16 cores)

Fig.1.On the left panel: speedup obtained with the QCT4M quasiclassical code on the Fermi machine, using 64 threads, as a function of the number of nodes used (each node has 16 physical cores). On the right panel: speedup obtained with the same sw using 64 nodes (1024 cores) and an increasing number of threads. Over 16 threads the computational gain slows down, because of the thread concurrency on 16 physical cores, anyway the speedup is more than twice passing from 16 to 64 threads, at no cost.

# threads	absolute timing (s)	speedup
1	3553.45	1.00
2	1714.29	2.07
4	798.59	4.45
8	396.97	8.95
16	200.63	17.71
32	137.82	25.78
64	95.49	37.21

Tab.2. QCT4M speedup as a function of the thread number (with 64 nodes).

Quantum Expresso (QE) code is designed to exploit the highly structured hardware hierarchy of modern supercomputers with very good performances. Obviously, the size and nature of the specific application set natural limits to the maximum number of processors for which the performances of the various codes are expected to scale. Simulations on systems containing a few hundred atoms, as in the object of this proposal, are now quite standard, and excellent

scalability up to 4800 processors has been demonstrated by the code authors [15]. An example of scalability is reported in the following figure:
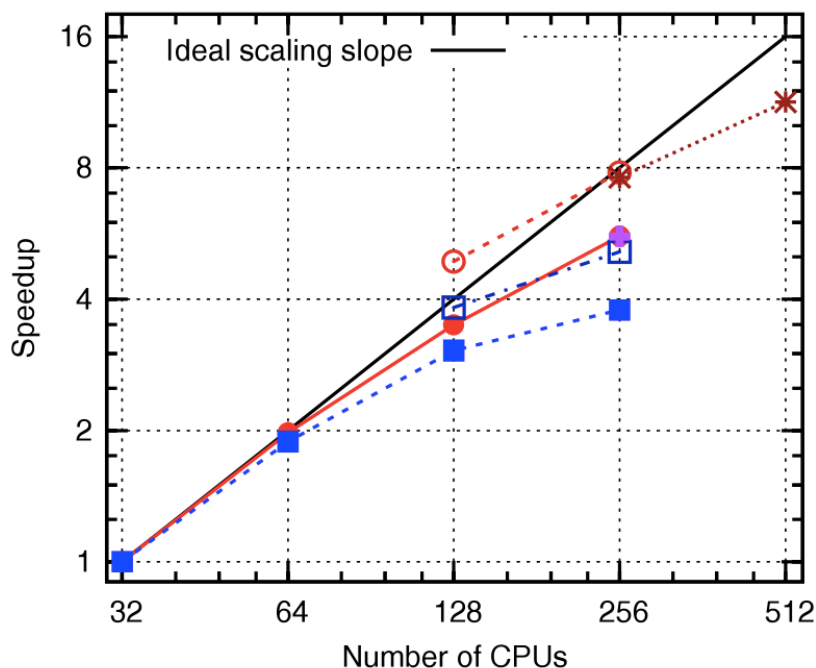

Fig.2. Scalability for medium-size calculations (CP code). Speedup with respect to 32 processors (right panel) as a function of the number of processors and for different numbers n_{task} of task groups, on an IBM BlueGene/P (BG/P) and on an SGI Altix. The system is a fragment of an Aβ-peptide in water containing 838 atoms and 2311 electrons in a $22.1 \times 22.9 \times 19.9$ °A3 cell, ultrasoft pseudopotentials, G point, 25 and 250 Ry cutoff for the orbitals and the charge density respectively [15].

On the Fermi machine the QE code shows a good scalability up to 65536 cores for a CP molecular dynamics calculation of a large system containing 1532 atoms.

Describe your experience of using HPC resources in the past and how you will manage using a Tier-0 system. What other experience do you and your team bring to this project?

Some experience about HPC resources has been gained by the project leader during his HPC Europa2 project in collaboration with Prof. Dmitri Sokolovski and Dr. Dario De Fazio at the University of Pais Vasco in Bilbao (Spain), during the last months of 2012. The title of the project is: "Classical and quantum complementarity in atom-diatom cross section calculations for plasma chemistry modeling". The MareNostrum as well as the Fermi machines were used during that project. The scope was to compare different approaches to reaction scattering of light species of astrochemical interest, with the goal of studying the computational feasibility of a complete rovibrational database in a collision energy range from ultracold to more than 10000K, necessary for astrochemical simulations. During that project the initial very simple parallel version of the QCT4M has been largely improved and adapted for a highly parallel

machine. The result of the comparison with very accurate quantum mechanical (QM) calculations on the same PES and machine has revealed a great accuracy and computational advantage of the quasiclassical code starting from collision energies as low as 0.1eV, where exact QM calculations start becoming very time consuming. This work gives strong support to the idea of using in a complementary way exact QM and QCT, in order to assemble complete databases in very large collision energy ranges.

In two subsequent ISCRA-C projects at CINECA [17] the QCT4M code has been refined by the project leader by using OpenMP threads. The final result is a hybrid code that distributes batches of trajectories to Fermi nodes using MPI calls, while trajectories in the batch are dynamically distributed among local threads in the node, with a clear gain visible in the speedup section of this document. This current version of the code has been tested against very accurate QM calculations for a light system, with excellent results that will be published on J.Phys.Chem. [19].

In the last years we performed calculations with Quantum Espresso on CINECA machines through different ISCRA calls. The first one was IsC10_MetInZeo in which we started the study on C-H bond breaking in the CH_4 interacting with ZSM-5zeolite. The study was continued by IsC17_MetZeo2 and in the last year by IsC29_USDinZeO, through which we obtained interesting results [14]. It was necessary to participate in three different calls because of the need for many calculations to characterize the reaction and the large number of electrons in the system.

Justify the number of core hours requested.

In the test runs presented in the scaling section of this document the statistical error on dissociation cross section is barely acceptable, a value *at least* five times larger of trajectories should used for massive production. Moreover, just 10 values of collision energies in the range 0.1-10 eV have been selected, a value largely insufficient for final results (at least 200 values are required in a range of 20 eV, or an equivalent continuous distribution).

Run type	# Runs	# Steps/Run	Walltime/Step	# CPU cores	Total core hours/Type Run
N+N2 parameters	200	1	0.5	1024	102400
N+N2 assessment	70	1	24	4096	6881280
N+N2 complete calc.	20	1	24	16384	7864320
opt-struct O,O2 graphene-substrate	10	100	0.05	1024	51200
iec O,O2 graphene-substrate	2000	30	0.10	1024	6144000
opt-struct H,H2 CsMo	10	100	0.05	1024	51200
iec H,H2 CsMo	2000	30	0.10	1024	6144000
					TOTAL 27238400

These test runs refer to just one initial state of the N_2 molecule: about 10000 rovibrational states should be considered, at least in an approximate way (one in 15, at least, in order to reliably reconstruct the whole set by interpolation).

Considering the computational time required by the tests, a minimum requirement of 2 million core hours are needed for one approximately complete rovibrational scan using the faster PES [7]. The other PES [6] is approximately one order of magnitude slower, and some experiments will be done to choose the most convenient PES to adopt for massive calculations. As a consequence a value of about 7 millions core hours in the first year and 8 millions in the second year will be sufficient.

Quantum Espresso runs will be of two types: 1) structure optimization (**opt-struct**) for clusters and slabs used as surface models; 2) interaction energy calculations (**iec**) for each considered configuration of gas-phase species and different sites on the surface. In particular, point 1) requires up to 100 points depending on the considered site while iec calculations imply the determination of minimum energy for each distance from the surface and in the case of molecule for each value of interatomic distance.

References
[1] M. Capitelli, R. Celiberto, G. Colonna, G. D'Ammando, O. De Pascale, P. Diomede, F. Esposito, C. Gorse, A. Laricchiuta, S. Longo, L. D. Pietanza and F. Taccogna, Plasma Physics and Controlled Fusion **53**, 124007 (2011).
[2] D. Bruno, A. Frezzotti and G. P. Ghiroldi, Phys. Fluids **27**, 057101 (2015)
[3] P. Tolias, S. Ratynskaia, A. Panarese, S. Longo and U. de Angelis, Journal of Plasma Physics 81, 905810314 (2015).
[4] F. Taccogna, P. Minelli, M. Capitelli, S. Longo, "Physics of Hall-effect thruster by particle model", AIP conference proceedings, 1501, 1390-1399, (2012)
[5] I. Armenise, F. Esposito, and M. Capitelli, Chem.Phys. **336**, 83 (2007).
[6] B.R.L. Galvão and A.J.C. Varandas, J.Phys.Chem.A **113**, 14424 (2009).
[7] J.D. Bender, P. Valentini, I. Nompelis, Y. Paukku, Z. Varga, D.G. Truhlar, T. Schwartzentruber, and G.V. Candler, J.Chem.Phys. **143**, 054304 (2015).
[8] F. Esposito, I. Armenise, and M. Capitelli, Chemical Physics **331**, 1 (2006).
[9] F. Esposito and M. Capitelli, Chem.Phys.Lett. **418**, 581 (2006).
[10] I. Wysong, S. Gimelshein, N. Gimelshein, W. McKeon, and F. Esposito, Physics of Fluids **24**, 042002 (2012).
[11] M. Rutigliano, C. Zazza, N. Sanna, A. Pieretti, G. Mancini, V. Barone and M. Cacciatore, J. Phys. Chem. A *113(*2009)*15366–15375*
[12] C.Zazza, M. Rutigliano, N. Sanna, V. Barone, M. Cacciatore, J. Phys. Chem. A 116 (2012) 1975−1983
[13] M. Rutigliano, P. Gamallo, R. Sayós, S. Orlandini, and M. Cacciatore Plasma Sources Science and Technology, 23 (2014) 045016
[14] Maria Rutigliano, Nico Sanna, Amedeo Palma, Computational and Theoretical Chemistry, 1074 (2015) 9-18
[15] P. Giannozzi et al., QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials, J. Phys.: Condens. Matter 21 (2009) 395502
[16] F. Esposito, Dinamica Quasiclassica di Processi Collisionali Inelastici e Reattivi in Sistemi $H + H_2$ e $N + N_2$ Rotovibrazionalmente Risolti, PhD Thesis, Università degli Studi di Bari, 1999.

[17] ISCRA-C projects HP10CIF3IN (2013), HP10C7RZR8 (2014)

[18] [1] F. Esposito, in *AIP Conference Proceedings*, edited by D.A. Levin, I.J. Wysong, and A.L. Garcia (2011), pp. 1357–1364.

[19] F.Esposito, C.M.Coppola, D. De Fazio, "Complementarity between quantum and classical mechanics in chemical modeling. The $H + HeH^+ \rightarrow H_2^+ + He$ reaction: a rigourous test for reaction dynamics methods ", to be published in J.Phys.Chem.A