

Italian National Agency for New Tecnologies,
Energy and Sustainable Economic Development

High Performance Computing on CRESCO infrastructure: research activities and results 2009-2010

Giugno 2011

This report was prepared by the New Projects and Web Applications Laboratory of ENEA's Technical Unit for Information Systems and ICT Developments (UTICT-PRA), from contributions provided by a selection of users of the CRESCO infrastructure.

Scientific Editor: *Delinda Piccinelli*, Enea, UTICT-PRA, Cr Frascati
Cover: *Amedeo Trolese*, Enea, UTICT-PRA, Cr Frascati

ISBN: 978-88-8286-242-8

INDEX

Preface

Nota sulla Utilizzazione del Sistema HPC CRESCO nell'Anno 2009.
G. Bracco, G. Guarnieri, S. Migliori, A. Santoro 7

Numerical Modelling of confined Particle Laden Flows.
F. Ambrosino, F.S. Marra 19

Integration of FLUKA Simulations for TOP-IMPLART
Treatment Planning System.
E. Basile D. M. Castelluccio, E. Cisbani, S. Frullani 27

Synchronization in Complex Biological Networks.
J. A. Almendral, R. Bajo, S. Boccaletti, J. M. Buldú, R. Gutiérrez, I. Leyva,
A. Navas, D. Papo, I. Sendiña-Nadal, M. Zanin 33

Biophysics at the atomic Scale.
L. Bongini 43

Meteorological and Air Quality Fields Production over Italy in the Frame
of the Minni Project.
G. Briganti, A. Cappelletti, L. Ciancarella, G. Pace, L. Vitali, G. Zanini 49

Simulation of chemical Reactions catalyzed by Transition Metals Complexes
and large-scale parallel Simulations of Biomembranes by Hybrid Particle-Field
Molecular Dynamics Technique.
L. Cavallo 59

Advances in LES of the HyShot II Scramjet Combustor.
D. Cecere, F. R. Picchia, N. M. Arcidiacono, E. Giacomazzi, F. Donato,
A. Ingenito, C. Bruno 65

Methodological Approach to study energetic and structural Properties on
nanostructured Cadmium Sulfide by using *ab initio* molecular Dynamics Simulation.
E. Burresi, M. Celino 73

Development of a Model for the numerical Simulation of Multiphase reacting
Flows by Large Eddy Simulation Technique.
F. Donato 77

Open FOAM computational Performances for large parallel distributed Simulations.
A. Funel, F. Ambrosino 85

Syngas turbulent Combustion Modelling and Simulation:
E. Giacomazzi, D. Cecere, F. Donato, F. R. Picchia, N. Arcidiacono 91

Materials for Hydrogen Storage: Numerical Simulation of Hydrogen Dynamics at a Mg-MgH_2 Interface using CPMD Code on CRESCO HCP.
S. Giusepponi, M. Celino 97

CFD Simulation of a Lean-burn Aero-engine Combustor with Low Pollutant Emission.
R. La Gala, F. Ambrosino, A. Funel S. Migliori,
P. Di Martino, S. Colantuoni, A. D'Anna 105

Crystalline and liquid Si_3N_4 Characterization by first-principles molecular Dynamics Simulations.
A. Mauri, M. Celino 111

Fullerene Derivatives as Electron Acceptors in Polymer solar Cells: a Quantum Chemical Study.
P. *Morvillo* 117

Numerical CFD Activities performed by AVIO.
F. Paglia 125

ARK3D – L'Infrastruttura ENEA-GRID per il 3D Remoto.
D. Abate, R. Ciavarella, G. Guarnieri, G. Furini, S. Migliori, S. Pierattini 133

Improving free Energy Perturbation Methods through Multiple State Sampling: the Histidine Tautomerism Case.
S. Marsili, P. Procacci 137

Ab- initio Calculations of excited State Properties of complex Systems.
M. Marsili, P. Gori, O. Pulci 145

A multiscale stochastic Volatility Model in mathematical Finance.
L. Fatone, F. Mariani, M. C. Recchioni, F. Zirilli 151

Exchange Flow through the Strait of Gibraltar as simulated by a non-Hydrostatic Model.
G. Sannino, L. Liberti, J. C. Sánchez-Garrido 157

A fast Method to determine the best High Performance Linpack (HPL) Parameters for a given Supercomputer.
G. Bracco, S. Migliori, S. Podda, S. Raia 163

PREFACE

This report contains a collection of papers presenting the main computational activities performed on CRESCO HPC system during its second year of operation, from the second half of year 2009 to the first half of year 2010.

This paper follows a similar report produced for the first year of operation 2008-2009 and now results from new domains as mathematical finance and medical physics add up to the previous list of fields as climate analysis, pollutant diffusion, nuclear technology, computational chemistry, computational fluid dynamics, plasma physics, biophysics and bioinformatics.

CRESCO (Computational center for research on Complex systems) project has been developed by ENEA in the framework of the Italian National Program for Research 2000-2006 and started its operation during 2008. CRESCO HPC system, located in Portici (NA) is now a facility with more than 3000 cores and about 300 TB raw data space, connected by a 2 Gbits link to the Italian academic network, managed by GARR. The system achieved rank 126 in the top 500 list of the most powerful HPC systems in the world at June 2008 and remained in the list until the end of year 2009. The facility is the main component of ENEA-GRID infrastructure which integrates all the ENEA computational resources in a single logical system following the computing GRID approach.

CRESCO HPC system has operated with high values of availability (for example 90% during 2009) and by a large number of users, some of them for more than 1 century of cpu elapsed time. The facility has provided computational services both to ENEA researchers and to external users, from academic and industrial communities, with a high utilization rate which almost saturates the available computer power. Some computational time has also been provided to the Virtual Organizations of the EGEE European project, as a fraction of the CRESCO infrastructure is accessible to EGEE by means of a gateway system developed by ENEA.

The success of the project assesses the role of ENEA in the HPC field and hopefully will trigger future enhancement of the facility so that ENEA researchers and their external collaborators can continue to rely on state-of-art computation resources, following the current fast evolution of the information technology.

CRESCO Team
Unità Tecnica Sviluppo sistemi per l'Informatica e l'ICT.

NOTA SULLA UTILIZZAZIONE DEL SISTEMA HPC CRESCO NELL'ANNO 2009

Giovanni Bracco[1], Guido Guarnieri[2], Silvio Migliori[3], Andrea Santoro[1].

[1]ENEA UTICT_HPC C. R: Frascati
[2]ENEA UTICT-HPC C. R. Portici (NA)
[3]ENEA UTICT Roma Sede

Introduzione

Questa nota riassume alcuni dati sul funzionamento del sistema HPC CRESCO [1] nell'anno 2009 sia dal punto di vista della disponibilità del sistema che del suo utilizzo effettivo da parte dell'utenza. Complessivamente nel corso dell'anno 2009 il sistema non è stato disponibile all'utenza per 38 giorni di cui 25 a seguito di fermi programmati e 13 per incidenti o malfunzionamenti. La disponibilità temporale del sistema, quindi, può essere valutata a circa il 90% nel corso dell'anno in oggetto.

Fermi Macchina Programmati

Nel 2009 ci sono stati due fermi programmati per interventi sull'ambiente software dell'infrastruttura:

1. **19/5/2009-3/6/2009** (16 giorni): migrazione del sistema operativo da RHEL5.1 a CentOS5.3. Il fermo ha riguardato l'aggiornamento dell'intero ambiente software inclusi i compilatori e le librerie.
2. **26/10/2009-3/11/2009** (9 giorni): migrazione per l'intero parco dei nodi di calcolo da indirizzi privati di rete a indirizzi pubblici per risolvere problemi di comunicazione su WAN. Installazione di una versione aggiornata kernel non afflitta da vulnerabilità.

In entrambi i periodi l'attività durante il fermo macchina per l'utenza ha incluso una fase di test del sistema per verificarne la funzionalità, effettuato utilizzando le utility standard di benchmark, come il codice HPL. Nei fermi macchina programmati la sottomissione di job utenti viene chiusa con 24 ore di anticipo rispetto alla data di inizio del fermo, tenendo conto che la durata massima temporale delle code di sottomissione standard su CRESCO è di 24 ore. In tali 24 ore però i job attivi possono arrivare a termine regolarmente.

Fermi Macchina dovuti a Incidenti o Malfunzionamenti

Le cause principali di malfunzionamento dell'infrastruttura nel corso dell'anno sono state problemi al sistema di condizionamento nel periodo estivo e malfunzionamenti della rete di interconnessione veloce InfiniBand (IB) [2], dal cui buon funzionamento dipende anche l'affidabilità del file system parallelo GPFS [3]. Non vengono qui riportati i malfunzionamenti della rete di connessione verso l'esterno del centro di calcolo di Portici nel quale il sistema HPC CRESCO è installato, perché in generale essi sono stati di breve durata e non hanno avuto impatto sull'operatività del sistema di calcolo. Gli eventuali problemi residui erano in generale transitori e non hanno provocato giornate intere di indisponibilità del sistema. I periodi significativi di indisponibilità per malfunzionamenti sono stati:

1. **26/7/2009-4/8/2009** (8 giorni): incidente dovuto a temperatura eccessiva. Il mancato funzionamento di 3 dei 5 chiller di condizionamento in sala calcolo ha costretto allo spegnimento di emergenza dell'intera infrastruttura e ad una successiva attività di test per eliminare i problemi ed i guasti causati dalle temperature eccessive ad alcuni dei componenti (ad es. fibre ottiche).
2. **15/9/2009-16/9/2009** (2 giorni): problemi IB/GPFS. Il cattivo funzionamento della rete IB, segnalato dagli utenti, ha richiesto lo spegnimento e la riaccensione dell'intera infrastruttura oltre al ripristino del corretto funzionamento di GPFS.
3. **6/10/2009-8/10/2009** (3 giorni): problemi IB/GPFS. Il cattivo funzionamento della rete IB, segnalato dagli utenti, ha richiesto di nuovo lo spegnimento e la riaccensione dell'intera infrastruttura, oltre al ripristino del corretto funzionamento di GPFS. In particolare sono stati individuati alcuni nodi aventi schede IB funzionanti in modo intermittente che causavano, di conseguenza, problemi all'intera infrastruttura.

Da notare che durante il periodo di chiusura estiva del centro ENEA di Portici e a valle dell'incidente verificatosi per il malfunzionamento del sistema di condizionamento della sala calcolo, l'operatività del sistema HPC CRESCO è stata limitata riducendo il numero di nodi di calcolo per il periodo 7/8/2009-24/8/2009. In tal modo si volevano limitare gli effetti di eventuali incidenti al sistema di condizionamento durante il periodo del maggior caldo estivo in condizioni di presidio limitato dell'infrastruttura. In tale periodo solo il 20% delle risorse di calcolo è stato lasciato a disposizione degli utenti.

Analisi dell'Utilizzo in Termini di Tempo di Orologio

L'analisi dell'utilizzo da parte degli utenti è stata fatta considerando i dati di accounting prodotti dalle utility contenute nel pacchetto psacct [4] tramite le quali vengono raccolti dati organizzati per giorno, mese ed anno da tutti i nodi sia di calcolo che di front-end. Il dato principale è quello del tempo di orologio totale usato da ogni utente su ognuno dei nodi del sistema . Ulteriori considerazioni sul tempo di CPU, sul numero di processi utilizzati ed altro includono anche qualche dato preso dal sistema di accounting di LSF [5], il gestore di risorse utilizzato sul sistema HPC CRESCO. Nella seguente *Tabella 1* vengono riassunti i dati globali di utilizzo del sistema HPC CRESCO nel corso del 2009.

Nprocess	RealTime(min)	RealTime(Anni)	CPUTime(min)	tCPU/Real	avM(kB)
508629075	1062529840.2	2021.56	631385782.52	0.59	2847008880

Tabella 1: ***Nprocess:*** *numero di processi lanciati,* ***RealTime:*** *tempo di orologio in minuti ed anni,* ***CPUTime:*** *tempo di CPU in minuti,* ***tCPU/Real:*** *rapporto tra tempo globale di CPU e tempo di orologio,* ***avM(kB):*** *utilizzo di memoria dei processi, mediato nel tempo.*

Il tempo d'orologio mostrato in colonna 3 è quello utilizzato dalle applicazioni e viene calcolato come il tempo di uso di un singolo core di CRESCO (nel corso del 2009 erano a disposizione in totale 2720 "core"), sommato sul numero totale di core utilizzati. ***Il dato mostrato vale 2021 anni di calcolo***, indicando un uso effettivo medio del sistema di circa il 75% (2021/2720=0.75). Tenuto conto dei periodi di chiusura del sistema, che riducono del 10% la disponibilità della macchina, il tempo effettivo di utilizzo totale è risultato essere di circa il 83 % (2021/(2720*0.9)=0.83). In colonna 5 della *Tabella 1* viene mostrato il valore globale del

rapporto tra tempo di CPU e tempo di orologio che è uguale al 59%. Il dettaglio di questo dato, che dipende dalla tipologia di applicazione e dal peso delle fasi di input/output, viene brevemente analizzato successivamente. In Tabella A della appendice vengono poi mostrati i dati disaggregati per utente. Nel 2009 sono state rilevate 261 diverse utenze delle quali 212 hanno usato il sistema per almeno 1 ora e la ***Tabella A*** si limita a mostrare queste ultime. Nella tabella gli utenti vengono identificati da un numero progressivo in modo da permettere l'analisi mantenendo un livello ragionevole di riservatezza. I quattro maggiori utilizzatori hanno utilizzato il sistema per più di un secolo ed in particolare il maggiore utilizzatore in assoluto ha totalizzato circa 3 secoli di tempo di calcolo. Gli utilizzatori che hanno usato almeno un anno di tempo di calcolo sono in totale 62.

Nella seguente *Tabella 2* i tempi di calcolo (sempre nell'accezione di tempo d'orologio) sono stati valutati raggruppando i dati per le aree tematiche a cui riferiscono i singoli utenti della ***Tabella A***.

AREA DI ATTIVITÀ	Tempo di calcolo (anni)	Frazione del totale	Istituzioni di appartenenza
Energia-combustione	621.87	0.31	ENEA
Materiali-innovativi	531.9	0.26	ENEA, INFN/RM, INFN/NA,UniFI,UniSaUniSS, Numonyx
Clima	357.57	0.18	ENEA,Ylichron
Nucleare-fissione	122.9	0.06	ENEA,ISS
Progetti-GRID	113.99	0.06	Progetto EGEE
Attività di supporto	103.11	0.05	ENEA
Ambiente	65.62	0.03	ENEA,AriaNet
Industria-aerospazio	32.77	0.02	Avio,AAPS,UniROMA1
Università	22.13	0.01	UniROMA1,CERI
Nucleare-fusione	19.99	0.01	ENEA,PoliTo
Biotecnologie	18.3	0.01	ENEA, CNR-ITB, Ylichron,CNR-ISA
Non definito	6.35	0	
Fusione-efda	4.69	0	ENEA,EFDA
Industria	0.37	0	NICE,CETMA
Totale	2021.56		

Tabella 2: *Tempo di calcolo (tempo di orologio) aggregato per l'area di attività degli utenti; per ognuna delle aree di attività vengono indicate nell'ultima colonna le istituzioni di provenienza degli utenti.*

I maggiori utilizzatori di CRESCO appartengono quindi alle tre aree Energia-Combustione, Materiali Innovativi e Clima, che nel loro insieme utilizzano circa il 75% del tempo di calcolo disponibile. La tabella riporta anche una lista approssimativa delle istituzioni di appartenenza degli utenti coinvolti nelle varie aree di attività.

Considerazioni relative ad altri Parametri di Utilizzo:

1. **Rapporto tempo di CPU e tempo di orologio**

In ***Tabella A*** alla colonna 5 viene mostrato il tempo di CPU utilizzato mentre in colonna 6 viene riportato il rapporto tra tale valore ed il tempo di orologio. Quest'ultimo dato indica l'efficienza con la quale l'utente fa uso della risorsa di calcolo e dipende sia dalla efficienza intrinseca del codice che dal peso del tempo di trasferimento dei dati rispetto a quello di calcolo. Per esempio, le utenze corrispondenti ad attività di interoperabilità di progetti GRID (utenze N. 5, 35, 36, 42, 52) hanno un valore molto basso di tale dato (<10%) corrispondente probabilmente ad un peso rilevante dei processi di trasferimento dei dati su WAN. Esistono poi chiaramente grandi variabilità dell'efficienza di utilizzo tra codice e codice, con un picco di massima efficienza che raggiunge l' 88% per l'utente N. 21. Esiste poi un dato anomalo superiore ad 1, nel caso dell'utente N. 39, il cui codice utilizza largamente OpenMP e probabilmente viene mal valutato dal sistema di accounting.

2. **Utilizzo della memoria**

In *Tabella 1* l'ultima colonna si riferisce al parametro relativo all'utilizzo di memoria da parte dei processi dell'utente, mediato sul tempo di durata dei processi. La somma su tutti gli utenti deve corrispondere come ordine di grandezza ad una frazione della memoria disponibile del sistema, ed infatti vale 2.8 TB su una memoria RAM totale del sistema HPC CRESCO di 5.9 TB. Tenuto conto del fattore di utilizzo del sistema, tale dato indicherebbe che l'utente in media cerca correttamente di dimensionare il proprio caso in modo da sfruttare in modo significativo la memoria disponibile sul sistema, tipicamente 2 GB per ogni core disponibile.

3. **Frequenza di utilizzo dei nodi**

In ***Tabella A*** alla colonna 8 viene indicato il numero dei file mensili di accounting per nodo nel quale l'utente compare e permette di valutare con che frequenza viene utilizzato il sistema nel corso dell'anno. Essendo i nodi circa 300 il massimo possibile utilizzo è dato da un valore pari a circa 3600, ovvero il numero dei nodi per i mesi nell'anno. Per esempio le utenze delle persone coinvolte nella amministrazione del sistema hanno un valore molto alto di tale parametro poiché sono state coinvolte nell'esecuzione di job di test che hanno riguardato l'intero sistema. Le attività di progetti GRID tendono invece ad utilizzare un numero ristretto di nodi per periodi limitati.

4. **Dati sulla tipologia dei job sottomessi**

Un dato di grande interesse che non emerge in modo semplice dai dati di accounting raccolti nodo per nodo tramite le utility psacct [4] è la taglia media del job parallelo sottomesso dall'utente e che può essere solo indirettamente correlato con il numero di processi riportati nella prima colonna della ***Tabella A***. Esistono poi dei casi riguardanti i progetti GRID in cui job seriali dedicati alla verifica del buon funzionamento della infrastruttura GRID producono un grande numero di processi. Per ottenere il dato relativo al numero di processori utilizzati è necessario accedere al sistema di accounting del gestore di code LSF [5] che fornisce anche la valutazione sul tempo di CPU, ma in modo meno accurato del sistema basato su psacct. Il risultato dell'analisi di accounting dai dati di LSF per il 2009 è nella seguente *Tabella 3*, che è allineata alla ***Tabella A*** dell'appendice per quanto riguarda gli utenti, ma si limita a mostrarne i primi 30. Si tratta di utenti che hanno utilizzato nel corso del 2009 il sistema per almeno 10 anni circa di tempo di calcolo.

USER	N jobs	Min NC	Max NC	Media NC	NC1	R1	NC2	R2	R Max	%DONE
1	1013	1	840	266.9	64	0.226	840	0.205	0.205	0.091
2	1920	1	2032	286.6	128	0.177	160	0.082	0.001	0.100
3	1027	8	1600	513.8	512	0.171	1024	0.144	0.002	0.039
4	1041	1	1152	279.2	272	0.228	144	0.117	0.007	0.666
5	11533	1	1	1.0	1	1.000	1	1.000	1.000	0.469
6	692	8	1024	265.6	160	0.214	320	0.079	0.010	0.074
7	2776	1	384	24.4	8	0.625	16	0.116	0.001	0.582
8	207	32	640	245.3	64	0.333	512	0.237	0.005	0.454
9	176	8	1152	316.0	216	0.545	576	0.097	0.034	0.625
10	2502	1	1536	120.4	8	0.324	32	0.230	0.005	0.796
11	685	8	1904	182.3	304	0.270	256	0.134	0.001	0.266
12	4436	1	520	16.8	1	0.785	64	0.184	0.000	0.806
13	3809	1	120	17.8	1	0.575	40	0.195	0.000	0.689
14	5340	1	1024	17.4	8	0.787	16	0.082	0.001	0.235
15	829	1	840	205.7	672	0.223	8	0.211	0.034	0.532
16	1725	1	1024	45.4	16	0.447	64	0.164	0.001	0.661
17	328	16	2000	224.0	56	0.357	80	0.101	0.006	0.384
18	582	32	128	37.7	32	0.935	128	0.057	0.057	0.741
19	219	1	520	74.8	8	0.297	64	0.256	0.014	0.374
20	354	1	2016	221.4	8	0.469	128	0.178	0.037	0.528
21	112	8	1024	404.4	256	0.321	800	0.205	0.036	0.045
22	304	1	2024	576.9	1024	0.247	256	0.174	0.003	0.326
23	267	2	64	19.4	16	0.670	32	0.165	0.026	0.300
24	216	1	1600	250.5	8	0.468	576	0.356	0.005	0.217
25	1371	1	1000	10.9	1	0.599	8	0.355	0.001	0.744
26	116	24	128	85.7	64	0.647	128	0.345	0.345	0.543
27	826	1	2	1.4	1	0.561	2	0.439	0.439	0.708
28	72	1	80	43.5	64	0.361	32	0.278	0.056	0.611
29	450	1	1024	87.3	8	0.333	1	0.127	0.009	0.613
30	1762	1	2	1.0	1	0.971	2	0.029	0.029	0.883

Tabella 3: *Per i primi 30 utenti di* ***Tabella A*** *vengono mostrati il numero di job eseguiti, il numero di core (**NC**) fornendo il minimo, il massimo, la media, la dimensione dei due casi più frequenti (NC1 NC2) con il loro peso sul totale dei casi (R1 ed R2), il peso sul totale dei casi con il massimo numero di core. Nell'ultima colonna viene poi mostrata la frazione di job terminati con codice di ritorno 0.*

Come si vede dalla *Tabella 3* i maggiori utilizzatori del sistema HPC CRESCO hanno eseguito nel corso dell'anno un numero di job che va dalle molte centinaia a qualche migliaio. Nella quasi totalità il sistema viene utilizzato per job paralleli con un numero medio di processori utilizzati che va dalle molte decine a qualche centinaio. Molti degli utenti hanno provato la scalabilità del codice fino a migliaia di core. Fa eccezione l'utenza N. 5 di un progetto GRID che ha utilizzato il sistema per soli job seriali. Nelle colonne da 6 a 10 della vengono fornite informazioni più di dettaglio sul numero di core utilizzati dai vari utenti, poiché il valore medio non è in realtà molto significativo. Infatti le modalità di utilizzo degli utenti sono spesso strutturate in un numero limitato di tipologie di job abbastanza diverse l'uno dall'altro in termini di numero di core utilizzati. Ancor meno significativo sarebbe poi il valore della varianza poiché la distribuzione del numero di core utilizzati è ben lontana dall'essere una distribuzione normale. Pertanto, nel seguente paragrafo si forniscono alcune informazioni quantitative relative a tale distribuzione. A tal proposito nelle colonna da 6 e 7 viene riportata la dimensione in numero di core (**NC1**) ed il suo peso relativo (**R1**) sul totale dei job sottomessi del caso che l'utente sottomette con maggior frequenza. In colonna 8 e 9 gli stessi dati (**NC2**) e (**R2**) vengono forniti per il caso che segue quello di massima frequenza. In colonna 10 viene poi riportato il peso sul totale del caso avente la dimensione in core pari al massimo assoluto di core utilizzati, riportato in tabella in colonna 4. Per esempio esaminando l'utente N. 3 si vede che nel 17.7% dei casi ha utilizzato 512 core, nel 14.4% ne ha usati 1024 e il caso a 1600 core ha riguardato solo una piccola frazione del totale par al 0.2%. Diverso il caso dell'utente N. 1 per il quale il caso di taglia massima da 800 core corrisponde al 20.5% dei job sottomessi e quindi si tratta di un tipico caso di produzione.

Le colonne indicate con **NC1** e **NC2** e i relativi pesi **R1** ed **R2** sul totale dei job sottomessi danno una idea precisa della tipologia dell'utilizzo del sistema indicando che molti degli utenti utilizzano in modo sistematico in produzione job di grande dimensione che fanno uso di molte centinaia di core superando anche il migliaio. Si deve poi sottolineare che le frazioni **R1** ed **R2** sono calcolate sul numero totale di job sottomessi mentre sarebbe forse più significativo utilizzare la frazione di tempo macchina che compete a tali job, poiché questo dato sarebbe più indicativo del loro peso sull'utilizzo totale del sistema. Purtroppo, però, LSF non valuta in modo preciso tale dato e quindi non è stato possibile includerlo tra quelli mostrati. Dalla tabella emerge poi il fatto che un numero limitato di utenti (6 per la precisione) tra i primi 30 maggiori utilizzatori sottomette job seriali che fanno uso di un solo core. Tra queste le utenze del progetto GRID a cui si è già accennato precedentemente.

5. Dati sull'esito dei job

Dall'analisi dei dati di accounting forniti da LSF [5] è possibile ricavare informazioni sull'esito dei job sottomessi, andando a differenziarli in base al codice di ritorno. I job che formalmente finiscono correttamente sono indicati da LSF con la label DONE e la frazione di job che terminano in tal modo è illustrata nell'ultima colonna di *Tabella 3*. L'esame di tale dato però deve tenere conto del fatto che le code di sottomissione su CRESCO hanno in generale durata limitata a 24 ore e che gli utenti, in particolare quelli che fanno maggior uso di alto parallelismo preferiscono gestire il codice con metodi di checkpointing [6], salvando regolarmente i dati nel corso del calcolo ed implementando la ripartenza dall'ultimo caso risultato valido. Tali utenti preferiscono lasciare andare il job fino al limite temporale della coda trascurando il fatto che formalmente il job termina in modo anomalo senza ritornare il codice DONE. Per tale motivo risulta impossibile stimare in modo corretto dal dato di LSF la fine corretta del job. Il dato può

essere più significativo nel caso di job seriali per i quali le tecniche di checkpointing sono applicate meno frequentemente. Per i job seriali il dato generale vale almeno il 70% dei casi fino a punte dell'88%.

APPENDICE

***Tabella A**: Dati di utilizzo disaggregati per utente*
*Le variabili: **Nprocess**, numero di processi lanciati, **RealTime:** tempo di orologio in minuti ed anni, **CPUTime:** tempo di CPU in minuti, **tCPU/Real:** rapporto tra tempo globale di CPU e tempo di orologio, **avM(k):** è la memoria utilizzata mediata sul tempo di esecuzione in kB. **Nfiles:** numero di file mensili di accounting per nodo in cui l'utente compare.*

NUSER	Nprocess	RealTime(m)	RealTime(Y)	CPUTime(m)	tCPU/Real	avM(k)	NFiles
1	4129635	154550203.58	294.045	95786504.26	0.62	76581341	2686
2	3983254	151782820.55	288.780	97756596.49	0.64	76190201	2759
3	1853712	110792970.75	210.793	68750844.71	0.62	82189125	1944
4	1182235	96724735.53	184.027	77025221.22	0.8	119818078	2057
5	27891610	51839844.72	98.630	2752275.05	0.05	1175883	91
6	778291	50048609.47	95.222	33329797.93	0.67	96004214	1989
7	446440	42460530.25	80.785	26011064.21	0.61	54482227	1643
8	201262	40062777.34	76.223	24477355.77	0.61	63717693	809
9	193567	33806391.06	64.320	22849435.57	0.68	92226865	1466
10	26137258	30659718.19	58.333	16254309.27	0.53	86684849	3328
11	287609	25834083.88	49.152	17882663.65	0.69	206877793	1484
12	5135012	24470829.65	46.558	14336979.09	0.59	53122560	1718
13	1310591	21887208.68	41.642	12738032.64	0.58	29763895	1206
14	1982841	19893146.03	37.848	14069327.81	0.71	288486756	1485
15	4052026	18990634.35	36.131	11134915.62	0.59	35729117	1151
16	1657633	15691215.69	29.854	3739560.06	0.24	32945807	1086
17	952777	13118982.67	24.960	5044733.06	0.39	39471421	1500
18	60824	12962549.4	24.662	9667152.27	0.75	338490116	629
19	193121	11856477.47	22.558	10048214.53	0.85	54614275	795
20	186127	11632055.61	22.131	7927954.62	0.68	360480213	2290
21	134179	11486526.36	21.854	10103915.22	0.88	26736766	506
22	1923495	11113433.09	21.144	7729320.02	0.7	78720189	3238
23	379944	10390382.87	19.769	5876759.85	0.57	1901862	124
24	247541	6811269.46	12.959	1778514	0.26	28999550	810
25	2799477	6734076.56	12.812	972086.59	0.14	40793367	560
26	26600	6506883.35	12.380	4990856.41	0.77	29502258	482
27	1452644	5994080.56	11.404	1269118.59	0.21	2354374	156
28	152075	4928504.73	9.377	3171797.06	0.64	2272325	144
29	1808025	4281070.19	8.145	2136904.05	0.5	28946232	1501
30	24894196	3873648.92	7.370	277719.33	0.07	1425151	106
31	21771	2683724.85	5.106	1866407.02	0.7	11855571	210
32	1289222	2615524.08	4.976	1240696.67	0.47	19012883	1113
33	17990	2463299.56	4.687	1561443.9	0.63	2604573	41
34	6351706	2280646.89	4.339	1193401.06	0.52	2880046	164

NUSER	Nprocess	RealTime(m)	RealTime(Y)	CPUTime(m)	tCPU/Real	avM(k)	NFiles
35	1656211	2185813.63	4.159	132142.56	0.06	584631	41
36	4051349	2015753.72	3.835	118122.71	0.06	806429	58
37	328635	2015018.68	3.834	560480.96	0.28	5902728	345
38	595906	1873506.79	3.565	908249.12	0.49	2905491	264
39	119258	1679405.99	3.195	4174008.52	2.49	2234024	123
40	13787218	1598290.82	3.041	223052.94	0.14	1705279	103
41	141321	1469397.98	2.796	176143.93	0.12	2955418	217
42	677707	1462521.59	2.783	98927.84	0.07	414457	30
43	138825	1450538.43	2.760	327010.25	0.23	5406806	89
44	449722	1446919.82	2.753	900970.1	0.62	5627899	208
45	2654695	1386685.41	2.638	104.41	7.53E-005	918313	76
46	21543	1314150.78	2.500	665566.93	0.51	10189358	87
47	88986	1303102.38	2.479	911057.49	0.7	26249093	545
48	227314	1278783.47	2.433	713504.58	0.56	34760070	241
49	29119	1139649.09	2.168	894929.65	0.79	15626899	258
50	28703	1091070.71	2.076	719580.64	0.66	13266052	250
51	138305	1060555.61	2.018	314672.7	0.3	28381445	189
52	307516385	925266.31	1.760	11569.67	0.01	2328478	156
53	101924	876289.5	1.667	235926.42	0.27	1926546	54
54	102772	871874	1.659	451588.42	0.52	709357	72
55	1642556	865986.11	1.648	21119.49	0.02	4363452	160
56	132063	815216.44	1.551	316427.19	0.39	1823764	54
57	18943	801223.48	1.524	207479.96	0.26	5594587	44
58	399551	785678.88	1.495	62923.94	0.08	874039	63
59	85222	757881	1.442	481179.15	0.64	13359498	528
60	10444173	713675.98	1.358	70406.59	0.1	33674018	3520
61	13147	700791.24	1.333	520596.27	0.74	7632610	157
62	324232	636196.26	1.210	119026.88	0.19	3115225	128
63	1077709	506474.89	0.964	774.9	0	27608864	2512
64	38411	433709.04	0.825	190644.41	0.44	9801883	211
65	61427	344922.93	0.656	20819.78	0.06	478916	41
66	71991	337737.02	0.643	168982.52	0.5	3656662	1095
67	218697	324538.58	0.617	16.54	5.10E-005	524026	20
68	38302	310333.26	0.590	108493.17	0.35	451756	29
69	201132	282240.83	0.537	12365.94	0.04	255503	18
70	7881	271721.08	0.517	231236.61	0.85	816667	45
71	17937	252232.11	0.480	6.02	2.39E-005	345913	47
72	3217	247939.2	0.472	71528.33	0.29	4474543	26
73	114609	231052.47	0.440	896.61	0	150405	16
74	19327	227714.38	0.433	45888.71	0.2	408675	33
75	2477856	193137.65	0.367	58627.43	0.3	212511	61
76	26683	180027.14	0.343	22792.65	0.13	1908883	97
77	165373	159748.53	0.304	94084.5	0.59	3456677	170
78	4625	155495.11	0.296	5379.89	0.03	138817	10
79	222627	150737.72	0.287	942.37	0.01	3656091	328
80	499371	142365.24	0.271	61.37	0	892782	76

NUSER	Nprocess	RealTime(m)	RealTime(Y)	CPUTime(m)	tCPU/Real	avM(k)	NFiles
81	101048	121855.64	0.232	499.53	0	266841	19
82	71990	119178.76	0.227	1790.23	0.02	332651	31
83	337716	115597.31	0.220	517.53	0	8759266	1150
84	210096	112786.56	0.215	61746.48	0.55	5649981	69
85	6373	112473.44	0.214	33558.49	0.3	7432498	36
86	119427	111149.12	0.211	316.91	0	253234	12
87	21480	103778.44	0.197	41979.85	0.41	465387	50
88	72326	92805.33	0.177	5422.85	0.06	370838	26
89	20555	88804.74	0.169	36695.04	0.41	315141	23
90	50416	83385.3	0.159	25.63	0	219387	24
91	15258	67554.47	0.129	2723.41	0.04	14552267	85
92	237389	60079.21	0.114	12198.12	0.2	387282	29
93	10569129	52820.8	0.100	153.5	0	677148	47
94	21979	50353.91	0.096	32.37	0	265731	24
95	6838	46506.92	0.088	49.05	0	126023	11
96	9968	39680.88	0.075	16975.59	0.43	302498	6
97	13820	35345.46	0.067	266.56	0.01	652186	53
98	6626	34761.5	0.066	0.01	2.88E-007	104799	13
99	19591	31027.27	0.059	1859.68	0.06	54425	4
100	8138	24398.89	0.046	0.21	8.61E-006	220132	20
101	2057	24397.5	0.046	1.68	6.89E-005	158855	15
102	11339	24238.3	0.046	17370.8	0.72	2644910	49
103	56472	23366.07	0.044	265.99	0.01	332184	29
104	4667075	23218.35	0.044	94.2	0	2118949	153
105	13226	23082.02	0.044	6758.14	0.29	805038	41
106	10125117	22860.09	0.043	203.85	0.01	522365	35
107	5469	20529.73	0.039	3.31	0	249198	18
108	17350	19464.31	0.037	217.29	0.01	4277987	29
109	859	19141.3	0.036	1.62	8.46E-005	34810	2
110	51249	17020.16	0.032	4707.13	0.28	663135	30
111	12410	17001.32	0.032	414.97	0.02	140689	15
112	63706	16909.46	0.032	4560.64	0.27	2487886	65
113	265	15823.78	0.030	2.06	0	7068	1
114	18402	13729.89	0.026	4.49	0	51822	5
115	24779	13375.32	0.025	0.14	1.05E-005	115204	14
116	39710	12841.49	0.024	72.46	0.01	478513	36
117	6541	10673.55	0.020	3795.88	0.36	1870078	74
118	4549	10399.78	0.020	238.95	0.02	143438	15
119	433	9927.5	0.019	1925.89	0.19	99229	8
120	20872	9779.17	0.019	0.25	2.56E-005	4381	3
121	18505	9308.07	0.018	172.25	0.02	310502	5
122	5131	8397.79	0.016	1.29	0	1186150	43
123	13632	7510.06	0.014	2401.14	0.32	180821	12
124	5552	7259.21	0.014	0.58	7.99E-005	86531	10
125	85697	7203.96	0.014	0.11	1.53E-005	19189	1
126	20888	6955.67	0.013	374.59	0.05	84705	6

NUSER	Nprocess	RealTime(m)	RealTime(Y)	CPUTime(m)	tCPU/Real	avM(k)	NFiles
127	86294	6924.47	0.013	8.33	0	30010	3
128	13165	6911.92	0.013	28.08	0	60622	10
129	732	6689.03	0.013	0.39	5.83E-005	33107	3
130	1121075	6592.03	0.013	23.39	0	914173	67
131	1482930	6251.69	0.012	33.02	0.01	1107735	81
132	4670	5435.85	0.010	592.89	0.11	61220	7
133	136751	4854.1	0.009	56.62	0.01	288879	22
134	66969	4219.71	0.008	249.07	0.06	125890	9
135	83798	4111.62	0.008	255.38	0.06	80195	5
136	543	3806.51	0.007	7.68	0	58281	5
137	1107	3688.41	0.007	0.02	5.42E-006	84335	11
138	183911	3299.56	0.006	14.69	0	12586	1
139	1732	3260.61	0.006	0.03	9.20E-006	130429	16
140	940	3163.61	0.006	2.81	0	27180	3
141	338	3085.7	0.006	18.27	0.01	42033	4
142	25522	2777.47	0.005	679.43	0.25	683912	51
143	1086	2695.16	0.005	0.48	0	79140	3
144	45331	2402.67	0.005	0.69	0	206371	15
145	7299	2337.82	0.004	8.97	0	134744	10
146	51343	2270.29	0.004	0.26	0	3889772	282
147	194	2132.03	0.004	662.78	0.31	26390	5
148	2251	2118.88	0.004	0.03	1.42E-005	105910	13
149	258	2018.98	0.004	0.02	9.91E-006	23196	2
150	25466	1771.19	0.003	30.75	0.02	175442	4
151	2519	1703.14	0.003	2.62	0	154228	16
152	185849	1685.25	0.003	61.02	0.04	639152	46
153	47474	1643.39	0.003	11.06	0.01	9968	1
154	288444	1478.96	0.003	14.26	0.01	1624154	118
155	1258	1457.14	0.003	49.86	0.03	29451	3
156	117485	1394.14	0.003	1.49	0	566152	41
157	511	1312.07	0.002	1.07	0	29971	2
158	665	1243.03	0.002	0	0	55778	7
159	74514	1220.59	0.002	0.83	0	21979	2
160	2513	1220.22	0.002	15.53	0.01	74851	6
161	331737	1087.9	0.002	4.23	0	1156786	83
162	257791	1074.91	0.002	3.28	0	1249847	90
163	72915	1066.82	0.002	1.41	0	504186	36
164	597	1016.06	0.002	84.02	0.08	12841	1
165	118	1005.03	0.002	0	0	7681	1
166	360	981.07	0.002	0.01	1.02E-005	12152	2
167	32069	835.31	0.002	2.25	0	19487	2
168	509	810.46	0.002	0.05	6.17E-005	32162	6
169	62	777.22	0.001	0	0	2157	1
170	130	753.09	0.001	156.15	0.21	595019	2
171	12815	728.74	0.001	0.22	0	308123	25
172	183585	659.31	0.001	2.39	0	1247881	90

NUSER	Nprocess	RealTime(m)	RealTime(Y)	CPUTime(m)	tCPU/Real	avM(k)	NFiles
173	5500	644.72	0.001	1.44	0	18240	3
174	182442	616.83	0.001	2.21	0	1386248	100
175	393	538.26	0.001	0.13	0	25309	4
176	22196	502.22	0.001	0.3	0	206480	15
177	79085	495.62	0.001	1.05	0	868438	63
178	147163	491.32	0.001	1.98	0	180821	13
179	91382	469.47	0.001	0.57	0	327036	24
180	78	436.68	0.001	0	0	10725	2
181	71706	394.15	0.001	0.89	0	844553	60
182	73	356.09	0.001	0.11	0	8708	2
183	6028	346.27	0.001	0.15	0	196712	16
184	74061	343.87	0.001	0.92	0	760968	55
185	238	340.01	0.001	0	0	19317	4
186	651	317.21	0.001	0.32	0	40974	4
187	18760	316.76	0.001	0.32	0	206477	15
188	12055	288.64	0.001	0.27	0	61570	4
189	13807	286.24	0.001	0.28	0	166496	12
190	38483	249.36	0.000	1.15	0	338953	25
191	699	223.61	0.000	0.01	4.47E-005	80860	10
192	30059	220.19	0.000	0.36	0	443031	32
193	62446	219.18	0.000	0.77	0	733998	53
194	3312	195.02	0.000	0.01	5.13E-005	155841	26
195	103	180.12	0.000	0	0	10423	1
196	14134	174.45	0.000	0.18	0	96744	7
197	7516	166.74	0.000	4.76	0.03	49277	5
198	97222	161.61	0.000	1.55	0.01	59005	4
199	49787	144.09	0.000	0.61	0	482746	35
200	236	137.04	0.000	73.18	0.53	82484	3
201	602	136.19	0.000	0.02	0	21628	2
202	10780	122.88	0.000	9.81	0.08	118098	2
203	30	100.03	0.000	0.01	0	3230	1
204	444	85.01	0.000	0.01	0	40060	6
205	23863	84.03	0.000	0.29	0	209313	15
206	12931	76.99	0.000	0.14	0	83448	6
207	18089	76.83	0.000	0.22	0	85072	6
208	9796	74.77	0.000	0.15	0	139531	10
209	15104	70.88	0.000	0.22	0	99605	7
210	1397	67.7	0.000	0	0	175458	14
211	126	64.27	0.000	0	0	7008	1
212	456	63.03	0.000	0.05	0	26567	2

BIBLIOGRAFIA

[1] G. BRACCO, S. MIGLIORI, S. PODDA et al. *"CRESCO HPC System Integrated into ENEA GRID Environment"* Proceedings – Final Workshop of GRID Projects PON Ricerca 2000-2006, Call 1575 pp. 151 - 155.

[2] http://www.infinibandta.org/

[3] http://www-03.ibm.com/systems/software/gpfs/

[4] http://www.gnu.org/software/acct/

[5] http://www.platform.com/workload-management/high-performance-computing

[6] http://en.wikipedia.org/wiki/Checkpointing

NUMERICAL MODELLING OF CONFINED PARTICLE LADEN FLOWS

Fiorenzo Ambrosino, ENEA - UTICT-HPC, C.R. Portici

Francesco Saverio Marra, Istituto di Ricerche sulla Combustione – CNR, Napoli

Introduction

This work is a collaboration between ENEA, in particular with the technical unit UTICT-HPC, and the Istituto di Ricerche sulla Combustione (IRC) of the CNR. It is also part of the PhD research work of the first author at the Chemical Engineering Department (DIC) of the University of Naples Federico II.

It has been promoted by the observation that coal combustion is considered to be among the main primary energy supply that will be used for very long time. Scenarios depict that, while oil and gas availability will soon decreases, a large share of the energy demand will be supplied by coal combustion for at least more than a century from now. In contrast, coal combustion is among the most pollutant energy technologies, thus requiring an huge research effort to introduce new, less pollutant combustion technologies.

Modelling of coal gasification processes is therefore an important step for the proper design of new primary energy transformation plants. Among these processes the evolution of particles, that can represent pulverized coal, soot or ash, as well as other pollutant species, immersed in a turbulent gas stream is of great interest for the combustion community and the energy supply actors.

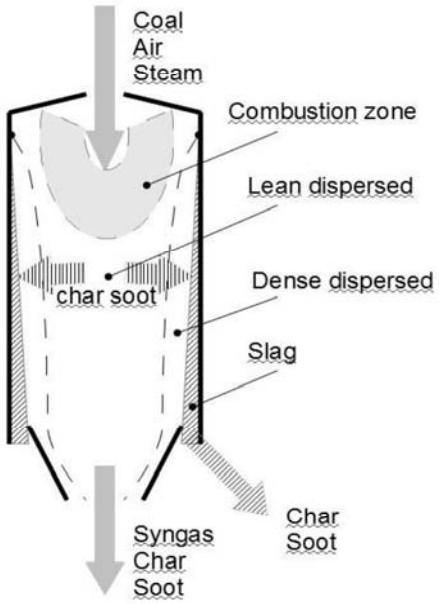


Figure 1: *Schematic diagram of the entrained-flow gasifier.*

There are many areas where a lack of knowledge on the fundamental mechanisms is present. One of these areas cover the transport mechanisms that govern the motion of pulverized coal in the combustion chamber, especially in proximity of the confining walls *(see Figure 1)* [1].

This problem arises not only in this specific application field but also in several other technologies involving particle-laden flows, making difficult to design devices used in many chemical engineering fields ranging from pharmaceutical to environmental. Even if the focus of this work is on the clean combustion technology, thus limiting the investigations to conditions usually applicable to a coal gasifier, it is expected that the knowledge and modelling tools that will be acquired on this phenomena could be easily extended to other field of applications.

The modeling is performed through the use of Computational Fluid Dynamic (CFD) techniques and tools in order to validate the methods and because there are no many experimental due to the excessive costs.

OpenFOAM is the numerical code adopted for simulating the physics involved in this study. All needed models have been implemented and validated with special care to the Large Eddy Simulation (LES) formulation adopted for the turbulence modelling and the multiphase formulation based on the Lagrangian Particle Tracking (LPT). The OpenFOAM (Open Field Operation and Manipulation) CFD Toolbox is a free, open source CFD software package produced by a commercial company, OpenCFD Ltd [2]. It has a large user base across most areas of engineering and science, from both commercial and academic organizations. OpenFOAM has an extensive range of features to solve anything from complex fluid flows involving chemical reactions, turbulence and heat transfer, to solid dynamics and electromagnetics. It allows for an effective cooperation with other research group towards the task of building a comprehensive model of a coal gasifier plant. The core technology of OpenFOAM is a flexible set of efficient C++ modules. These are used to build a wealth of: solvers, to simulate specific problems in engineering mechanics; utilities, to perform pre- and post-processing tasks ranging from simple data manipulations to visualization and mesh processing; libraries, to create toolboxes that are accessible to the solvers/utilities, such as libraries of physical models.

OpenFOAM uses finite volume discretization of the basic model equations, the Navier-Stokes balance equations for mass, momentum and energy, valid for a multi-component reactive mixture. The resulting system of partial differential equations can be solved on arbitrary 3D domains, partitioned through unstructured meshes of polyhedral cells. The fluid flow solvers are developed within a robust, implicit, pressure-velocity, iterative solution framework, although alternative techniques are applied to other continuum mechanics solvers. Domain decomposition parallelism is fundamental to the design of OpenFOAM and integrated at a low level so that solvers can generally be developed without the need for any parallel-specific coding.

Open source code has been considered a good choice to investigate very computationally expensive CFD problems due to the high number of computational cores required, number usually limited with the adoption of CFD commercial software licences.

Activity Description

The Euler-Lagrangian (E-L) approach for the gas and particle phases respectively is a good compromise to model predictively the particle dispersion in gas and focus on the mechanisms of wall-particle interaction. This class of problem usually needs a very big computational power due to both the high number of computational cells required to properly select the cut-off

threshold of the LES filter and the high number of particles to track. Turbophoresis effects of the fluid phase on particle motion are very sensitive to the dimension of the resolved scale so to model correctly the particle-wall interaction phenomena there is the need of a good resolution of the mesh size near the wall and an appropriate turbulence model. Direct Numerical Simulations (DNS) for the fluid phase are too computational expensive for industrial geometries and, on the other hand, Reynolds Averaged Simulations (RAS) can't get the effect of physical turbulent eddies on the particle motion.

The simulations performed are involving a lot of particles, from 10^5 to 10^6 particles with diameter not too big compared to the dimension of the computational cell.

A one-way coupling approach is considered. It means that the particle phase motion is affected by fluid phase flow conditions but the contrary does not hold. This is considered a valid assumption if the particle volumetric concentration is not too high (dispersed regime).

Turbulence has been modelled with a LES approach. The one equation dynamic model for Sub Grid Scale (SGS) terms is used. Gravity force is not considered in the momentum source term. The gas model used is the perfect gas.

At particle injection time, fixed in conditions of fully developed flow, N=100000 particles are placed in the channel. The injection adopted is not physical. Indeed the particles were placed instantly in the overall channel domain, randomly spaced along the x, y and z direction respectively. The initial particle velocity is the same of the fluid at particle positions.

Since the number of particles very often exceed the computational tracking capabilities, it is common to represent a group of physical particles by a single computational unit called "parcel". The physical particles inside a parcel are assumed to have the same position, diameter, velocity, temperature and other properties. The concept of parcel is useful when the size of the particles is very small compared to the computational cell sizes. Indeed the parcel dimension must be smaller of the computational cell size. Particles in the parcel are not supposed to interact each other.

In the simulations presented in this work each particle is tracked independently so each parcel has got only one particle inside. This assumption is due to the fact that near wall, the mesh size required to correctly reproduce by the LES approach the viscous sub-layer becomes very small, of the same order of magnitude of the particle diameter. This has been one of the main reason demanding for the computational power available at CRESCO.

The methodology used to solve the equation of motion of the particles is the Lagrangian Particle Tracking (LPT). Tracked particles are considered point-wise. They are assumed to be rigid hard spheres. Relevant forces acting on particles are Stokes drag and buoyancy. The drag coefficient is assumed to be function of Reynolds particle number. Gravity is not considered.

The algorithm implemented in OpenFOAM for the LPT is described in [4] in which is shown the TrackToFace method. This method practically consists of moving the particle between faces it crosses in the time step duration. In *Figure 2* is shown an example of motion of a particle during a single time step.

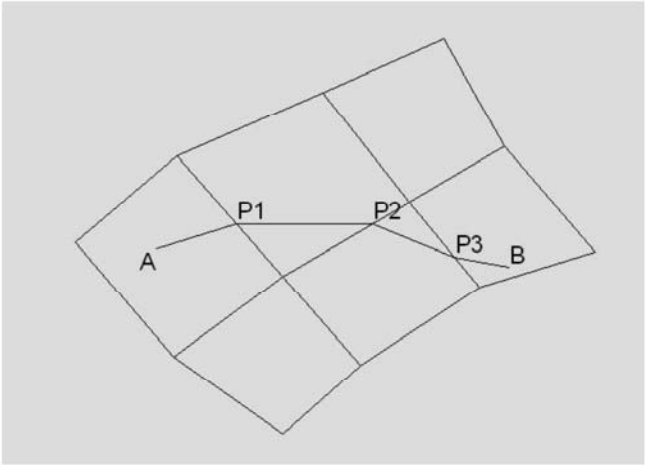


Figure 2: *Lagrangian Particle Tracking: sketch of TrackToFace method.*

It is a particle-laden turbulent gas flow in a flat channel. The case study is the same to the one discussed in [3] in which the effect of main particle parameters, like their density and diameter, globally grouped into the Stokes non-dimensional number [5], is observed. The aim of this activity was to verify and validate the numerical tools and methodologies selected. The capabilities of the CFD toolbox OpenFOAM to handle this class of problems have been so assessed. Moreover, being a simple study case, it is possible to consider it as a starting point on this topic in addressing problems of gradually increasing complexity.

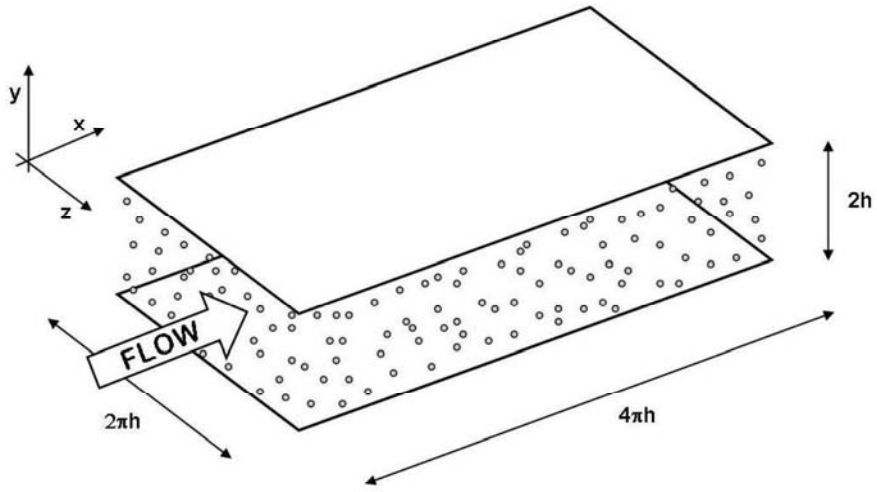


Figure 3: *Periodic channel case geometry.*

The physical domain is showed in *Figure 3*. It is a simple three-dimensional flat channel with two infinite flat parallel no-slip walls on the top and on the bottom. This walls have the function to confine the jet of particle-laden flow. Periodic boundary conditions are assumed at the other four sides of the box.

Two different clouds of particles are considered in simulations, characterized by a different diameter: 45.6 and 102 μm respectively. This leads to two different values of the non-dimensional Stokes number (St).

Periodic boundary conditions are imposed on particle moving outside the computational domain in stream-wise and span-wise directions, so they are reintroduced via periodicity.

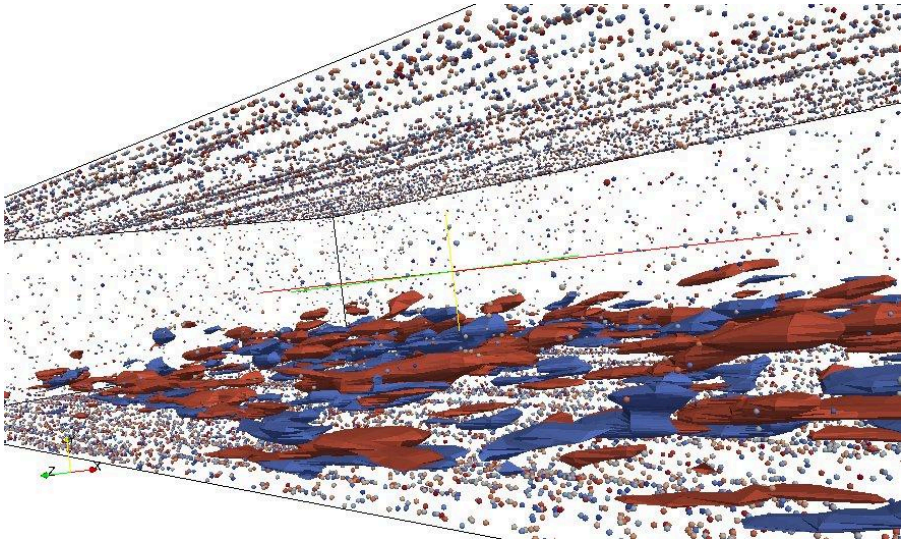


Figure 4: *Turbulent structures and particle interactions in boundary layer.*

The complex interaction between the particle phase and the fluid flow give rise to the formation of structures in the boundary layer as shown in *Figure 4*. Surfaces of iso-Q values (red corresponding to 1000 and blue to -1000 values), where Q is the second invariant of the velocity gradient, are reported only for the lower half of the channel. Very intense turbulent structures are formed in the turbulent boundary layer, allowing turbophoresis to promote an accumulation of particles on the wall. This accumulation, however, is not uniform: particles tends to allineate along the streamwise direction. This kind of observations is very important to characterize the accumulation and segregation of coal particles onto the surface of a gasifier, being these phenomena responsible of the conversion degree and the emission level of polluttant species. Further investigations will be conducted to include the effects of the properties of the bounding walls on the particles-wall interaction.

In *Figure 5* is shown a top view of the bottom wall of the channel for particle cloud with St=25. In this figure only particles close to the wall (distance lower than 4mm) are shown so it is possible to recognize the particle and boundary layer interaction. In that figure it is possible to see different particle concentration along the span-wise direction. This is due to the presence of turbulent eddies near the wall that are affecting to the particle concentration. This phenomenon is not possible to be caught by a RANS method.



Figure 5: *Top view of the channel; particles agglomeration for Stokes=25.*

The high performing computing demand is notable. An open source CFD code allows to increase the number of computational nodes, without increasing the license cost, up to the number of computational nodes limited by the saturation of the speed-up increase. Actually, despite of the simple geometry of the channel case and only 230000 computational cells required for the Eulerian phase, the problem has been simulated with about one hundred computational cores working in parallel. Several simulation tests have been done to obtain the optimum domain decomposition method necessary to reach good values of parallel efficiency and speed-up. The optimum has been obtained by decomposing the domain of 48 x 96 x 48 cells in 96 sub domains (partitioning 16 times along x direction, 2 times along y direction and 3 times along z direction).Also different CRESCO [6] architecture machines have been tested: the "cresco1" group (IBM x3850-M2 each with 4 Xeon Quad-Core Tigertown E7330 processors (2.4GHz, 1066MHz, 6MB L2), 32/64 GB RAM), "cresco2" (IBM HS21 each with dual Xeon Quad-Core Clovertown E5345 processors (2.33GHz, 1333MHz, 8MB L2), 16 GB RAM) and the "cresco nehalem" group (dual Intel Xeon Quad-Core Turbo Boost Nehalem E5530 processors (2.40 GHz, 8MB L3), 16 GB RAM).The 96 cores simulation on "cresco2" architecture machines have reached a speed-up value of about 76 that corresponds to an efficiency of about 80%. Absolute speed performances are worst on "cresco1" and the best on "cresco nehalem" machines but simulations have been done prevalently on "cresco2" because of the larger availability of these machines.

A typical simulation of the particle laden flow in the channel that cover 1800 simulated seconds has required about 400 hours of calculation time (about 2 weeks) with 96 parallel working cores.

REFERENCES

[1] Montagnaro, Fabio, & Salatino, Piero. *"Analysis of char-slag interaction and near-wall particle segregation in entrained-ow gasi_cation of coal"*. Combustion and Flame **157** (2010), pp. 874-883.

[2] OpenFOAM The Open Source Computational Fluid Dynamics (CFD) Toolbox website, http://www.opencfd.co.uk/openfoam/

[3] Marchioli, C., Soldati, A., Kuerten, J.G.M., Arcen, B., Tanire, A., Golden-soph, G., Squires, K.D., Cargnelutti, M.F., & Portela, L.M. *"Statistics of particle dispersion in direct numerical simulations of wall-bounded turbulence: Results of an international collaborative benchmark test"* International Journal of Multiphase Flow **34** (2008) pp. 879-893.

[4] Macpherson, Graham B., Nordin, Niklas, & Weller, Henry G. *"Particle tracking in unstructured, arbitrary polyhedral meshes for use in CFD and molecular dynamics"*. Communications in Numerical Methods in Engineering, **25** (2009), pp. 263-273.

[5] Sommerfeld, M. *"Overview and Fundamentals. In: Theoretical and Experimental Modelling of Particulate Flows"*. Lecture Series 2000-06. von Karman Institute for Fluid Dynamics.

[6] Enea CRESCO website, http://www.cresco.enea.it

INTEGRATION OF FLUKA SIMULATIONS FOR TOP-IMPLART TREATMENT PLANNING SYSTEM

E. Basile, D. M. Castelluccio, E. Cisbani, S. Frullani

Istituto Superiore di Sanità – Roma

Activity

The ENEA GRID Cresco system resources have been used for simulating therapeutic proton beams using the FLUKA code **[1] [4]**.

This is the preliminary work of several years of research program for the development of a custom Treatment Planning System (TPS), which will be realized for the "*Terapia Oncologica con Protoni and Intensity Modulated Proton Linear Accelerator for Therapy*" (TOP-IMPLART) project, conducted by ENEA, ISS and IFO-Regina Elena. The TOP-IMPLART project aims to realize an innovative proton therapy facility in Italy, based, for the first time, on a linear proton accelerator. A fully active 4D treatment mode, which allows a much better dose conformity and homogeneity over the Planning Target Volume (PTV) respect to conventional techniques for many tumors, will be accomplished by this accelerator.

Introduction

Proton therapy offers an enhanced possibility of effective tumour control for several tumour categories **[4][5]** respect to the conventional radiotherapy techniques. The rationale for the use of proton beam in radiotherapy stands on the physical characteristics of energy loss when protons (and heavier ions) penetrate matter: the entrance integral dose is low respect to the dose released at the end of the path (Bragg Peak) with a rapid distal dose fall-off.

In the last twenty years, a progressive development of hadron therapy took place and more than 60000 patients have been treated worldwide. Many hospital based centers are nowadays active in the world (at least 40) and many centres are under construction or in a phase of advanced project**[6].**

In Italy the research "*Centro di AdroTerapia e Applicazioni Nucleari Avanzate*" (CATANA) centre is already in operation. The centre allows the treatments of shallower tumours (4 cm max) like uveal melanomas **[7]**.

Recently the "*Centro Nazionale di Adroterapia Oncologica*" (CNAO), has been inaugurated and soon the first patient is expected to be cured.

The TOP_IMPLART project aims to realize an innovative proton therapy facility, based for the first time, on a linear accelerator. The TOP-IMPLART project benefits of previous studies, designs and tests done under the TOP project carried on in the years 1995-2005 by "*Istituto Superiore di Sanità*" (ISS), "*Agenzia Nazionale per le nuove tecnologie, l'energia e lo sviluppo economico sostenibile*" (ENEA) and "Istituto Regina Elena" for the realization of a proton facility. The TOP-IMPLART project is going to be refunded by the Department of Innovation of Regione Lazio.

The TOP-IMPLART project

The TOP-IMPLART facility will be very innovative, since it will be the first worldwide which will use a compact proton linear accelerator which permits an active 3+1D scanning. Unlike proton linear accelerators used for research purposes, which operate with intense beams at low frequencies, the TOP-IMPLART facility is planned to generate very low intensities at moderately high frequencies, as requested in clinical applications. Moreover the TOP-IMPLART current intensity must be widely varied on a pulse to pulse basis. These features allow the fully active 3+1D scanning. In the 3D+1 scanning mode the beam is actively longitudinally and transversally scanned with current intensity modulation on a pulse to pulse basis (active longitudinal scanning is allowed by active energy modulation). The current modulation at high frequency allows to achieve a better dose homogeneity over the PTV, as long as the TPS is adequately detailed and realistic.

The development of a custom TPS is of great importance for a facility. In fact, although commercial TPS are available, an accurate dose calculation requests the knowledge of the specific beam delivery modalities of the accelerator which will be used. Then it is recommended to develop a tool, customized and optimized for the specific facility.

Monte Carlo simulations of energy deposited by a proton beam

The development of a TPS needs several years of research and development, and the work done so far represents the preliminary stage of this program. The approach chosen for the development of the TPS is aimed to match the precision of Monte Carlo simulations and the rapidity of analytical calculation in order to estimate dose distributions; this is done defining an analytical optimized model derived from and constructed on Monte Carlo simulation of the dose distribution in heterogeneous media, using simulations in water as benchmark.

Several Monte Carlo simulations, performed with the FLUKA **[1] [1]**code, have been run over the Cresco system of the ENEA GRID. The use of Cresco system has allowed the simultaneous run of different simulations in parallel and a great reduction of the computing time. The simulations are discussed in details in **[3]**.

FLUKA code is developed for serial running, but parallelization has been performed in order to greatly reduce the computing time.

The following two assumptions taken by this code have been allowed "parallelizing" the run:

- particles from a radiation source interact with matter but not with each other;
- radiation histories do not perturb each other.

In order to reduce the running time the *embarrassingly parallel* approach **[8]** has been adopted, that is multiple processors have been used concurrently to solve the problem, executing different batches of histories independently. This approach provides multiple output files which have been combined together, so to obtain results identical to that of a serial computation of all histories, one after another, on a single processor: instead of running a long simulation treating all histories on a single processor, many short simulations treating a reduced number of histories in parallel on many processors have been run. In this work the following steps have been done in order to "parallelize" the FLUKA run:

- modifying the input file, by introducing a variable that allows initializing different and independent random number sequences;
- writing a bash script, that submit a parallel job on Cresco;
- writing a bash script, that, combines the output files from different GRID processors, calling some FLUKA post-processing routines.

Simulations of a proton beam impinging on a homogeneous target, have been performed varying the beam energy, the target material and the incidence angle. For every simulation a job composed by 100 parallel cycles of 10 runs of 100.000 primary particles, equivalent to a single run of 100.000.000 of primary particles, has been submitted. The Cresco serial queues *cresco_serh48* and *cresco_serh144* have been used for an average time of 17 hours per job and a storing space of 33 Gb per job for an equivalent total time of 21.2 years, which shows that without the Cresco facility (or equivalent) this development cannot be carried on.

Discussion and results

The work aimed at an analytical way of describing the dose distribution in a heterogeneous material, starting from Monte Carlo simulations of a proton beam impinging in water and in other materials. The presence of heterogeneity affects the proton range penetration, so water equivalent density is generally used for computing the dose distribution. The water equivalent density accounts for the different stopping powers of the crossed media. Data in water can be used as benchmark in order to reproduce dose distributions in other materials, using the water equivalent density. In order to find the analytical function for dose description, several steps have to be done and this work **[3]** is the starting way to face this problem:

1) the deposited dose along the beam axis (Bragg curve) in several materials has to be represented in terms of analytical Bragg curve. A suitable analytical way of reproducing the deposited energy integrated over the transverse plane in bone and "lung" in terms of the deposited energy integrated over the transverse plane in water has been presented for several beam energies and generalized for any energy. For example, in *Figure 1* is represented the analytical model of the integrated deposited energy in bone for a 200 MeV proton beam, obtained starting from the integrated deposited energy FLUKA data in water.

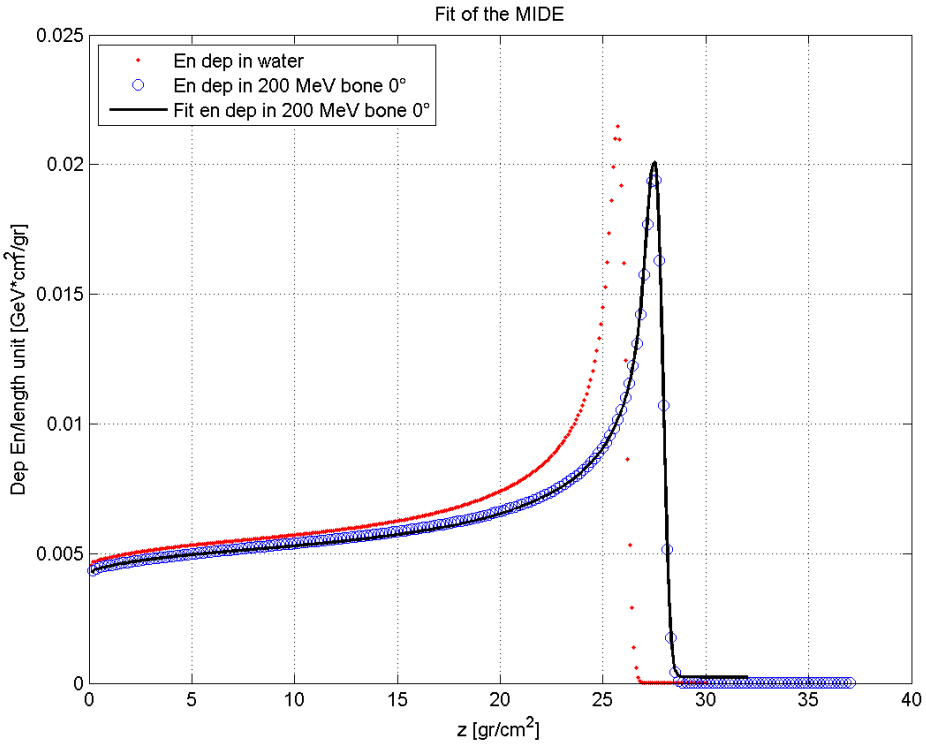


Figure 1: *Analytical model (in black) of the integrated deposited energy along the beam axis (in blue); for comparison the integrated deposited energy in water (in red) is also shown.*

2) multiple scattering in water and in other materials has to be reproduced. This is accomplished by fitting the dose distribution in the plane orthogonal to the beam direction at several depths along the beam axis for beams of various energies impinging in different materials. It has been found that a suitable fit function could be a sum of two Gaussian functions, whose significant parameter is the sigma parameter, which accounts for the beam spread. An analytical expression should be found for the sigma values relative to different target. Moreover the sum of the heights of the two Gaussian at each penetration depth represents the corresponding deposited energy along the beam axis. The fit over the transverse plane can then give the solution to the problem of analytical formulation.
3) several simulations of proton beam impinging at various angles respect to the surface normal for different energies have been accomplished and the results have been shortly described.

The results of these works have to be harmonized in an all-in analytical solution of the problem.

Starting from the CT (Computed Tomography) scan data, suitably converted in water-equivalent density, it should be possible to evaluate the dose distribution accounting for the heterogeneities, the multiple scattering and the beam incidence not orthogonally to surface. This evaluation has to be compared with experimental measure in order to validate the method.

The final step is the use of the analytical function derived from the Monte Carlo simulation to obtain dose distributions which can be used as kernel in an optimization model implemented in a TPS.

Perspectives

In future further simulations will be run in the Cresco system in order to improve the results so far obtained and to find the researched general analytical expression for a proton beam in a material other than water. When this function will be found, the analytical calculation will be validated by comparing it with the simulation results first for a simple heterogeneous medium composed of homogeneous regions, and then for a more complex CT data set.

Corresponding author: emilia.basile@iss.infn.it

BIBLIOGRAPHY

[1] A. Ferrari, P.R. Sala, A. Fassò, and J. Ranft, "*FLUKA: a multi-particle transport code*", CERN 2005-10 (2005), INFN/TC_05/11, SLAC-R-773.

[2] G. Battistoni, S. Muraro, P.R. Sala, F. Cerutti, A. Ferrari, S. Roesler, A. Fassò, J. Ranft "*The FLUKA code: Description and benchmarking*", Proceedings of the Hadronic Shower Simulation Workshop 2006, Fermilab 6-8 September 2006.

[3] Basile E. "*Integration of Monte Carlo simulation in radiotherapeutic treatment planning with proton beam*", Specialization Thesis, Università degli studi "La Sapienza", Roma, advisors Frullani S., Gigante G. E., (2010).

[4] Rosenberg, A. E., Nielsen, G. P., Keel, S. B., Renard, L. G., Fitzek, M. M., Munzenrider, J. E. and Liebsch, N. J. "*Chondrosarcoma of the base of the skull: a clinicopathological study of 200 cases with emphasis on its distinction from chordoma*" Am. J. Surg. Pathol. **23**, (1999), 1370–1378.

[5] Gragoudas, E., Li, W., Goitein, M., Lane, A. M. and Egan, K. M. "*Evidence-based estimates of outcome in patients irradiated for intraocular melanoma*" Arch. Ophthalmol. **120**, (2002), 1665–1671.

[6] "*Scientific and technogical development of Hadrontherapy*", 11th ICATPP Conference Como 5-9 October, 2009.

[7] G. A. P. Cirrone, G. Cuttone, P. A. Lojacono, S. Lo Nigro, V. Mongelli, I. V. Patti, G. Privitera, L. Raffaele, D. Rifuggiato, M. G. Sabini, V. Salamone, C. Spatola, and L. M. Valastro "*A 62-MeV Proton Beam for the Treatment of Ocular Melanoma at Laboratori Nazionali del Sud-INFN*", IEEE TRANSACTIONS ON NUCLEAR SCIENCE, VOL. **51**, NO. 3, (2004), pp 860-865.

[8] M. Chin, J. Giddy, D. Lewis, D. Walker, "*An embarassingly parallel framework for running EGSnrc/BEAMnrc/DOSXYZnrc, FLUKA, MCNP/MCNPX, GEANT4 and PENELOPE on grid and cluster computers*", ICCR 2007.

SYNCHRONIZATION IN COMPLEX BIOLOGICAL NETWORKS

Juan A. Almendral[1,2], Ricardo Bajo[1], Stefano Boccaletti[1], Javier M. Buldú[1,2], Ricardo Gutiérrez[1], Inmaculada Leyva[1,2], Adrián Navas[1], David Papo[1], Irene Sendiña-Nadal[1,2], Massimiliano Zanin[1]

[1] *Centre for Biomedical Technology, Technical University of Madrid, Spain*
[2] *Complex Systems Group, Universidad Rey Juan Carlos, Madrid, Spain*

1. Unveiling protein functions through the synchronization dynamics of the protein-protein interaction network (PIN).

One of the fundamental challenges that biologists are facing nowadays is to elucidate the function of the millions of proteins that are being added to the databases due to the great progress in sequencing the genome [1]. Knowing the protein or proteins involved in one specific molecular function is essential to understand the biological processes within the cell and to ultimately design new proper drugs to cure diseases. The problem is only going to get worse: determining the sequence of a protein becomes easier and cheaper everyday, but determining the function remains hard, slow and expensive.

The recent availability of the networks of the physical interactions between proteins has made possible to introduce computational methods to tackle this problem, allowing to predict the function of a protein based on the structure of the network of the neighboring proteins whose function is known. With this motivation, we intended to offer a completely new strategy, based on the established knowledge on the organization of synchronization properties in a modular network of oscillators [2]. The results we have obtained with the protein interaction network of the yeast are groundbreaking, insofar as they show with clear evidence how a dynamical approach to the analysis of biological networks is able to identify proteins whose function was wrongly assigned or to predict multiple functions for proteins exhibiting a particular dynamical behavior within the network (see *Figure 1*). Moreover, our approach is able to furnish a novel description of the meta-organization of a cell [3] as shown in *Figure 2*.

2. Node vulnerability under finite perturbations in complex networks.

In the study of complex networks, a very important concept is that of vulnerability of the global behavior or performance under the action of external perturbations. The classical approaches to vulnerability focus on how certain topological graph properties change after accidental (random) or intentional removal of nodes or links. In a recent work, we introduce an alternative approach, based on the *dynamics* taking place on the network [4]. Specifically, we study how fast the network dynamics abandons a collective (synchronized) state under the action of finite size perturbations applied on individual nodes, and how this depends on the centrality of the perturbed node. In simple words, the question we posed to ourselves is the following: if an attack, failure or large fluctuation in a node abruptly disrupts the collective behavior of a large ensemble of dynamical systems, how is the destructiveness of the perturbation related to the topological importance of the node in the network?

To address this issue, we consider three dynamical systems of chaotic Rössler oscillators coupled following three well-known different topologies (an Erdös-Rényi random graph, a Barabási-Albert scale-free network, and a Configuration Model scale-free network). The system is

synchronizable, and the coupling parameter is chosen so as to guarantee that the synchronized dynamics is slightly linearly unstable. We define local and global divergence rates of the perturbed to the unperturbed dynamics to measure the vulnerability of the system under the perturbations.

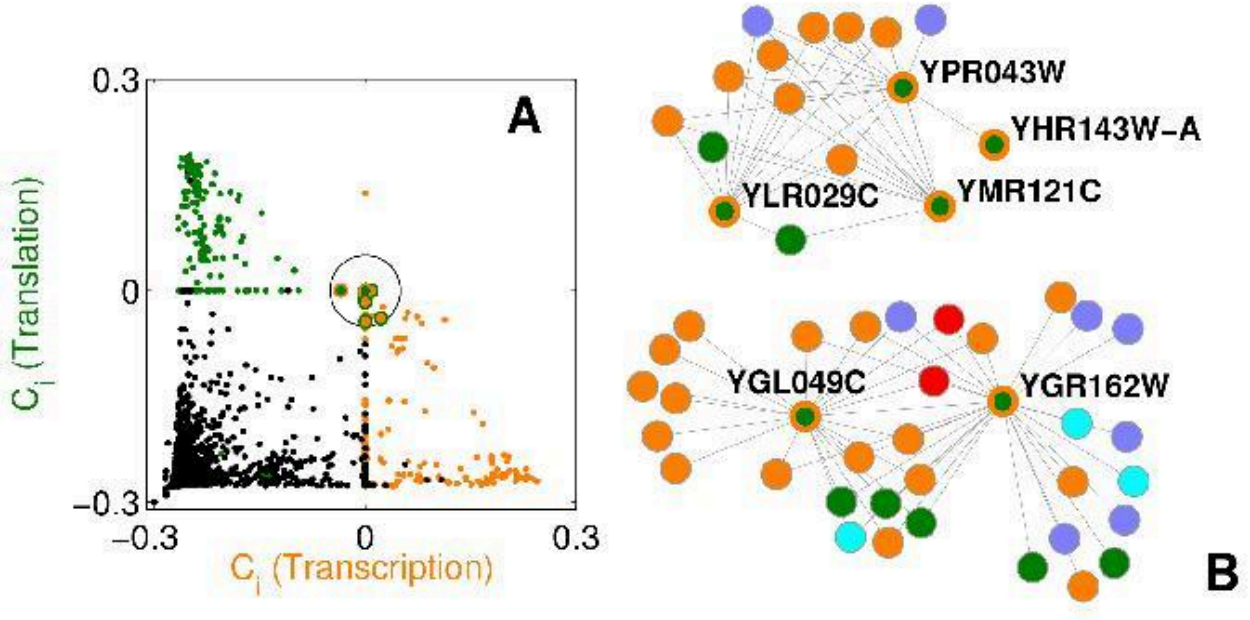


***Figure 1**: **Identification of multi-functional proteins.**(A) Dynamical indicators for the 2,049 proteins in the PIN of the yeast. The color indicates the functional module initially assigned to each protein (orange=transcription, green=translation, black=rest). The 30 proteins located inside the circle remain there after re-assignation to the predicted function, and are depicted as circles bordered with the color of that function. (B) Visualization of the network backbone made of 6 (out of 30) of the multi-functional proteins in (A).*

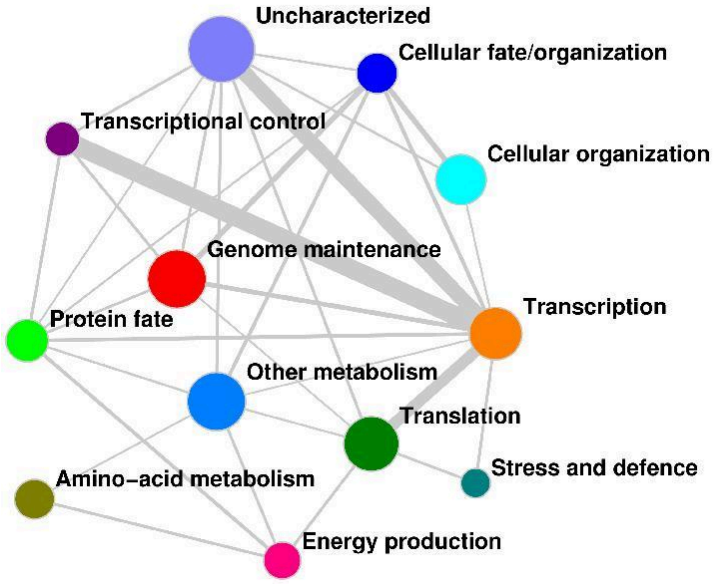


***Figure 2: Coarse grained representation of the PIN in terms of cell functioning and coordination.** The size of nodes is proportional to the total number of proteins participating to the corresponding function, the width of the links is proportional to the size of the corresponding overlapping interface.*

Surprisingly, the results turn out to be quite independent of the topological details. The global divergence rates show a very distinct yet non-trivial dependence on the centrality (see *Figure 2*), which is further defined by the degree centrality, so can be said to be largely dependent only on local (first neighbors) properties. On the one hand, the most isolated nodes (ISOL, in the figure) are the most vulnerable, whereas more central nodes (MEDIUM) are less and less vulnerable as we go along the centrality axis (degree centrality in *Figure 3 A*, eigenvector centrality in *Figure 3 B*). At some point, however, the trend is reversed, and vulnerability and centrality become positively correlated variables. The hubs of the network are more vulnerable than the nodes of intermediate connectivity.

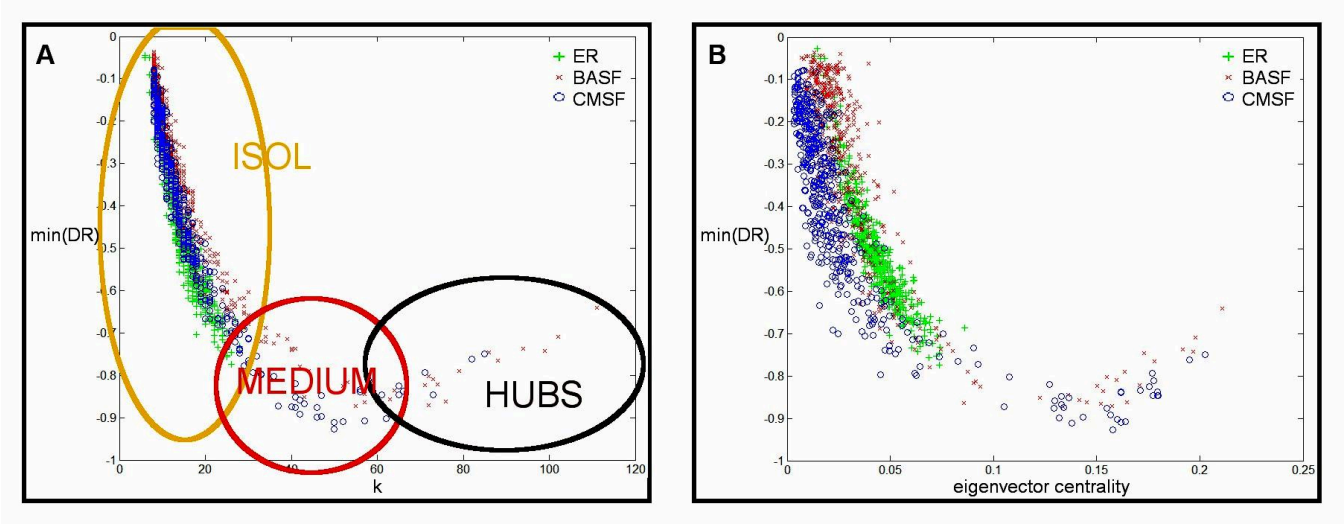


***Figure 3**: Dependence of the dynamical vulnerability on the centrality of the perturbed node. Minimum of the divergence rate as a function of degree (A) and eigenvector centrality (B) for all three networks under study.*

The mechanism underlying such dependence of the vulnerability on the centrality must be related to the way the perturbations propagate over the network. The local divergence rates were used to further elucidate the issue. The high vulnerability of the isolated nodes is seen to be due mainly to the fact that the perturbed node hardly suffers from any damping, so even though the propagation is relatively slow and irregular its effect does not dissipate and is very destructive since the beginning. For nodes of intermediate connectivity the damping becomes stronger, and, whereas the propagation is also more effective and homogeneous, still the stabilizing influence of the rest of the network prevails. For the hubs, the propagation becomes so efficient that it outweights the stabilizing effect of the many neighbors they have, and then the perturbation applied on them is more destructive than applied on nodes of intermediate connectivity. A schematic illustration is shown in *Figure 4*.

Relevant applications of our approach and results can be found in technological or infrastructural networks, where a practical issue if often to desing the better protection strategy for each on of the units to avoid the spreading over the system of an occasional breakdown or intentional attack, or in the disruption of pathological synchronized activity in the central nervous system by external action (applied or magnetically induced current pulses).

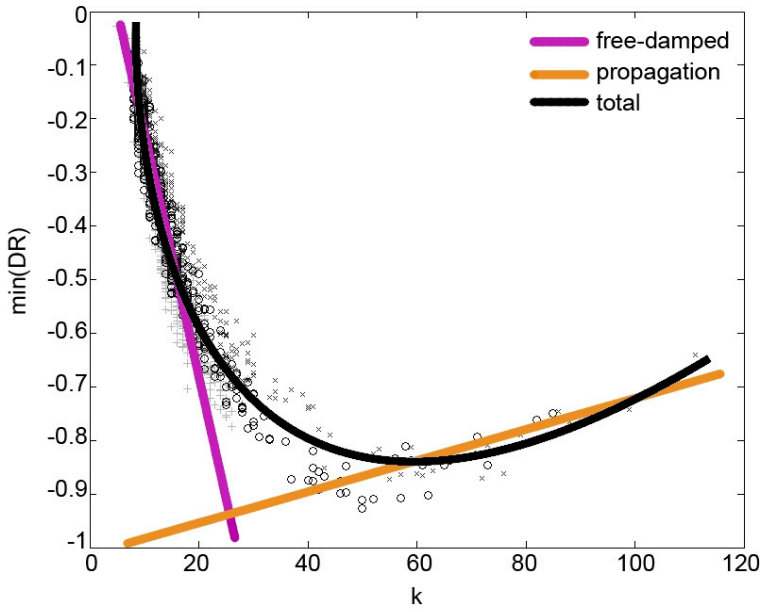


Figure 4: Interpretation of results. *The observed relationship between dynamical vulnerability and centrality is interpreted to result from the interplay between two opposing forces: damping (the stabilizing effect of the unperturbed elements in the network) and propagation (the spreading of the perturbation), assumed to vary in importance across the centrality axis roughly as schematically shown in the figure.*

3. Emergence of structural patterns in networks of synchronized oscillators with competitive interactions.

Synchronization occupies a privileged position in the study of emergent collective phenomena because of its ubiquity in the natural world. Furthermore, many real world networks, such as biological neural networks or social networks, are known to be adaptive: links between nodes are enhanced or depressed depending on the dynamics. Hence our interest in the structural properties emerging from synchronized dynamics in adaptive networks.

To address this issue, we study a weighted network of Kuramoto oscillators topologically regular whose coupling weights evolve in time following two opposing adaptive mechanisms [5]. The first one, known as Hebbian learning in neuroscience and as homophily in sociology, consists in the reinforcement of link weights for dynamically correlated neighboring oscillators. As reinforcement without bound is not realistic (for instance in networks of neurons, the enhancement of some synapses is compensated by the weakening of others), the opposing mechanism, homeostatic plasticity, is also implemented in the weight equations. Two parameters are considered in the system equations: the coupling strength, λ, and the time of integration, T, of the dynamical correlation.

As soon as the adaptive mechanism is switched on, the global synchronization is rapidly enhanced in the network. Most interestingly, after the system is allowed to reach a stationary state, for a large combination of parameters λ and T, between the two extreme phases of the classical Kuramoto model (incoherent evolution of the oscillators, global synchronization), there emerges through adaptation another phase corresponding to a large region of parameter space in which the global synchronization is still low, but the local synchronization of each oscillator with

its neighbors is almost perfect (see *Figure 5*). In this phase, a detailed analysis of the network resulting from adaptation reveals a distinct self-organization of the network into communities of highly synchronized oscillators (*Figure 6*). Moreover, a distinct scale-free distribution of the weights is found for small λ, which gradually turns into a homegeneous distribution for increasing coupling (*Figure7*).

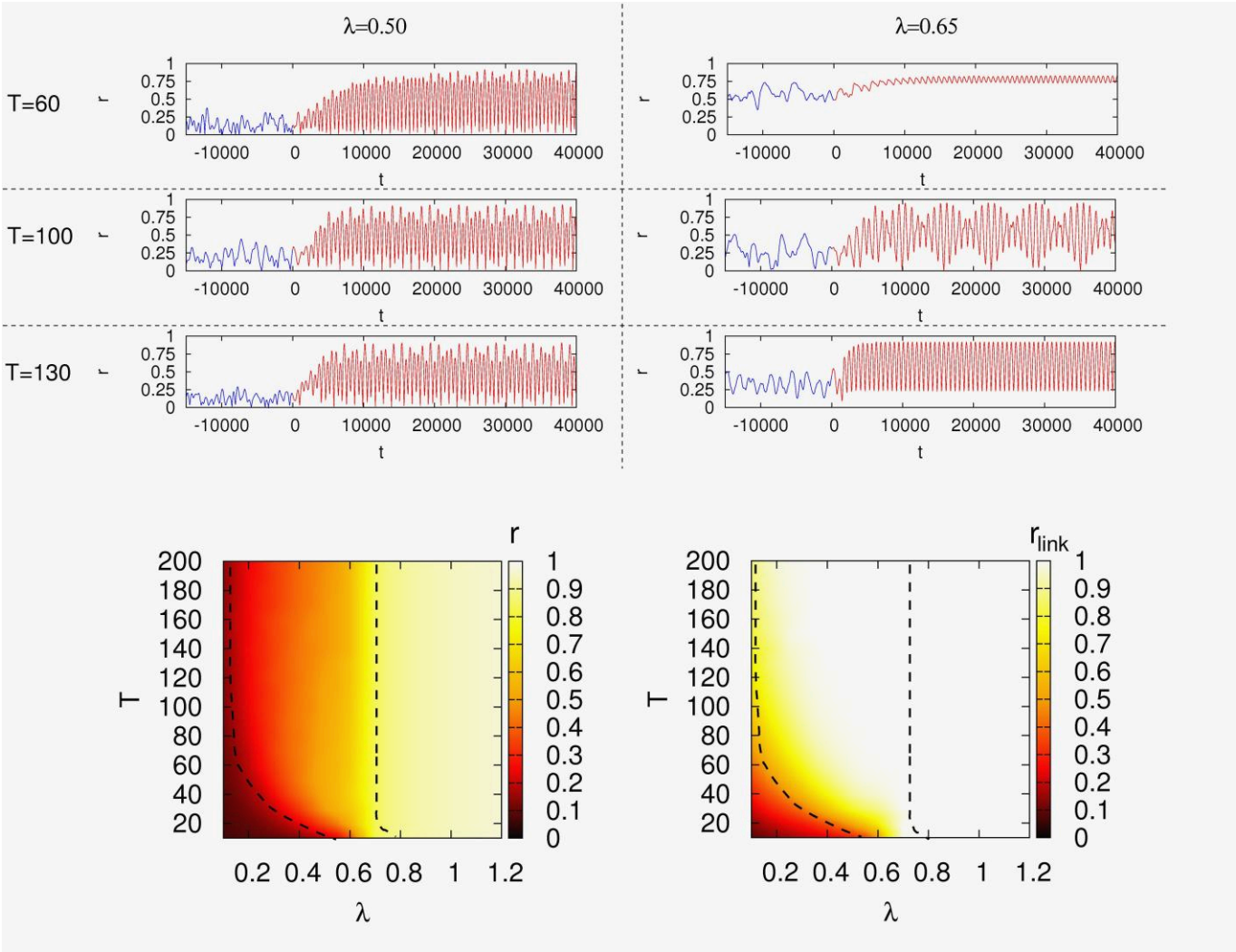


Figure 5: Global and local synchronization. *In the top panels, the global synchronization, r, (Kuramoto order parameter) is shown as a function of time for several coupling strengths λ and integration times T. At time t = 0 the adaptive mechanism is switched on. Below, the mean global synchronization and local synchronization, r_{link}, are shown as a function of both parameters. The region bounded by the two dashed lines correspond to the phase of high local synchronization and low global synchronization.*

In conclusion, as a result of adaptation, two properties, such as modularity and scale-freeness, that have been observed in many real world networks, are seen to emerge from synchronization. Preliminary work with a similar model suggest that the emergence of these properties remains essentially unaltered under changes in the model equations as long as the weight equations implement the interplay between the two opposing adaptive mechanisms.

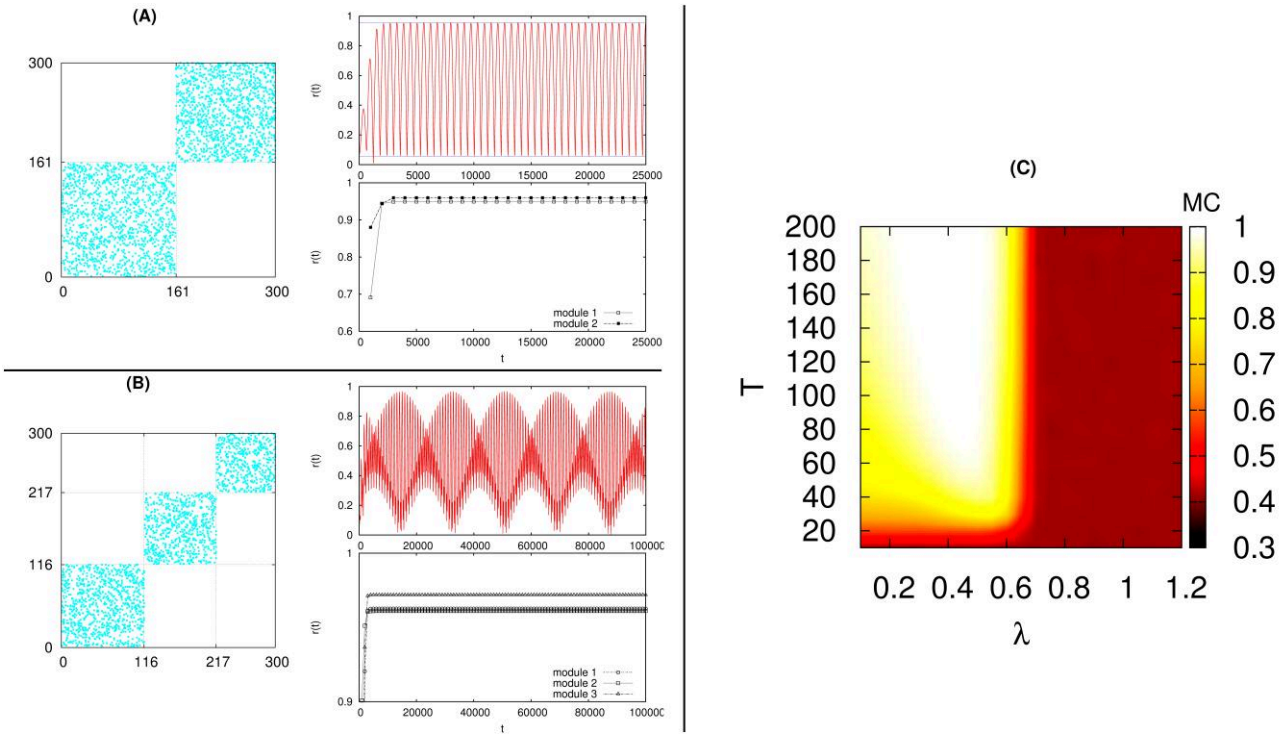


Figure 6: Modularity as a result of adaptation. *The network is seen to be clusterized into communities of tightly synchronized oscillators in the phase of high local synchronization and low global synchronization (on the left, two communities are shown in (A) and three in (B)). The modularity in the network as measured by the modular cohesion throughout parameter space is shown in (C).*

4. Synchronization of neurons in geometrical complex network.

Recently, the study of complex networks has been devoted a great deal of attention, since an increasing number of real world systems have been seen to be successfully described by them. However, most of these studies do not consider the fact that, in many cases, real complex networks are embedded in geometric space. Examples can be found in all scientific and technical areas, such as the Internet, power grids, epidemic spreading over contact networks, social relationships or neural connectivity. Very recent studies include space-dependent generative models of networks to explain the origin of experimentally observed connectivity in biological neural networks, specifically in the topology of the *C. elegans* neural network.

The additional special features that the network acquires when the spatial positions of nodes is considered are relevant for the global properties of the network, but they have been studied on very few occasions. For instance, the impact of modularity in the behavior of complex networks is receiving much attention on the part of scientists, but the fact that in most cases the formation of modules has its origin in spatial proximity is not properly considered. As for dynamics on networks, synchronization has proved to be a most relevant kind of collective dynamics in complex network as a mechanism for information transmission. Again, so far very few works have focused on the importance of the spatial structure for the dynamical behavior of complex systems in general, and their synchronization in particular.

In this work we study the synchronization properties of a spatial network of neurons [6]. We consider the topology of the network generation spatial model, its dynamical properties, and the relation between topology and function. We observe the emergence of synchronization waves in the system as a distinct dynamical feature of the spatial network (*Figure 8*).

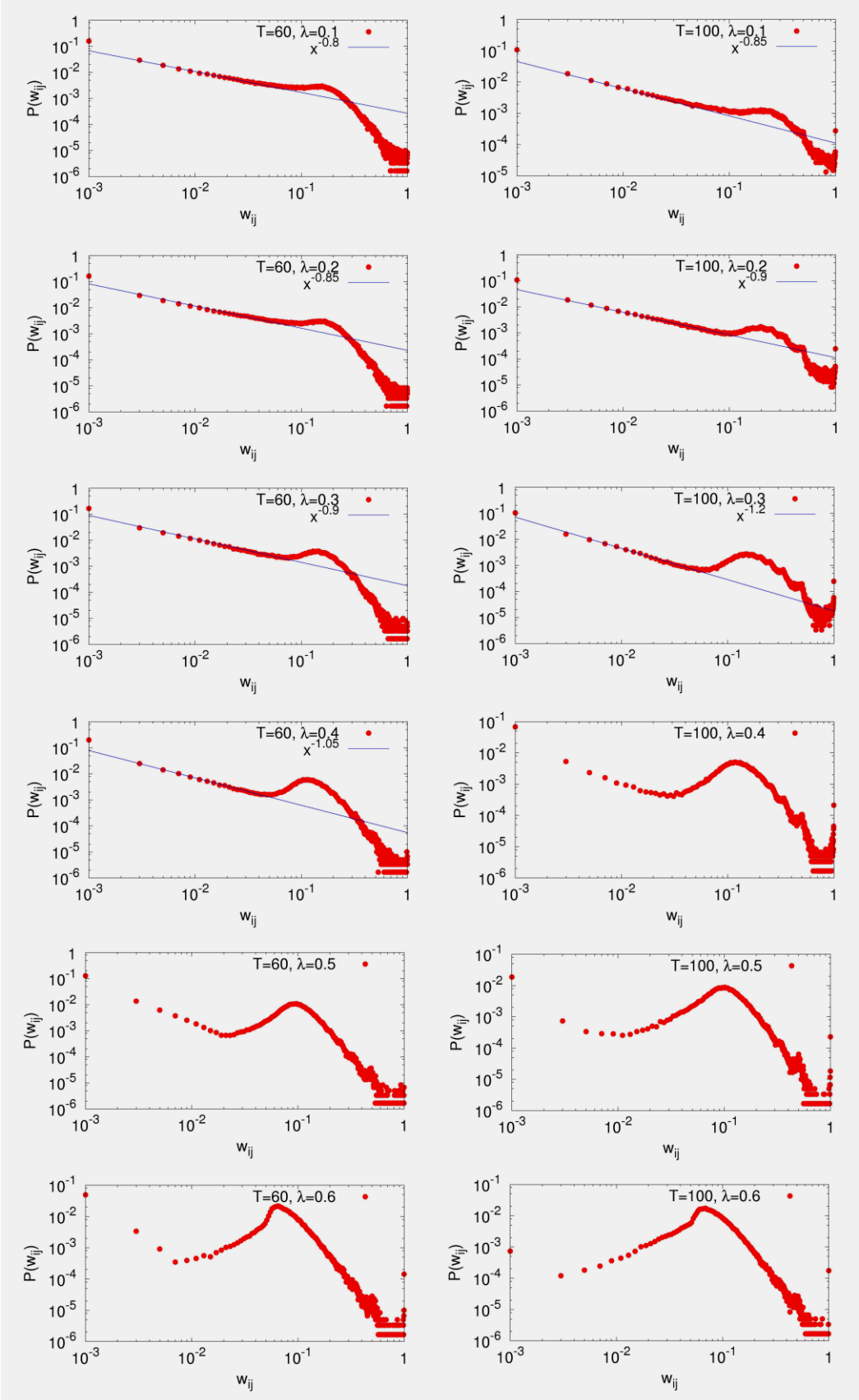


Figure 7: Weight distribution resulting from adaptation. *Link weight distribution for integration time T = 100 and different coupling strengths λ. For weak coupling, the distribution is scale-free. For increasing coupling, a maximum emerges and the distribution gradually changes into a homogeneous network for very high λ.*

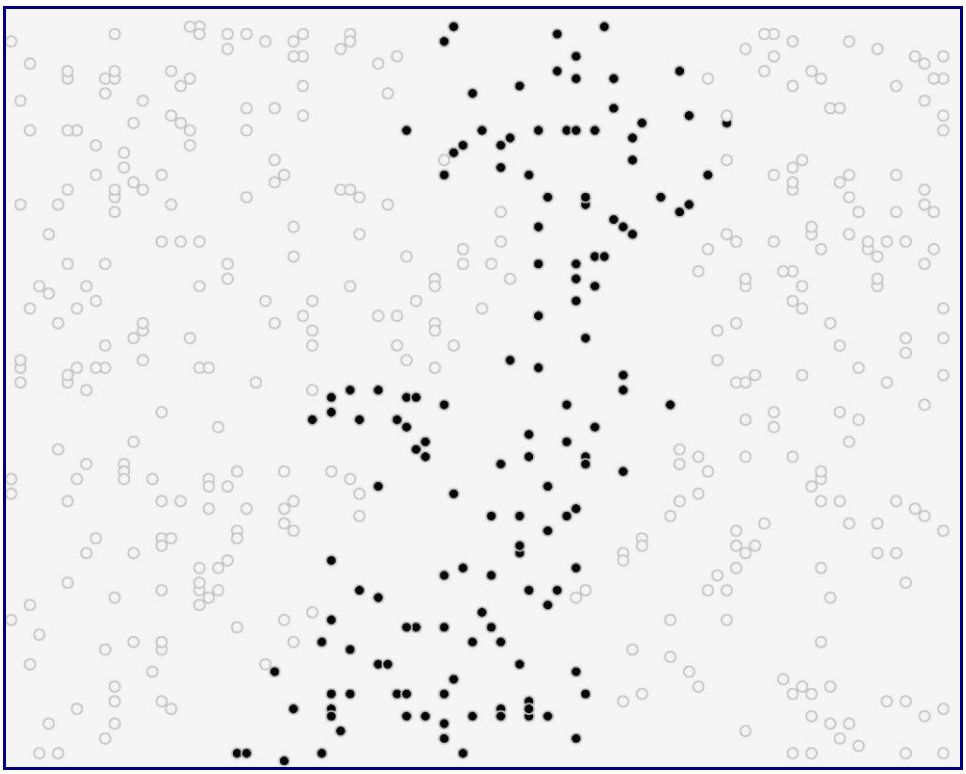


Figure 8: Apparition of waves in the system. *Over a sample networks generated with N=500, < k >= 10 and connectivity range σ=5, instantaneous frame of the activity when coupling constant d = 0.2. The inactive neurons are empty circles, whereas full circles represent the spiking neurons.*

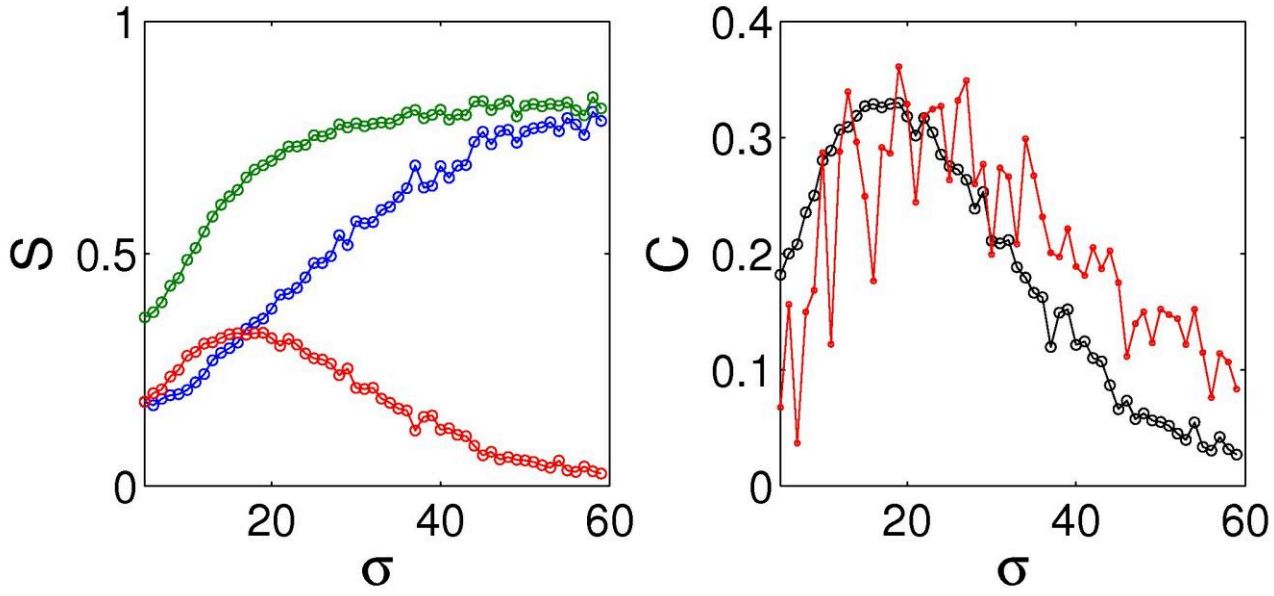


Figure 9: Correlation between dynamics and topology. *(Left) S (blue), Sr (green; local synchronization) and Sw=Sr-S (red) vs. σ (range for connections) plot for coupling constant d= 0.2 and < k >=15. (Right) Average correlation between the topological and functional centrality vectors. We see that there is a maximum correlation when Sw is maximum, that is, with synchronization waves dynamics.*

These waves allow a global mapping of the network and the search for relevant topological information from dynamical activity *(Figure 9*). Additionally, we study the robustness of the model to external perturbations and we try to characterize neurons according to the dynamical behavior observed.

REFERENCES

[1] PUNTA M., OFRAN Y. (2008) *"The rough guide to in silico function prediction, or how to use sequence and structure information to predict protein function."* PLoS Comput Biol 4: e1000160.

[2] LI D., LEYVA I., ALMENDRAL J.A., SENDIÑA-NADAL I., BULDÚ J.M., HAVLIN S., BOCCALETTI S. (2008) *"Synchronization interfaces and overlapping communities in complex networks."* Phys Rev Lett 101: 168701.

[3] SENDIÑA-NADAL I., OFRAN Y., ALMENDRAL J.A., BULDÚ J.M., LEYVA I., LI D., HAVLIN S., BOCCALETTI S. (2011) *"Unveiling Protein Functions through the Dynamics of the Interaction Network."* PLoS ONE 6: e17679.

[4] GUTIÉRREZ R., DEL-POZO F., BOCCALETTI S. "*Node Vulnerability under Finite Perturbations in Complex Networks."* PLoS ONE (in Press).

[5] ASSENZA S., GUTIÉRREZ R., GÓMEZ-GARDEÑES J., LATORA V., BOCCALETTI S. "*Emergence of structural patterns out of synchronization in networks with competitive interactions"* (in Preparation).

[6] LEYVA I., NAVAS A. *"Synchronization of neurons in geometrical complex network"* (in Preparation).

BIOPHYSICS AT THE ATOMIC SCALE

Lorenzo Bongini

Small Biosystem Lab, Departament de Fisica Fondamental
Facultat de Fisica, Barcelona

Triggering protein aggregation: a matter of vibrational entropy

Aggregation, the formation of large supramolecular aggregates, is a rather common phenomenon for all sorts of polypeptide chains. Importantly, severe neurodegenerative diseases such as Alzheimer's disease, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis and prion diseases are all associated with the accumulation of large amounts of aggregates in the brain. Moreover, it is increasingly being realized that such pathologies have common cellular and molecular mechanisms [1]. In particular, these seem to be linked to the early stages of the aggregation process, where elementary, cytotoxic proto-fibrilar assemblies of different nature start forming [2].

Many natural proteins tend to aggregate, typically following a nucleation mechanism, in a temperature range that might extend from the temperature of thermal unfolding down to their physiological temperature of exercise [3].

The generality of the process is further reinforced by the remarkable structural uniformity of the final aggregation products, that, in most cases, share the same β-spine architecture characteristic of amyloid fibrils: an assembly of beta sheets perpendicular to the fibril axis.

The lack of clear signatures of the aggregation propensity at the sequence level [4] suggests that the phenomenon should be governed by rather general laws, uniquely relied to the dynamical properties of poly-peptide chains in solution. In particular, in view of the rather slow time-scales characteristic of aggregation.

In order to investigate the above hypothesis, in the dependence of vibrational entropy on the content of secondary structure in a large database of protein structures has been investigated using the all-atom molecular force field AMBER [5]. The crystal strucctures where downloaded in from the Protein Data Bank and minimized with the sander module in implicit solvent and finally the hessian in the minimun of the energy was computed and diagonalized with the nmode module providing the vibrational spectrum needed to compute vibrational entropy. Our results suggest that vibrational entropy of protein fluctuations about their native structures are correlated with their content of secondary structure. Our main point is that such observation might be re-interpreted in the context of aggregation as a postulate for the existence of a thermodynamic force guiding the aggregation process towards β-rich structures: according to our finding, this might reflect the effort of the developing aggregate to maximise its vibrational entropy.

In view of the positive correlation existing between the α and β contents, we shall consider in the following the correlation of entropy with a global structural indicator defined as (β -content minus α-content), which we shall refer to as β-preference. Positive correlation of this quantity with entropy leads to a negative correlation with free energy at high temperatures, that will in turn favour the formation of β-like structures. Importantly, in order to compare proteins of different lengths, we shall always consider intensive observables, namely entropies per residue.

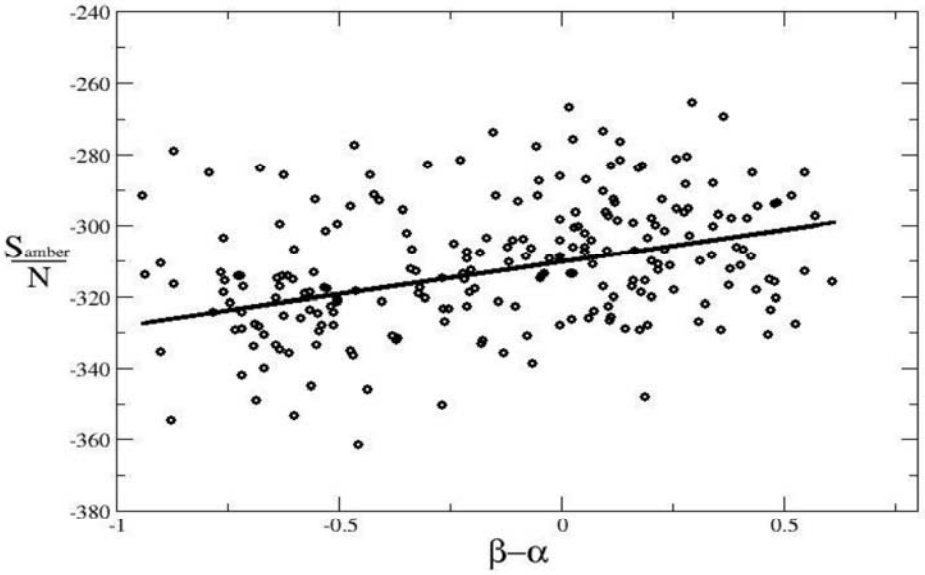


Figure 1: *Correlation of β-preference (β-content minus α-content) with vibrational entropy per residue S/N in the PDB select as calculated at the all-atom level with the AMBER force-field.*

Figure 1 shows that the vibrational entropy computed using an atomistic force field has a slight positive correlation with β-preference (correlation coefficient=0.30). Unfortunately, the high complexity of a full atomistic description makes it extremely difficult to isolate the principal physical and structural properties at the core of the observed correlation. Therefore, in order to simplify the analysis, we proceeded by introducing different sources of structural and dynamical details starting from a simple coarse-grained elastic network description of residue-residue interactions.

The model chosen is an evolution of the Anisotropic Network Model (ANM) originally proposed by Tirion [6] at the all-atom level and successively reconsidered by Bahar and co-workers in the C_α coarse-grained approximation [7]. In particular we chose to investigate the effect of three elements of detail to the basic ANM modeling scheme:

1) additional degrees of freedom corresponding to coarse-grained representations of side chains,

2) a more realistic hierarchy of force constants, reproducing the (at least) three different strengths of bonds featured in a protein molecule, i.e. covalent bonds (such as the peptide bond along the chain), hydrogen bonds and long-range forces such as hydrophobic or screened electrostatic interactions,

3) appropriate values for the amino-acid masses, correctly reproducing the protein sequence.

Only if all level of details are introduced the resulting Hierarchical Network Model (HNM) shows to reproduce the observed correlation with the β-preference while at the same time correlating with the atomistic molecular entropies (*Figures 2 and 3*).

For further details refer to [8].

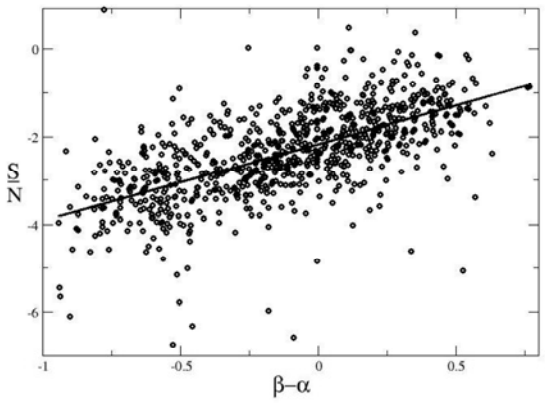


Figure 2: *Correlation of the total entropy per residue S/N with β-preference (β-content minus α-content) while switching on different interactions in the HNM C_α model.*

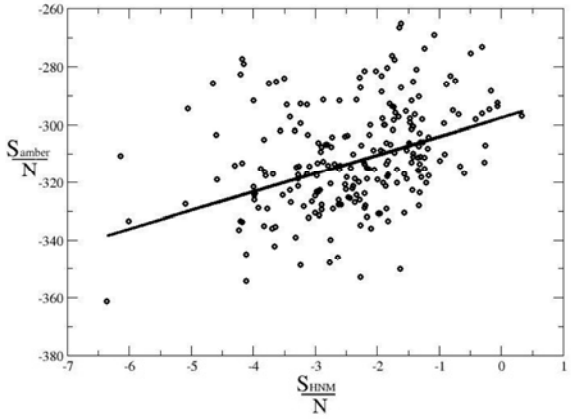


Figure 3: *Vibrational entropies as calculated within the all-atom description of the structures versus results of the full HNM coarse-grained model (weighted HNM with side-chains).*

Modeling DNA overstretching

In the last decade optical tweezers and atomic force microsopes have allowed to develope a brand new field of biophysical investigation: single molecule force spectroscopy. By measuring the elastic response and the breaking force of single biomolecules under tension one can get an insight on their free energy landscape, at least along the reaction coordinate which describes the stretching of the molecule (generally the end-to-end-distance). One of the first molecules to be investigated with these techinques has been DNA. The elastic properties of DNA have been characterized at the end of the last century and toghether with the characterization of the entropic elasticity opposing longitudinal stretching another phenomenon became appearent: DNA overstretching. When stretched above 60 pN double stranded DNA molecules whose extremities are free to rotate undergo a highly cooperative phase transition to an ovestretched form which is characterized by an average rise (distance between consecutive base pairs) 70% bigger than in standard B-DNA and an elicity of 1 turn every 30 base pairs instaed of the customary 1 every 10 [9-10]. It has long been debated whether this new conformation is just partially melted DNA or has a more rigorously defined structure. Early molecular dynamics studies suggested an "S-ladder" structure where base pairing is conserved and the traditional elix is flattened.

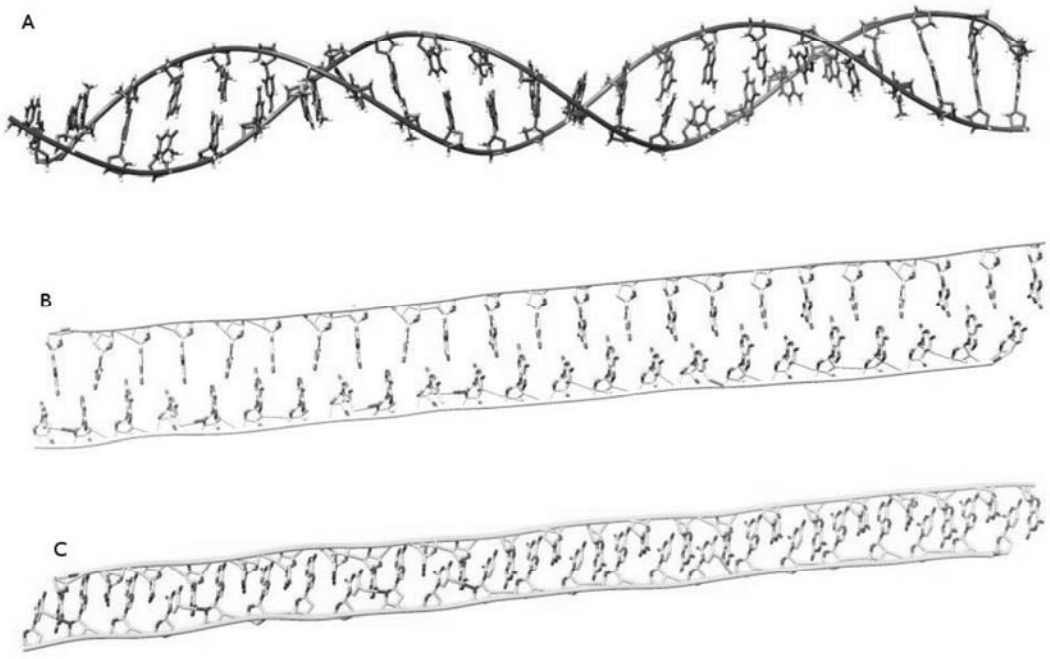


Figure 4: *Different starting conformations for the constraint molecular dynamics runs.*

Several constrained molecular dynamics symulation of a double stranded portion of DNA of 24bp were performed in AMBER starting from different competing conformations depicted in *Figure 4*. The constraint was set on the end to end distance of each strand separately so as to allow strand separation while maintaining the total molecular extension compatible with S-DNA. Configuration A corresponds to a helix with the strandard helicit of B-DNA (10 bp per turn) with the same rise (distance between consecutive base pairs) as overstretched DNA (0.58 nm). Configuration B corresponds to a S-ladder (no helicity and rise of 0.58 nm) while configuration C (which we will call the S-zipper) is characterized by having each strand in the same configuration as in the S-ladder but the two strands shifted and brougth nearer so as to stick each nitrogen base between two of the opposing strand. The S-zipper conformation maximizes base staking which is absent in the other two conformations.

Simulation were performed for 10 ns in implicit solvent with the generalized Bore/sufrace area model at a ionic strenght of 100 mM using the f99 force field. Temperature was set at 300 K by means of a Langevin dynamics. A quick comparison between the configuration attained after 10 ns shows that the dynamics is still far from comverging to a unique conformation (see *Figure 5*).

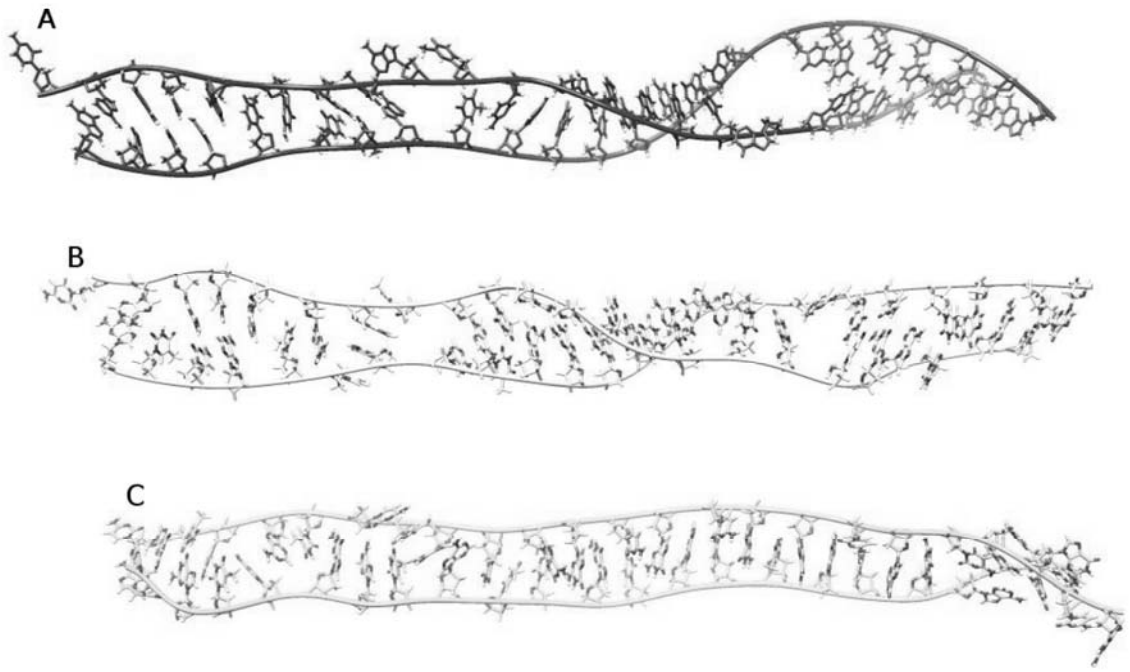


Figure 5: *Snapshots of constrained MD trajectories corresponding to the inital configuration of Fig. 4 after 10 ns.*

The results, however, share some remarkably homogeneus structural features that to be conserved in the final S-DNA conformation. After 10 ns, regardless of the starting structure, DNA had attained a conformation which is characterized by low helicity and the lack of regular base pairing or base staking pattern. Even in case where no base staking between nitrogen bases belonging to different strands was initially present, as in the caser of the starting configuration A and B, some staked pairs can easily be spotted in the final configuration. On the contrary base pairing is efficiently disrupted especially when starting from the S-ladder configuration. Base-pair disruption, however does not seem to lead to systematic melting, although some denaturation bubbles show up in various spots in along the different final configuration. The two opposite strands appear, however, to remain united probably due to the presence of inter-strand hydropobic or Van der Vaals interations between opposite nitrogen bases. Such interaction might rise from a well formed staking or more frerquently from an aspecific contatc between the hydrophobic sides of teh nitrogen bases.

These preliminary results suggest therefore yet another interpretation of the overstretched state of DNA: an unstructured conformation of essentially two single strands held toghether by staking and Hydrophobic interactions. Undoubtedly further simulations are needed to confirm this hypothesys. Constrained simulations at intermediate molecular extensions have already been performed whith the aim of veryfying, by means of umbrella sampling simulations, that the suggested conformation correctly reproduces the experimental free energy profile characterizing the B-S transition in DNA.

BIBLIOGRAPHY

[1] M. Bucciantini, E. Giannoni, F. Chiti, F. Baroni, L. Formigli, J. Zurdo, N. Taddei, G. Ramponi, C.M. Dobson, M. Stefani *"Inherent toxicity of aggregates implies a common mechanism for protein misfolding diseases"* Nature **416** (2002) pp. 507-511.

[2] H.A. Lashuel *"Membrane Permeabilization: A Common Mechanism in Protein-Misfolding Diseases"* Sci. Ageing Knowl. Environ. **38** (2005) p. pe28.

[3] J.F. Smith, T P.J. Knowles, C.M. Dobson, C.E. MacPhee, M.E. Welland *"Characterization of the nanoscale properties of individual amyloid fibrils"* PNAS **103** (2006) pp. 15806-15811.

[4] S.Ventura *"Sequence determinants of protein aggregation: tools to increase protein solubility"* Microbial Cell Factories **4** (2006) pp. 11.

[5] J.W. Ponder, D.A. Case, V. Daggett *"Force Fields for Protein Simulations; Advances in Protein Chemistry; Protein Simulations"* Advances in protein chemistry **66** (2003) pp. 27-85.

[6] M.M. Tirion *"Large Amplitude Elastic Motions in Proteins from a Single-Parameter, Atomic Analysis"* PRL **77** (1996) pp. 1905-1908.

[7] P. Doruker, A.R. Atilgan, I. Bahar *"Dynamics of proteins predicted by molecular dynamics simulations and analytical approaches: Application to alpha-amylase inhibitor"* Proteins: Structure, Function, and Genetics **40** (2000) pp. 512-524.

[8] L. Bongini, F. Piazza, L. Casetti, P. de LosRios *"Vibrational entropy and the structural organization of proteins,"* Accepted in: The European Physical Journal E - Soft Matter & Biological Physics, 2010.

[9] S.B. Smith, Y. Cui, C.P. Bustamante *"Overstretching B-DNA: the elastic response of individual double-stranded and single-stranded DNA molecules"* Science **271** (1996) pp. 795-799.

[10] J.F. Léger, *et al. "Structural Transitions of a Twisted and Stretched DNA Molecule"* PRL **83** (1999) pp. 1066

METEOROLOGICAL AND AIR QUALITY FIELDS PRODUCTION OVER ITALY IN THE FRAME OF THE MINNI PROJECT

Gino Briganti[1], Andrea Cappelletti[1],
Luisella Ciancarella[2], Giandomenico Pace[3], Lina Vitali[2], Gabriele Zanini[4]

[1] ENEA-UTVALAMB-AIR. Territorial Office of Pisa
[2] ENEA-UTVALAMB-AIR. Bologna Research Centre
[3] ENEA-UTMEA-TER. Casaccia Research Centre
[4] ENEA-UTVALAMB. Bologna Research Centre

Introduction

Since 2002, on behalf of the Italian Ministry of the Environment, ENEA has been leading a national Project, named MINNI (National Integrated Modelling system for International Negotiation, www.minni.org, [1]), for the development of an Integrated Assessment Modelling system. The objective of the project is to support policy makers in the elaboration and assessment of air pollution policies at international, national and local level, by means of the more recent understandings of the atmospheric processes.

The MINNI Project consists of two main components: a multi pollutant Eulerian Atmospheric Modelling System (AMS), and the GAINS-Italy Integrated Assessment Model (Greenhouse Gas and Air Pollution Interactions and Synergies), developed within a joint research project ENEA-IIASA (International Institute for Applied Systems Analysis).

The AMS includes the meteorological model RAMS [2], pollutants emission processor EMMA [3] and the chemical transport model FARM [4]. The first AMS version has been applied to estimate deposition and air concentrations fields over Italy, with a spatial resolution of 20x20 km^2 and a hourly time step; the 1999 was selected as reference year. In 2005 the Italian Ministry of the Environment commissioned to ENEA a new AMS version with a spatial resolution of 4x4 km^2 that was applied to years 1999 and 2005.

In July 2008 a further 3 years agreement between the Italian Ministry of the Environment and ENEA was signed. According to this new agreement, deposition and air concentrations fields over Italy will be calculated for years 2003 and 2007. Moreover this agreement foresees and schedules several improvements of AMS that are important within the perspective of the implementation of the new air quality directive (2008/50/EC).

The meteorological and air quality simulations (for one year on the whole Italian domain with a hourly time step and a spatial resolution of 4x4 km^2) demand huge computational and storage resources. So it was chosen to take advantage of the availability of the ENEA-GRID infrastructure; in particular more recently CRESCO HPC system provides all the computational and storage resources that MINNI project needs.

Several ENEA-GRID users perform activities in the frame of MINNI projects. In particular Gino Briganti carries out the chemical-dispersion simulations, using as inputs the pollutants emissions data preprocessed by Andrea Cappelletti and the meteorological data calculated by Giandomenico Pace and Lina Vitali. A shared code and data organization has been set up and several dedicated shell procedures have been developed for the automatic execution of yearly runs.

In particular from May 2009 to June 2010 all meteorological runs have been completed for both years 2003 and 2007. As far as it concerns air quality simulations, several tests have been carried out in order to test different code versions and identify the more numerical stable one. Moreover pollutants concentration fields have been calculated for year 2003.

Producing and evaluating meteorological fields

The 2003 and 2007 meteorological fields at national scale have been reconstructed by means of the prognostic non-hydrostatic model RAMS [2] and in order to realize an hourly representation of meteorological fields over the whole Italian peninsula, with a resolution of 4x4 km^2, the target domain has been subdivided into 5 sub-domains including respectively Sicily, Sardinia and Northern, Central and Southern Italy. Initial, boundary conditions and data assimilation have been based on mesoscale analyses produced by mean of the RAMS pre-processor ISAN (ISentropic ANalysis). Nudging technique has been employed by RAMS to assimilate data analyses during the whole model simulation.

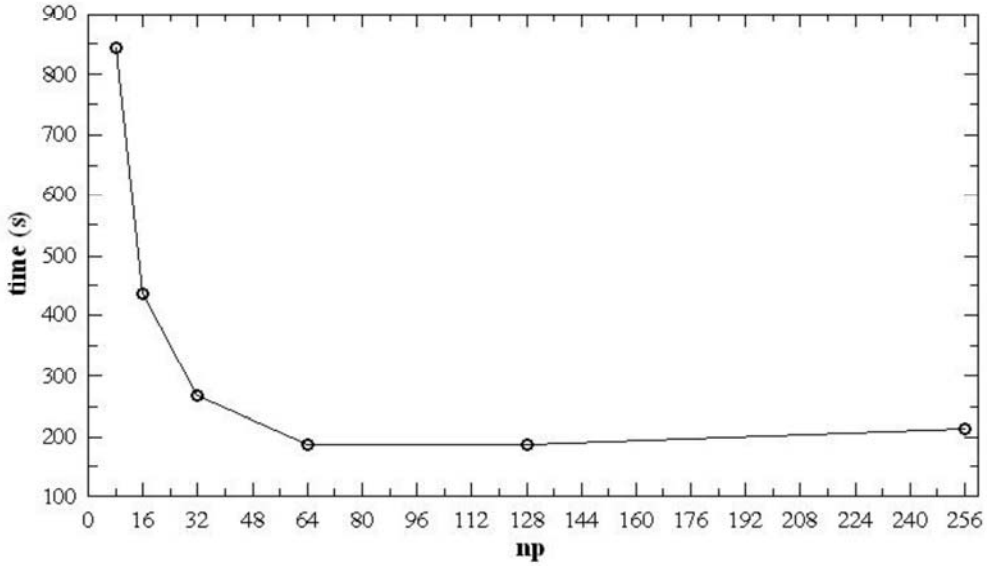


Figure 1: *Execution time as function of the processor number.*

RAMS code was compiled in its parallel version and implemented in MVAPICH environment. In order to identify the optimal processor number minimizing execution times, some different processor numbers were tested and 64 was chosen as optimal one (see *Figure 1*).

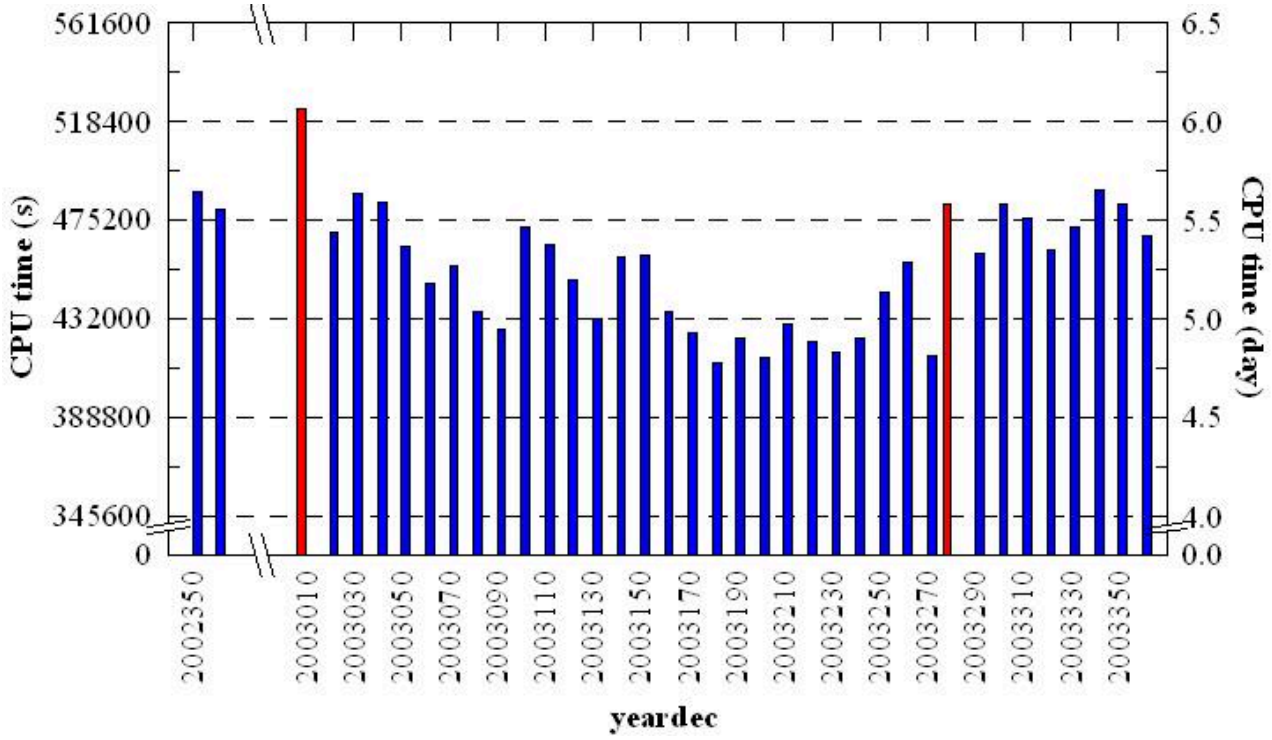


Figure 2: *Jobs CPU times for an yearly meteorological simulation. Blue means that job was dispatched on cresco2 machines, red means that job run on cresco1.*

In *Figure 2* jobs CPU times for an yearly meteorological simulation are shown. The yearly run was divided into 10 days simulation jobs. Parallel jobs were principally dispatched on cresco2 machines. Only few jobs run on cresco1, with slightly longer CPU times.

In *Figure 3* an example of calculated meteorological fields is shown. As a first test of RAMS results, calculated fields were compared, over all the Regional sub-domains, with monthly and annual indicators calculated by Italian Superior Institute for Environmental Protection and Research (ISPRA) in the framework of SCIA database (http://www.scia.sinanet.apat.it). Comparison was made for main meteorological variables and a general good agreement between observed and simulated data was observed.

During 2010 a systematic validation of all meteorological fields (1999, 2003, 2005, 2007) calculated in the frame of MINNI project was planned and observed data were collected from Italian Regional Meteorological Networks. An exhaustive validation of 2005 MINNI meteorological fields has been carried out on meteorological data set from Regional Center of Environmental Modeling of ARPA Friuli Venezia Giulia. Validation results were presented in June 2010 at the 13th International Conference on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes [5].

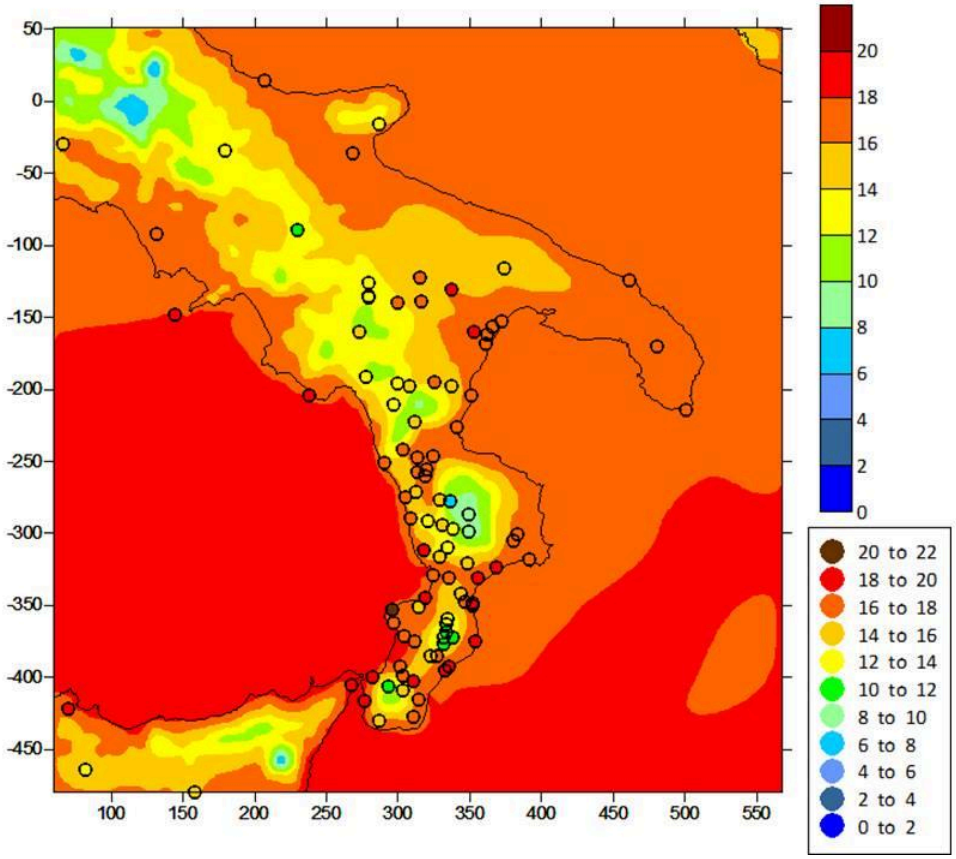


Figure 3: *Annual mean temperature (°C): comparison between calculated fields (4x4 km^2 resolution) and observations from the ISPRA-SCIA database (colored symbols).*

Inventory and emission pre-processing: pollutants emission processor EMMA

The used anthropogenic emission inventory has been derived from the emissions from major point sources and from the diffuse sources at provincial level provided by national emission inventories [6]. These emissions are classified according to activity level CORINAIR/SNAP (CO-oRdinated INformation on the Environment in the European Community AIR / Simplified Nomenclature for Air Pollution). EMEP emission inventories for years 1999 and 2005 have been employed to describe the anthropogenic sources located in other countries included in the computational domain. The biogenic emissions had also two sources: APAT 2000 over Italy and a global Guenther database [7] for the other countries. The emissions fields also include the maritime activities, the ship emissions on the national and international sources and the port areas. The diffuse emissions and the minor point sources are distributed in the lowest model layers of FARM (below 50 m) with 80% in the first 20 m above ground. The point sources such as industries, power plants, volcanoes, etc., are treated individually in FARM, considering the plume rise effects.

The time modulation has been accounted for all sources as well: we assigned at each SNAP activity its characteristic modulation curve, allowing the determination of hourly mean emission. We provided three curves for each SNAP sector, representing daily, weekly and monthly variations.

With respect to the past simulations, we introduced two other sources of particulate matter (PM): the re-suspension due to road and highway traffic and the emissions from farmlands soils; for the last one, we used an estimate of the emission factors performed through a targeted measurement campaign [8].Moreover, we changed the spatial disaggregation of vehicular emissions, by taking more account of the population distribution through some suitable weights; in such a way, traffic emissions nearby urbanized area look more realistic.

A map of NO_2 emissions, averaged on the whole 2005 year, has been depicted in *Figure 4*, where we can see the impact of both most relevant urban areas and highways.

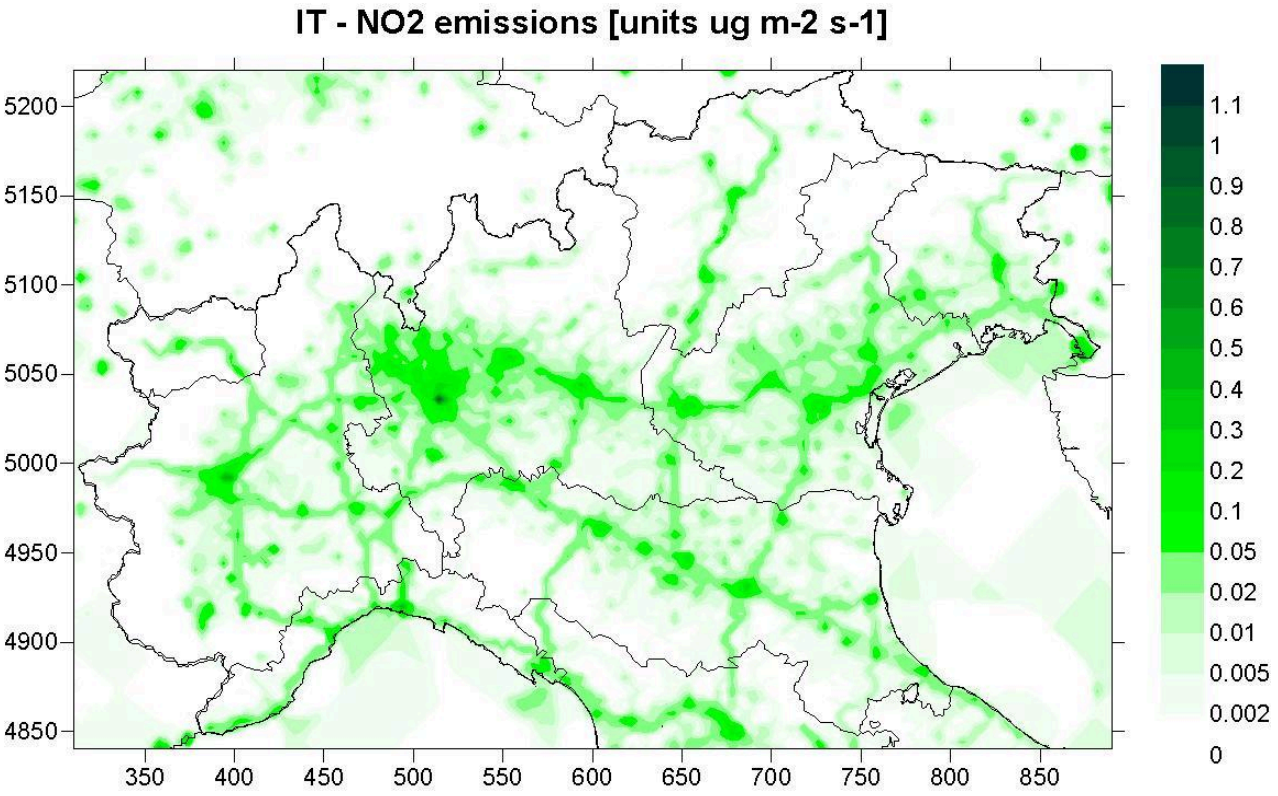


Figure 4: *Yearly averaged (2005) emissions of NO_2.*

Application of the chemical dispersion model FARM

We used the model FARM (Flexible Air quality Regional Model) [4], which is a three-dimensional Eulerian model that accounts for the transport, chemical conversion and deposition of atmospheric pollutants, by assuming a K-type turbulence closure. The code has been derived from STEM [9]. Gas-phase photochemical reactions are described by means of SAPRC-90 chemical scheme [10] and aerosol dynamics is described with the modal AERO3 scheme, which is the same module implemented in the Community Multiscale Air Quality (CMAQ) modeling system [11].

For both 2003 and 2005 years, we conducted one simulation over the whole Italy (IT) with a horizontal resolution of 20x20 km^2 and 16 vertical levels up to 10 km, densest near the ground. Both the initial and boundary conditions have been provided by EMEP fields at 50 km horizontal

resolution [12]. Next, we have conducted a nesting from 20 to 4 km over five subdomains: north of Italy (NI), centre of Italy (CI), south of Italy (SI), Sardinia (SA) and Sicily (SC).

We simulated both 2003 and 2005 years. The emission fields for 2003 year was achieved through a suitable interpolation from 1999 and 2005 inventories. The turbulence fields and the other boundary layer parameters were calculated starting from RAMS outputs.

The yearly run was divided into 30-day long simulation jobs. Actually, it was realized that the 30-day period is long enough to neglect transient regimes, and short enough to allow a sort of "parallelization" because we have twelve independent jobs.

With a view to future simulations, aimed to determine Atmospheric Transfer Matrices [13], which are computationally very relevant, we decided to compile the FARM source code with the OpenMP directives and test it before to perform 2003 and 2005 simulations. The *Figure 5* shows the trend of maximum CPU consumed for each processor to simulate 1 day period; typically, the CPU demand is maximum (about +15%) for intense photochemical activity, i.e. strong UV radiation. From *Figure 5* we see that the optimal choice is to use 4 processors; in fact, such a choice guarantees both a significant reduction of the specific CPU time (about 2.4 times) and the employment of a limited amount of resources, thus ensuring a greater chance of accessing free resources.

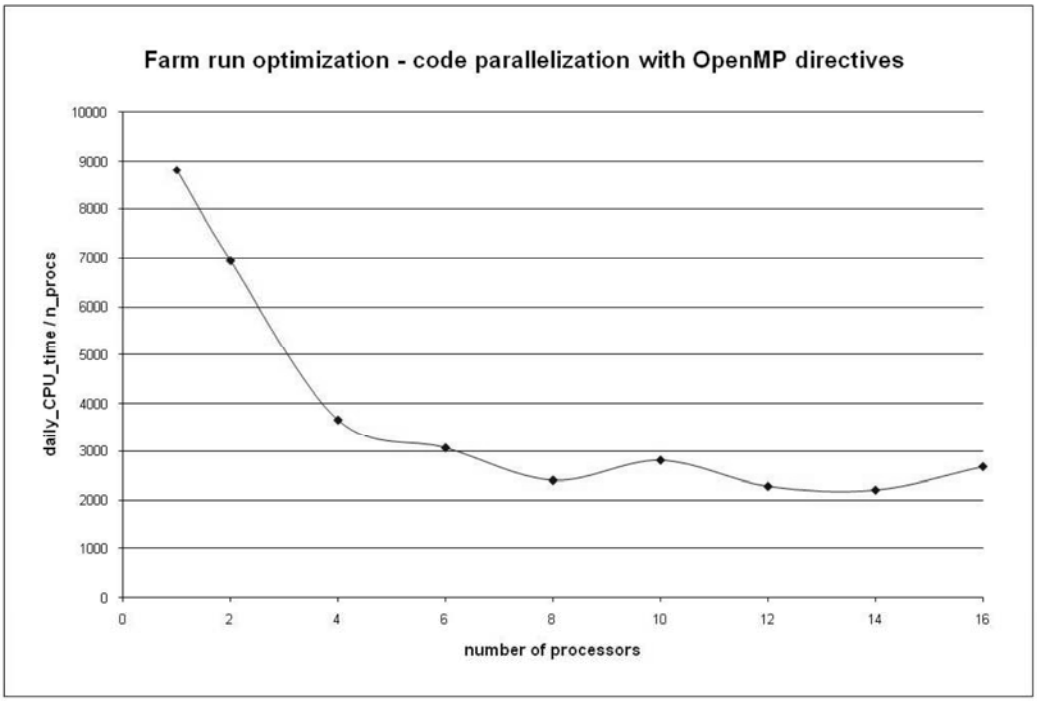


Figure 5: *Maximum execution time (1day CPU time per processor) as function of the processors number.*

We performed the output validations by using the Italian stations of the AirBase 2005 dataset (v.4) [14]. We calculated most of the statistical model performance indices, based on standard air quality model evaluation methodologies [15] and EU regulation. For example, ozone index for forests protection AOT40F (annual sum of the excess of hourly concentrations over the cut-off of 40 ppb during light hours, from April to September for forest trees) is shown in *Figure 6*.

The AOT40Fs calculated from observed AirBase data are overlapped on the simulated AOT40F field; different kind of observation stations are represented with different symbols indicating the standard classification. *Figure 6* shows an overall acceptable performance of the model, that is pretty good for industrial and traffic measure points.

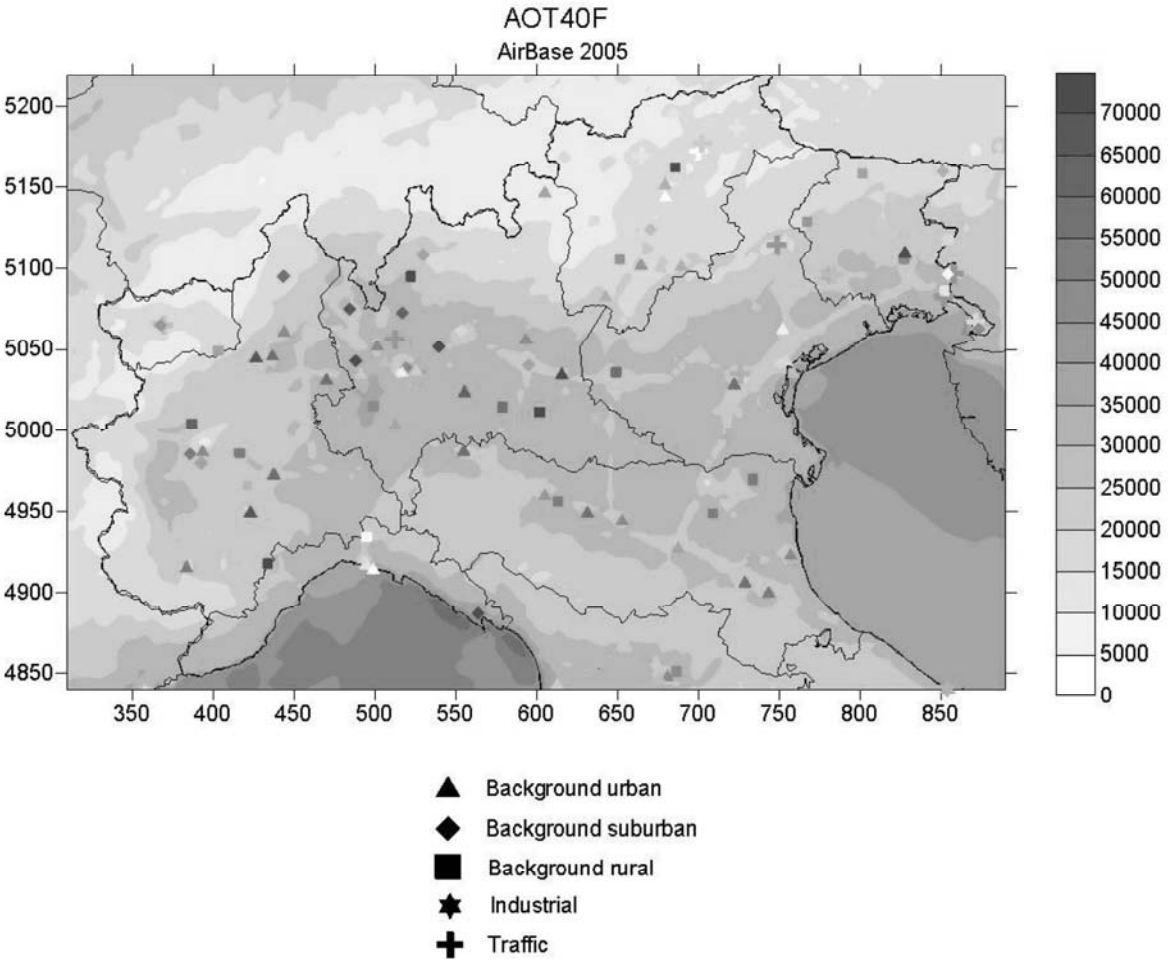


Figure 6 : *Map of zone index AOT40F [µg m^{-3} h] and measured data. Observation points are classified according to station types. White symbols mean lacking of observed data.*

Some of the results of the simulations were presented during this year at International Conferences: in February 2010 at the COST 728 final workshop [16]; in June 2010 at the 13th International Conference on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes [17,18] and in September 2010 at the International Aerosol Conference of Helsinki [19].

REFERENCES

[1] ZANINI G., PIGNATELLI T., MONFORTI F., VIALETTO G., VITALI L., BRUSASCA G., CALORI G., FINARDI S., RADICE P., SILIBELLO C. *The MINNI Project: An Integrated Assessment Modeling System For Policy Making*. In Zerger, A. and Argent, R.M. (eds) MODSIM 2005 International Congress on Modelling and Simulation. Modelling and Simulation Society of Australia and New Zealand, Melbourne, Australia, 12-15 December 2005, pp. 2005-2011. ISBN: 0-9758400-2-9.

[2] COTTON W.R., PIELKE R. A., WALKO R. L. , LISTON G. E., TREMBACK C. J., JIANG H., MC ANELLY R. L., HARRINGTON J. Y., NICHOLLS M. E., CARRIO G. G. AND MCFADDEN, J. P. *RAMS 2001: Current status and future directions*. Meteorology and Atmospheric Physics, 82, (2003) 5-29.

[3] ARIANET, EMMA (EMGR/make) – User's guide – Version 3.5, 2005, Arianet R2005.08

[4] SILIBELLO C., CALORI G., BRUSASCA G., GIUDICI A., ANGELINO E., FOSSATI G., PERONI E., BUGANZA E. *Modelling of PM10 concentrations over Milano urban area using two aerosol modules*. Environmental Modelling & Software, 23 (2008), 333-343.

[5] VITALI L., FINARDI S., PACE G., PIERSANTI A., ZANINI G. *Validation of simulated atmospheric fields for air quality purposes in Italy*. Poster presentation. Proceedings of the 13th International Conference on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes. Paris, June 2010.

[6] ISPRA – Istituto Superiore per la Protezione e la Ricerca Ambientale - www.sinanet.apat.it

[7] GUENTHER A., HEWITT C.N., ERICKSON D. ET AL. *A global model of natural volatile organic compound emissions*. Journal of Geophysical Research, 100 (2005), 8873-8892.

[8] BERICO M., MALAGUTI A., LORENZELLI R., SERRA F., TONELLI C. *Nuova metodologia per la stima del fattore emissivo delle polveri da lavorazioni in agricoltura*. ENEA Report RT/2010/37/ENEA.

[9] CARMICHAEL G.R., UNO I., PHADNIS M.J., ZHANG Y., SUNWOO Y. *Tropospheric ozone production and transport in the springtime in east Asia*. Journal of Geophysical Research, 103 (1998), 10649–10671.

[10] CARTER W.P.L. *A detailed mechanism for the gas-phase atmospheric reactions of organic compounds*. Atmospheric Environment, 24 A, 3 (1990), 481-518.

[11] BINKOWSKI F.S., ROSELLE S.J. *Models-3 community multiscale air quality (CMAQ) model aerosol component 1. Model description*. Journal of Geophysical Research, 108, D6 (2003), 4183.

[12] http://www.emep.int/

[13] International Institute for Applied Systems Analysis, http://www.iiasa.ac.at/

[14] AirBase - the European Air quality database, http://air-climate.eionet.europa.eu/databases/airbase/query_retrieval.html

[15] CHANG J. C., HANNA S.R. *Air quality model performance evaluation*. Meteorol Atmos Phys 87 (2004), 167–196.

[16] FINARDI S., BRIGANTI G., BRUSASCA G., CALORI G., CAPPELLETTI A., D'ALLURA A., MIRCEA M., PACE G., PEDERZOLI A., RADICE P., SILIBELLO C., VITALI L., ZANINI G. *Air quality assessment over northern Italy for the reference year 2005*. Mesoscale modelling for air pollution applications: achievements and challenges - COST 728 final workshop, Geneva, 25-26 February 2010.

[17] MIRCEA M., CAPPELLETTI A., BRIGANTI G., PEDERZOLI A., VITALI L., PACE G., MARRI P., SILIBELLO C., FINARDI S., CALORI G., ZANINI G. *Impact of horizontal grid resolution on air quality modeling: a case study*. Poster presentation. Proceedings of the 13th International Conference on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes. Paris, June 2010.

[18] BRIGANTI G., CAPPELLETTI A., MIRCEA M., PEDERZOLI A., VITALI L., PACE G., MARRI P., SILIBELLO C., FINARDI S., CALORI G., ZANINI G. *Impact of horizontal grid resolution on air quality modeling: a case study*. Poster presentation. Proceedings of the 13th International Conference on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes. Paris, June 2010.

[19] MIRCEA M., CAPPELLETTI A., BRIGANTI G., PEDERZOLI A., VITALI L., PACE G., MARRI P., SILIBELLO C., FINARDI S., CALORI G., ZANINI G. *Impact of horizontal grid resolution and emission inventory on fine and coarse particulate matter: a case study over Italy.* International Aerosol Conference, Helsinki, 29 August - 3 September 2010.

SIMULATION OF CHEMICAL REACTIONS CATALYZED BY TRANSITION METALS COMPLEXES AND LARGE–SCALE PARALLEL SIMULATIONS OF BIOMEMBRANES BY HYBRID PARTICLE-FIELD MOLECULAR DYNAMICS TECHNIQUE

Luigi Cavallo

Chemistry Department, University of Salerno.

Activities made by the MoLNaC (Molecular Lab for Nanomaterials and Catalysis)

Modern theoretical methodologies, aided by the advent of high speed computing and also by impressive efforts in massive parallel processing, are enabling to calculate directly the mechanical properties of novel materials, to follow the dynamics of chemical reactions, to sample the conformational landscape of biomolecules, and more. This makes computer science somehow eclectic, *i.e.* it draws upon multiple theories, instead to be hold just to one paradigm or a set of assumptions and allows a net gain of complementary insights into a subject.
In this perspective activity at the MoLNaC (Molecular Lab for Nanomaterials and Catalysis at the University of Salerno, http://www.molnac.unisa.it) spans from the simulation of chemical reactions catalyzed by transition-metal complexes to large-scale parallel simulations of biomembranes by Hybrid Particle-Field Molecular Dynamics Technique.

1) Static DFT study of the mechanism of homopolymerization of polar vinyl monomers and of their copolymerization with non polar olefins.

Within a collaboration with Prof. Chen (Colorado State University) on the acrylate coordination-addition polymerization, DFT calculations have been performed to clarify the mechanism of methylmetacrylate homopolymerization into highly syndiotactic polymers catalyzed by group 4 metallocenes.[1] Experimental results obtained by Chen and co-workers show that the cationic *ansa*-metallocene ester-enolate catalyst derived from the activation of the precatalyst $[Ph_2C(Cp)(2,7\text{-}t\text{Bu}_2\text{-Flu})]Zr[OC(O\text{i}Pr)dCMe_2]_2$ with $[Ph_3C][B(C_6F_5)_4]$, stood out as the best catalyst in *all* aspects of the MMA polymerization at room temperature, including the highest activity, efficiency, syndiotacticity, and control. Mechanistic results are consistent with a catalyst-site-controlled, monometallic coordination-addition mechanism, involving fast intramolecular addition within the catalyst-monomer complex leading to the resting eight-membered ester enolate chelate, followed by the rate-limiting ring-opening of the chelate to regenerate the active species. Several unique features of this polymerization system, as the constant syndiotacticity of poly methyl methacrylate produced over a wide polymerization temperature range and the insensitivity of its high activity, degree of control, and stereoselectivity to solvent polarity and structure of weakly coordinating anions; and deviation from a pure site-control mechanism at high [MMA]/[catalyst] ratios, were rationalized by DFT calculations. The computational results on real size models, which have been possible due to the CRESCO computational power, provided theoretical support for the proposed monomer-assisted catalyst-site epimerization, after an enantiofacial mistake, to a thermodynamically more stable resting state. This accounts for the observed higher than expected [*mr*] contents based on a pure site-controlled mechanism. Moreover, DFT calculations rationalized why the $Ph_2C<$ bridged catalyst exhibits higher stereoselectivity than other catalysts with the $Me_2C <$ or $Me_2Si <$ bridge: the bridge rigidity pushes the η^3-bound Flu ligand closer to the growing chain and the monomer,

thereby increasing ΔE_{stereo} between the competing transition states for the addition of a monomer molecule to the opposite (correct and wrong) enantiofaces of the enolate growing chain. The relative polymerization activity of this catalyst series is shown to correlate with the relative energetics of the back-biting of the penultimate unit and ion-pair formation.
DFT calculations were performed also to study the mechanism of copolymerization of hydrocarbon monomers with polar monomers (acrylates and vinyl acetate), catalyzed in this case by neutral Pd(II) phosphinosulfonate complexes. This work arises from a collaboration with the experimental group of Prof. Mecking (University of Konstanz).[2] Due to its relevance, and with aim to shed light on the underlying reaction steps, organometallic polymerization intermediates from multiple consecutive acrylate insertions were isolated and their reactivity studied by combining experimental and theoretical results. DFT calculations rationalized the homopolymer and copolymer low molecular weight based on the high stability of the six membered resting state formed by chelating coordination of the penultimate inserted acrylate unit. The penultimate acrylate unit inserted in the polymer chain is κ-*O* coordinate to the metal center via the carbonyl group. Further chain growth requires opening of this highly stable chelate by the monomer (ethylene or acrylate): this is the key limitation in acrylates homopolymerization and their copolymerization with hydrocarbon monomers.

2) Dynamic DFT study of the mechanism of olefins metathesis promoted by Ru-catalysts.

Within an FP7 project (Eumet project, grant agreement n° CP-FP 211468-2), we performed advanced static and dynamics DFT simulations of a deactivation reaction promoted by CO coordination to the Ru center of a N-heterocyclic carbene based catalyst. Static DFT calculations indicate that CO binding to the Ru center promotes a cascade of reactions with very low energy barriers that lead to the final crystallographically characterized product, in which the original methylidene group has attacked the proximal aromatic ring of the N-heterocyclic carbene ligand leading to a cycloheptatriene through a Buchner ring expansion, see *Figure 1*.

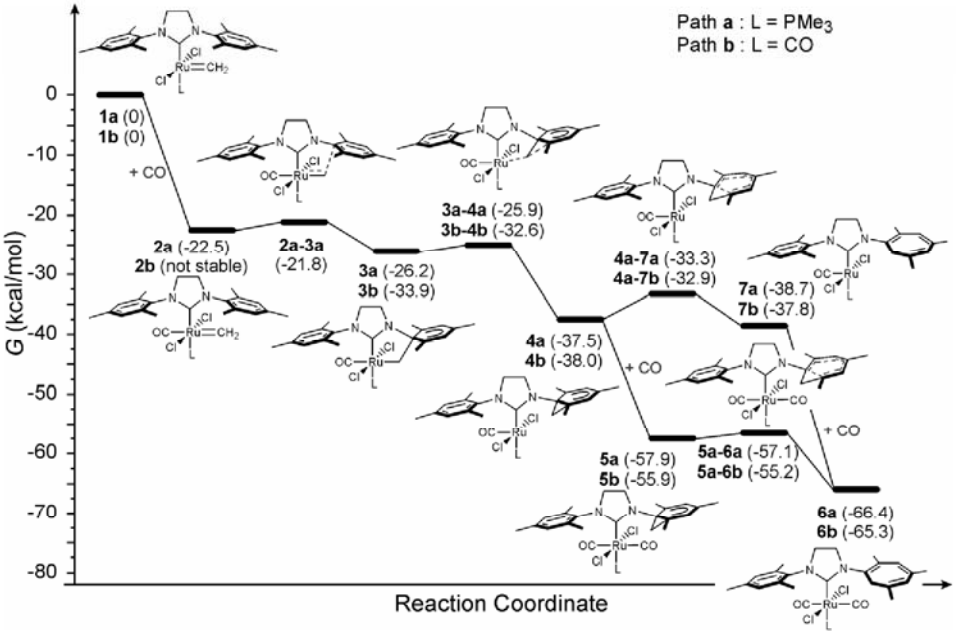


Figure 1: *Energy diagram of the complete* ***1*** *to* ***6*** *deactivation pathway. In parenthesis is the energy of the various species, in kcal/mol, relative to the (pre)catalyst* ***1a.***

Analysis of the relevant molecular orbitals, supported by *ab initio* molecular dynamics simulations, illuminate the key role of the π-acid CO coordinated trans to the Ru-methylidene bond to promote this reactivity. Based on this result, we investigated to which extent a large set of π-acid groups could promote this deactivating reaction. Results clearly indicate that almost any s-acid group that can approach the Ru center in the sterically crowded position trans to the Ru-methylidene bond can promote this deactivation route.[3]
Still with a DFT approach we performed a detailed static and dynamics characterization of eleven N-heterocyclic carbene (NHC) ligands in Ru-complexes of general formula $(NHC)Cl_2Ru{=}CH_2$. Analysis of the dynamic trajectories indicates that the nature of the N-substituent can result in extremely different flexibility of the Ru-complexes. In almost all the cases the N-substituent trans to the Ru-ylidene bond is severely folded so that it protects the vacant coordination position at the Ru-center. Limited flexibility is instead associated to the N-substituent on the side of the Ru-ylidene bond. NHCs with a single *ortho* substituent, either a simple Me or a bulkier *i*-Pr group, have a preferential folding that bends the unsubstituted side of the ring towards the halide-Ru-halide plane. Analysis of the dynamics trajectories in terms of buried volume indicates that the real bulkiness of these systems can be somewhat modulated and this flexibility is a key feature that allows NHCs to modulate their encumbrance around the metal in order to make room for bulky substrates. Analysis of the buried volume in terms of steric maps illuminated that NHCs with mesityl or 2,6-diisopropylphenyl N-substituents shape quite different reactive pockets. Rather flat with a constant pressure on the halide-Ru-halide plane the former, vault shaped with higher pressure on the sides the latter. Regarding the NHCs with an *ortho* tolyl or *i*-Pr group on the N-substituent, the steric maps quantify the higher impact of the unsubstituted side of the ligand in the first coordination sphere of the metal, and evidence the overall C_S and C_2-symmetric reactive pockets of the corresponding complexes. We believe that the detailed characterization of the differently shaped reactive pockets is a further conceptual tool that can be used to rationalize the experimentally different performances of catalysts bearing these ligands or to devise new applications.[4]

3) Large-scale parallel simulations of biomembranes by Hybrid Particle-Field Molecular Dynamics Technique

Recently, in collaboration with Prof. T. Kawakatsu (Tohoku University) we developed a hybrid simulation technique where self-consistent field theory (SCF) and molecular dynamics (MD) simulation are combined (MD-SCF).[5] The idea is to obtain a strategy, as far will be possible, having the main advantages and avoiding the main disadvantages of both SCF and atomistic approaches. MD simulations are widely applied in the both fields of synthetic materials and biomolecules. Due to the possibility of the straightforward chemically consistent models such as full atomistic or specific coarse-grained (CG) models these methods are very useful to characterize structural and dynamical properties of polymeric materials and biomolecules. The main disadvantage is the limitation to small length and time scales. Also in the best case of specific coarse-grain models the actual computer power limits the length scale to few tens of nanometers and few thousands of nanoseconds. In contrast, SCF approaches assure accessibility to definitely larger length and time scales but having models with very low chemical specificity.
The computational efficiency of this method is due to the fact that the calculation of the intermolecular forces (the most demanding part of a simulation) that in classical MD is calculated by double loop over particle pairs is completely skipped.

These interactions are calculated at mean field level considering the interactions, at mean-field level, between single particles in a density field. Parallel MD-SCF program using atom decomposition method with the aid of Message Passing Interface (MPI) library has been developed. This program will make full use of the advantage of High Performance Computing (HPC) and greatly contribute to the investigation of synthetic materials and biomolecules with larger length scales and longer time scales.

One interesting application of these methods is the understanding of interactions of biocompatible block polymers with biological interfaces. This research has important technological applications in industry and in medicine. In particular, triblocked polymer based on polyethylenoxide (PEO) and polypropilenoxyde (PPO), are broadly used for biotechnological and biomedical applications. Despite these applications, the mechanism interactions of polymers with the biological system are not yet completely understood. The main factor influencing the efficiency of the proposed scheme is the low amount of communication of the computational problem due to the particular formulation having single molecules in a density field. In *Figure 2* is depicted a large system that can be already handled in the current implementation.

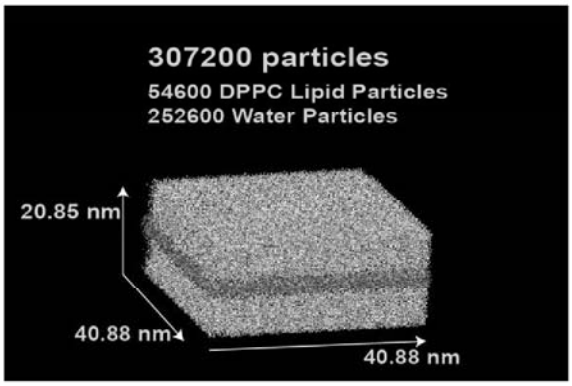


Figure 2: *The system modelling a biomembrane made of Dipalmitoylphosphatidylcholine (DPPC) lipids and water.*

This system size would allow a preliminary study of the interaction of a polymeric micelle with a biomembrane. More in general, system size of million of particles are considered as the largest systems that can be simulated. [5]

REFERENCES

[1] ZHANG Y., NING Y., CAPORASO L., CAVALLO L., CHEN E. Y.-X. "*Catalyst-Site-Controlled Coordination Polymerization of Polar Vinyl Monomers to Highly Syndiotactic Polymers*" Journal of the American Chemical Society **132** (2010) pp 2695-2709.

[2] GUIRONNET D., CAPORASO L., NEUWALD B., GÖTTKER-SCHNETMANN I., CAVALLO L., MECKING, S. "*Mechanistic Insights on Acrylate Insertion Polymerization*" Journal of the American Chemical Society **132** (2010) pp 4418-4420.

[3] POATER A., RAGONE F., CORREA A., CAVALLO L. "*Exploring the Reactivity of Ru-Based Metathesis Catalysts with a π-Acid Ligand Trans to the Ru−Ylidene Bond*" Journal of the American Chemical Society **131** (2009) pp 9000–9006.

[4] RAGONE F., POATER A., CAVALLO L. "*Flexibility of N-Heterocyclic Carbene Ligands in Ruthenium Complexes Relevant to Olefin Metathesis and Their Impact in the First Coordination Sphere of the Metal*" Journal of the American Chemical Society **132** (2010) pp 4249–4258.

[5] MILANO G., KAWAKATSU T. "*Hybrid particle-field molecular dynamics simulations for dense polymer systems*" The Journal of Chemical Physics **130** (2009) pp 214106-214114.

ADVANCES IN LES OF THE HYSHOT II SCRAMJET COMBUSTOR

D. Cecere, F.R. Picchia, N.M. Arcidiacono, E. Giacomazzi, F. Donato

ENEA, Casaccia Research Centre, Rome, Italy

A Ingenito, C. Bruno

University of Rome "La Sapienza", Department of Mechanics and Aeronautics, Rome, Italy.

With the overall goal to clarify the physics of compressible (supersonic) combustion, a 3D LES of the HyShot supersonic combustor has been performed and is reported in this paper. HyShot is an (originally) Australian program to assess feasibility of supersonic combustion by means of a ballistic test flight. The HyShot combustion chamber is shaped as a box 75x9.8 mm in cross section and 300 mm long. Hydrogen is injected at 90 degrees with respect to the supersonic airstream 40 mm downstream from the combustor inlet by means of four 2 mm diameter choked orifices. Air enters the channel at a Mach number that, in the actual test, depended on the flight trajectory; in this simulation, the trajectory point is that at height = 28 km, where the Mach number was 2.79, P=82.11 kPa and T = 1229 K. A structured grid of about $52x10^6$ nodes discretizes the actual combustor shape, where hydrogen-air combustion is treated by means of a detailed chemical kinetics model including 9 species and 37 reactions. The present LES also reproduce the large-scale dynamics of the flow reported in the experiments, but the richness of data provided by LES allows a much deeper exploration of the flow physics. Numerical results indicate that hydrogen penetrates in the air stream generating 3D bow shock structures upstream of the injection orifices as seen in experiments. In these regions recirculation zones upstream and downstream of the fuel injection orifices are observed as expected; the OH predicted by LES indicates that a flame starts already in the upstream recirculation zone. Interactions among the essentially 1D airstream entering the combustor, the heat released and the 3D jets produce large vorticity rates and therefore enhance and accelerate turbulent mixing. Combustion is predicted very fast and efficient: only 0.5% of hydrogen is found unburned at the combustor exit.

Introduction

Thanks to advances in accuracy of numerical algorithms and computational power, the application of Large Eddy Simulation (LES) to compressible turbulent combustion is the focus of significant research due to his superior capabilities than Reynolds-averaged Navier-Stokes simulations in providing more realistic turbulent flowfield such as mixing mechanisms due to unsteady eddy structures, heat transfer and combustion. In these supersonic combustors in fact, the enhancement of supersonic turbulent mixing of jet fuel and crossflow air is a critical point for combustion, due to the limited residence time of flow in the combustion chamber.

Lack of experimental data due to difficulties in setting up test facilities (Bardina et al, 1997) and in measuring thermo-physics flow proprieties has made understanding the physics of compressible flows a very difficult topic: the choice of geometry and diagnostics is sometimes a compromise and may be critical. Much of the physics of supersonic combustion is still unclear; more recently, novel experimental data shown by Nishioka [[i], [ii], [iii]] and past data by Swithenbank et al [[iv], [v], [vi]] indicate mixing may be even faster than commonly assumed or predicted, and the flame short, compact and with low ignition delay. These findings were also known from US work

done in the '60s for aeronautical applications and [mostly] still classified. In fact, data on mixing in supersonic flows has always been considered scarce. This is due to the fact that supersonic flows are fully 3D and mixing is mainly due to streamwise vortex structures generated by fuel jet impingement. Mixing between two *2-D* parallel supersonic jets can be drastically slow with respect to the same subsonic jets [vii]: in fact, in *2-D* experiments eddies may be only *spanwise*, streamwise vorticity is suppressed, and so is the powerful mixing enhancement due these 'helicoidal' structures.

In this context, LES simulations can be of support, but until recently, numerical simulations were still far from realistically predicting the (few) experimental results. This is due to the fact that for these high flow Reynolds numbers, a fairly dense grid is required to properly predict flame anchoring and vortex structures, implying large computational resources. At the same time, it is important for LES simulations to adopt high order schemes with low-diffusion numerical schemes to accurately reproduce complex shock interactions, contact surfaces without dumping artificially resolvable scales of turbulence present in high Reynolds number flows.

In the present study, LES simulations of Hyshot combustion chamber are performed by means of a in-house code (S-HeaRT, Supersonic Heat Release and Turbulence). A hybrid method has been implemented to properly capture shocks while at the same time solving the transport equations away from discontinuities via a low dissipation, central scheme with fourth order accuracy.

HyShot Geometry

The SCRJ combustor was not designed to produce thrust, only to show that the supersonic hydrogen-air flame could ignite and be anchored. In the present work, the flight conditions at 28 km and M=7.6 of HyShot I are simulated. The HyShot scramjet geometry consists of a rectangular air intake 305 mm long x 100 mm wide, a combustor 300 mm long x 75 mm wide, and a thrust plate 200 mm long x 75 wide. The intake is a 17° inclined wedge compressing the incoming hypersonic flow (see *Figure 1*).

The combustor is 300 mm long and has a constant rectangular area, 9.8 mm high and 75 mm wide, hosting 16 pressure transducers starting from 90 mm downstream the combustor leading edge. Distance between pressure transducers is 13 mm. The thrust plate is at 12° with respect to the airstream path and is equipped with 11 pressure transducers 11 mm downstream from the combustor exit. Distance between pressure transducers is also 13 mm. The fuel is injected in cross flow with respect to the incoming air by means of four 2 mm diameter injectors located 40 mm downstream of the combustor inner surface leading edge. The SCRJ inlet conditions assumed in these simulations are those from telemetry or estimated at altitude 28 km and 0° angle of attack; the equivalence ratio was 0.426. The combustor inlet pressure and temperature are respectively 82.11 kPa and 1229 K. Air enters the combustor at Ma=2.79. Fuel is injected transversally at Ma=1, T=250 K and P=303.340 kPa. *Figure 1* shows (in red) the portion of the combustor simulated in the present work.

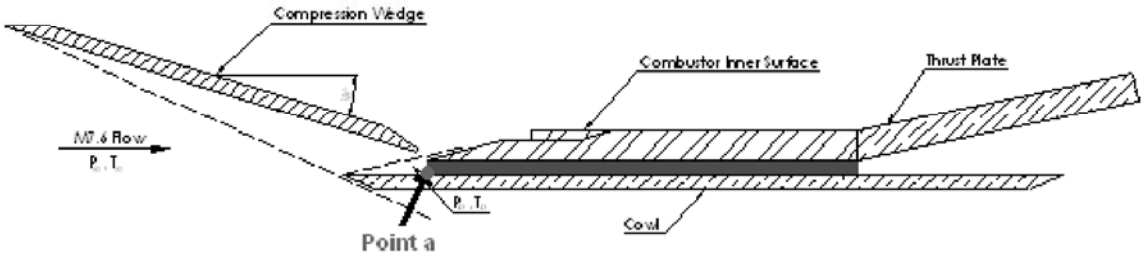


***Figure 1**: HyShot Geometry.*

The HyShot inlet conditions are reported in *Tables 1-2* below [viii]. In order to simulate properly the flow velocity at the combustor entrance, and to avoid simulating a shock wave departing from the bottom of the combustor (Point a) a velocity profile has been imposed at the combustor inlet. In fact, the boundary layer is already well developed and must be accounted for. This velocity profile has been obtained from previous simulations by Jeung et al. [ix], including also the intake flowfield.

	Combustor inlet
Pressure [Pa]	82110
Mach No.	2.79
Density kg/m3	0.2358
Temperature [K]	1229
Sound speed	682.9 m/s
Flow velocity	1905.291 m/s

***Table 1**: Combustor inlet flow conditions*

	Combustor inlet
Pressure [Pa]	307340
Mach No.	1
Density kg/m3	0.3020
Temperature [K]	250
Sound speed	1204.4 m/s
Flow velocity	1204.4 m/s
Φ	0.426

***Table 2**: Fuel injector flow conditions*

The SCRJ combustor geometry of Figure 1 has been mapped by a 3-D cartesian nonuniform grid. The three dimensional computational domain, defined by a multi-block Cartesian orthogonal structured grid, is partitioned into a number of sub-domains using a genetic type algorithm. The combustion chamber grid includes 880x130x450 = 51.480M grid points.

Numerical Model

3D LES of the HyShot combustor have been performed with the S-HeaRT code developed in ENEA.

A finite difference method on a Cartesian not uniform grids in a co-located cell-centered variable arrangement, together with an explicit, fully compressible solver is used to solve the reactive, Navier-Stokes (N.-S.) equations. In writing the numerical scheme, the focus was on the numerical approximation of the derivatives in the advection terms at the resolved scales. The interface fluxes are evaluated by a hybrid method capable of capturing shocks and at the same time resolving the turbulent structures away from discontinuities with minimal dissipation. This goal was achieved by adopting a 4^{th}-order central scheme. In the shock capturing scheme the reconstruction of the Riemann problem at the cell interface is performed by means of a WENO 53 scheme. To obtain the interface fluxes from the reconstructed states the approximate hybrid HLLC/HLLE Riemann solver is implemented. Time-integration is by means of the fully explicit third order accurate TVD Runge-Kutta scheme of Shu and Osher. The N.-S. equations are thus fully coupled. At the inlet, all quantities are prescribed, except density in the subsonic regions.

Partially non-reflecting boundary conditions have been implemented (following the NSCBC technique) to reduce numerical reflection of acoustic waves back into the computational domain. Because of the short testing time, all walls are assumed adiabatic. The perfect gas law is assumed. The fitness function implemented in the genetic algorithm aims to balance the computational cost among the processors, and to minimize the amount of data transferred across processors boundaries. A Domain Decomposition approach and MPI (Message Passing Interface) are used for the parallel implementation of the S-HeaRT code that can be run on different platforms (e.g., clusters of single or multiple-core machines, SMP machines and so on). For numerical stability the time step was about 10^{-9} s.

Turbulence SGS and Chemistry Model

The unsteady simulations performed are based on the "eddy viscosity" Fractal Model, FM subgrid model. This model turns itself off in laminar regions of the flow. The detailed kinetic mechanism of Warnatz, involving 8 species and 37 chemical reactions, is adopted in the present simulation as it was already implemented in the HeaRT code. A future simulation will evaluate differences with the scheme by Jachimowski. Molecular processes not taken into account at the resolved scale are Dufour and Soret effects, cross-diffusion, pressure gradient diffusion, and diffusion due to body forces. Preferential diffusion is included, and the species diffusive fluxes are modeled following Hirschfelder and Curtiss. All molecular properties for individual chemical species, except their binary mass diffusivities, are calculated a priori by the software library EGlib provided by Prof. Ern. In particular, kinetic theory is used for dynamic viscosity and thermal conductivity. The values calculated are then stored in look-up tables from 200 to 5000K every 100K. Values for intermediate temperatures are calculated at run-time by linear interpolation. The mixture average properties are estimated at run-time. In particular, the simulations used in this work implement Wilke's formula with Bird's correction for viscosity, and Mathur's expression for thermal conductivity. The effective diffusion coefficients, D_i, of species i inside the mixture are estimated by assuming individual Schmidt numbers, Sc_i, calculated as the median of Sc_i vs T distributions for non-premixed flames.

Results

3D LES were run for 2 ms starting from the initial condition. The snapshots shown in what follows were taken at 1.4 ms and are discussed below. Statistics will be presented in a future paper.

Figure 2 shows the density field at X= 0.0093 m and X= 0.0375 m. The bottom slice shows the H2 distribution at the upper wall. This figure shows the bow shock wave located at 10 mm from injection and a train of shock waves reflecting from the upper wall and impinging on the H2 jet core. Due to the bow shock, there is a recirculation zone upstream the fuel injection where hydrogen diffuses.

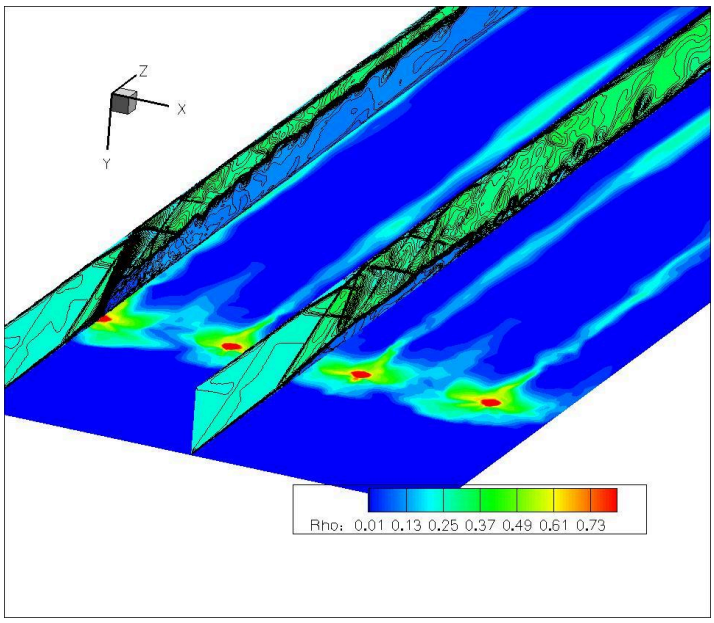


***Figure 2**: Density at X= 0.0093 m and X= 0.0375 m.*

By inspecting the flowfield, hydrogen is seen to expand rapidly, partially 'blocking' the supersonic air crossflow, and forcing development of three-dimensional shock structures (see Fig. 3a). In fact, the interaction between fuel jet and supersonic crossflow produces coherent structures such as horseshoe vortices, jet-shear layer vortices, contra-rotating vortex pairs and wake vortices. These structures are very important because of their role in enhancing fuel-air mixing. In *Figure 3a* color mapping shows H_2O already present upstream the fuel injection. The 3D features of vortex structures in its side view are visualized by the instantaneous isosurfaces of the streamwise in *Figure 3b*. These isosurfaces are colored in red and green, and represents clockwise and counter-clockwise rotating vortices with axis in the streamwise direction.

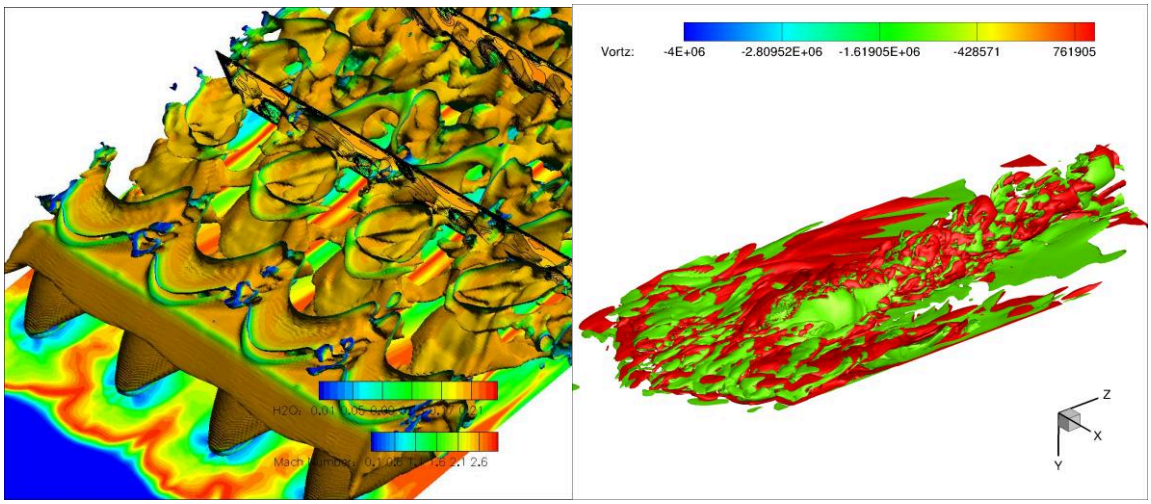


***Figure 3 a)**: Isocontour of pressure at 181 KPa, slice of H_2O; b) Isosurfaces of the z component of filtered vorticity.*

Due to these structures, hydrogen is convected outwards, toward the two eddies in-between the two H_2 streams. *Figure 4a* shows the two-dimensional pressure contours from the side view near the H_2 inlet region. It's clearly evident, the barrel shock, the Mach disk and the bow shock, that due to the high pressure region in front of the jet hole bends forward in the near wall region.

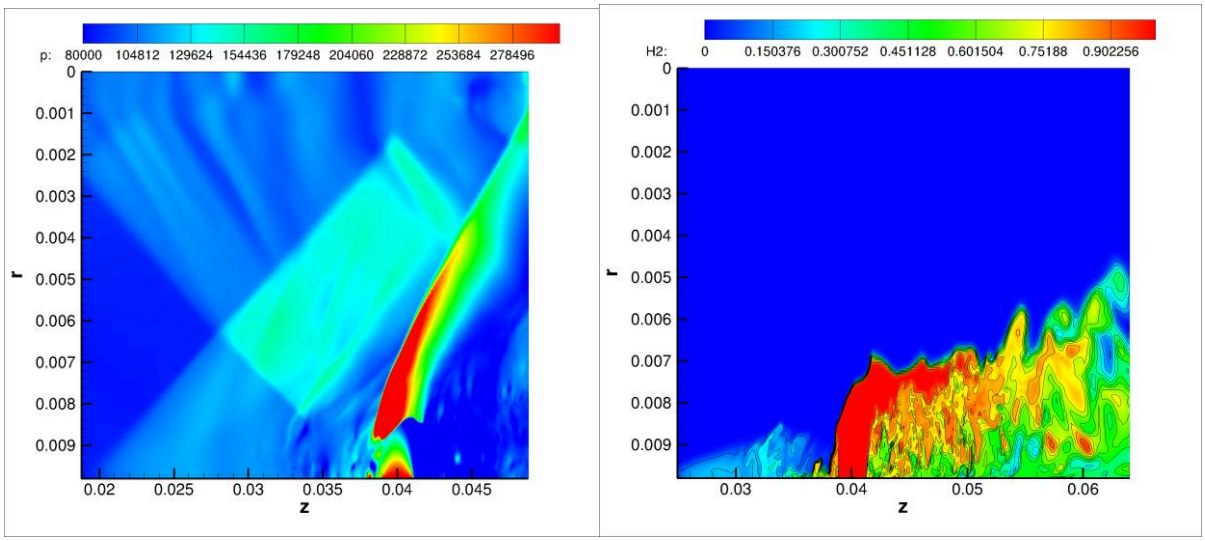


***Figure 4 a)**: Contour of pressure; b) Contour of H_2 mass fraction.*

Figure 4b shows the mass fraction contour of the H_2 that penetrates in the flow-stream and propagate upstream along the boundary layer.

Preliminary Conclusions

In this work, a 3D LES of the HyShot scramjet combustor has been done by means of a hybrid numerical scheme, a highly refined grid and a detailed kinetic scheme. Results point out combustion is very efficient: the flame anchors upstream the flow injection, within the recirculation zone between the bow shock and the fuel injection: the bow shock is located about 10 mm from the H_2 orifices. Mixing is very predicted to be very fast, in particular contra-rotating vortices within the H2 core flow improve the turbulent diffusion of H_2 while eddies between the [initially] separate fuel streams are responsible for the fast air/ H_2 mixing. The flame structure is shown to be driven by large scale eddies, suggesting a possible "flame in eddies" structure. Favored also by the high initial temperature, kinetics is very fast. Combustion efficiency calculated by the unburned H_2 mass fraction (only 0.5% at the combustor exit) is ~ 99.5%. Future work will include testing other kinetic schemes.

REFERENCES

[i] Nishioka, M., Sakaue, S., Komada, K., Sakoshi, H., Furukawa, I., "Mixing transition in supersonic streamwise vortices", *Fluid Mechanics And Its Applications*, Vol. 79, 2006, pp. 249-258.

[ii] Nishioka, M., Hiejima, T., Sunami, T., and Sakaue, S., "Streamwise Vortices as a Powerful Means for Supersonic Mixing Enhancement", *Proc. Int. Symp. On Dynamics and Statics of Coherent Structures in Turbulence: Roles of Elementary Vortices*, ed. by S. Kida, pp. 217–228, 2002.

[iii] Sunami, T., Wendt, M., and Nishioka, M., "Supersonic mixing and combustion control using streamwise vorticity", *AIAA paper 98-3271*, 1998.

[iv] Swithenbank, J., Eames, I.W., Chin, S.B., Ewan, B.C.R., Yang, Z., Cao, J., and Zhao, X., "Turbulent Mixing in Supersonic Combustion", High Speed Flight Propulsion Systems, ed. by S.N.B. Murthy and E.T. Curran, *Progress in Astronautics and Aeronautics*, Vol. 137, 1991, pp. 341-382.

[v] Swithenbank, J., Chigier, N. A., Vortex Mixing for Supersonic combustion, *21st (International) Symposium on Combustion*, The Combustion Institute, Pittsbutgh, PA., 1969. pp. 1154–62.

[vi] Swithenbank, J., "Flame Stabilization in High Velocity Flow", *Combustion Technology*, ed. by H. Palmer and J. M. Beer, Academic Press, N Y, 1974, pp. 91-127.

[vii] Papamoschou, D., and Roshko, A., "The Turbulent Compressible Shear Layer: An Experimental Study", *Journal of Fluid Mechanics*, Vol. 197, 1988, pp. 453-477;

[viii] M. Frost, A. Paull and H. Alesi, *Report on the Hyshot Scramjet Experiments in The T4 Shock Tunnel*, Australia, 2001.

[ix] S.H. Won, I. S. Jeung and J. Y. Choi, "Turbulent Combustion Characteristics in HyShot Model Combustor with Transverse Fuel Injection", AIAA 2007-5427, *43rd AIAA/ASME/SAE/ASEE Joint Propulsion Conference*, 8 - 11 July 2007, Cincinnati, OH

METHODOLOGICAL APPROACH TO STUDY ENERGETIC AND STRUCTURAL PROPERTIES ON NANOSTRUCTURED CADMIUM SULFIDE BY USING AB INITIO MOLECULAR DYNAMICS SIMULATION

E.Burresi[1], M.Celino[2]

[1]ENEA, C.R. Faenza, via Ravegnana 186, Ravenna, Italy
[2]ENEA, C.R. Casaccia, via Anguillarese 301, Rome, Italy

Introduction

Nanostructures semiconductors II-VI like cadmium sulfide quantum dots (CdS QDs) have attracted much attention for their novel optical and electronic properties employed in different technological fields. Due to particular band gap, CdS present high luminescent characteristics and is being used in photovoltaic field, emitting and optoelectronic devices. In addition, novel and interesting properties have been obtained for CdS nanostructure due to the quantum confinement[1] that appears when the particles size is on the order of the excitonic Bohr radius [2],[3]. In this way the band gap is tunable by size, obtaining modulated luminescent properties from synthesis process [4],[5],[6].
The stability on the CdS phase, and in general the stability of phase between bulk and nanostructure material, is still now discussed and not entirely understood [7],[8],[9],[10]. In fact,in the specific case, two main crystallographic forms for CdS bulk are wurtzite and zincblende and wurztite is the most stable at ordinary pressures and temperatures. As a general result for nanoparticles, the crystallographic structure is size dependent and a more probable stable wurtzite phase is found for a length greater than 6 nm. However it is not yet completely understood the exactly crystallographic form for small CdS nanoparticles and several data and results are not in agreement between themselves.
It will illustrate a calculation procedure to describe the geometrical and structural properties of CdS-QD by means of ab initio molecular dynamics simulation.

Computational details

The results presented in this paper were carried out by using Car-Parrinello Molecular Dynamics (CPMD) (Car and Parrinello 1985) [11] in density functional theory framework. We used a generalized gradient approximation BLYP for the exchange-correlation functional already implemented in CPMD package. In addition, we applied a norm-conserving pseudopotentials approach in order to describe the interaction between valence electrons and core ions: in the specific case, a Martins-Trouiller type of pseudopotentials in a Kleinman-Bilander scheme for S and Cd were used. Moreover for Cd atom, nonlinear core correction was applied whereby the exchange-correlation functional is calculated by adding a model core density to the valence electron. All structures, atoms, dimers and cluster were studied like isolated molecules. The temperature range of annealing was set from 100 K to 400 K. Furthermore, for 10K, a geometrical optimization was supplementary done. CPMD molecular dynamics was performed in increments of 50 K; a Nosé-Hoover chain thermostat was used to control the temperature.

Results

Preparatory calculations on atoms and dimers of cadmium and sulfur then on CdS dimers had initially requested to set box size and cutoff energy. In this context we also tested box and cutoff values calculating binding energy and bond length for cadmium, sulfur and cadmium sulfide dimer, checking the theoretical results with experimental data as reported in *table 1*, obtaining a satisfactory accord. In order to carry out calculations on large size systems as CdS cluster and to reduce computational costs, the cutoff energy was set at 30 a.u.

Data	CPMD(blyp ct=30)	Experimental
R (Cd-Cd) Å	3.79	4.07-4.33*
R (S-S) Å	1.93	1.89
R (Cd-S) Å	2.47	2.52**
BE (Cd2) Kcal/mol	0.2	2.7
BE (S2) kcal/mol	159.8	101.9-100.7††
BE (CdS) kcal/mol	44.2	91 / ≤47†

***Table 1**: Optimize bond length and binding energy (BE) for Cd_2,S_2 and CdS dimer. The experimental BE data are from Handbook of Chemistry and Physics 53rd Edition, page F-183.*

On the initial structure (*Figure 1*) we performed a CPMD molecular dynamics calculation at 10 K to relax the geometry. Subsequently, the annealing of the cluster in the range 100-400K was done.

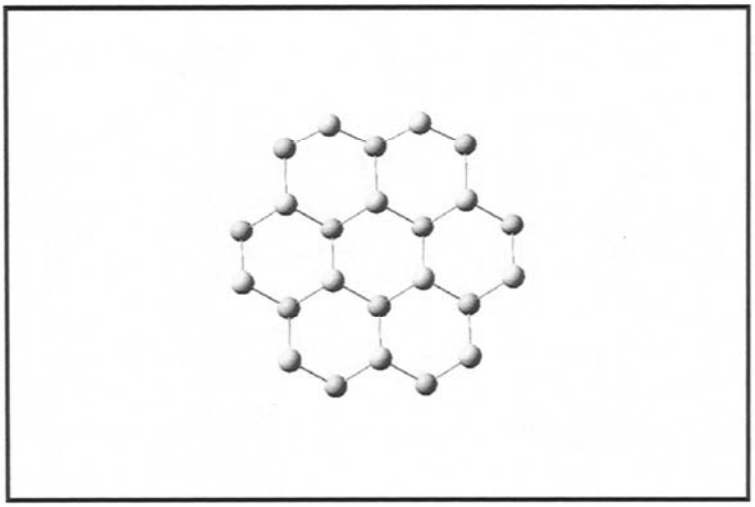


***Figure 1**: Top views of original CdS wurtzite cluster.*

In *Figure 2* is reported the energy formation per cadmium sulfide unit as a function of temperature. In this context, it has been showed a general linear trend for low temperatures range, from 100 K to 250 K where the behavior is slightly rising. At 300 K the figure shows a slight modification of this behavior, probably due to the growth of small atomic surface modification (*Figure 3* where views of CdS structures at 300 K are reported).
This tendency could suggest for next temperatures novel conformation of the CdS cluster that could be more stables than previous ones. Moreover, Wen [12] reported a graphic showing that at 350 K the most stable phase seems to be the rocksalt for nanoparticles with size less than 2 nm. In our case at 350 K is evident a strong decrease of the formation energy.

This change suggests that from 300 K started a modification of the wurtzite phase towards a more stable form.

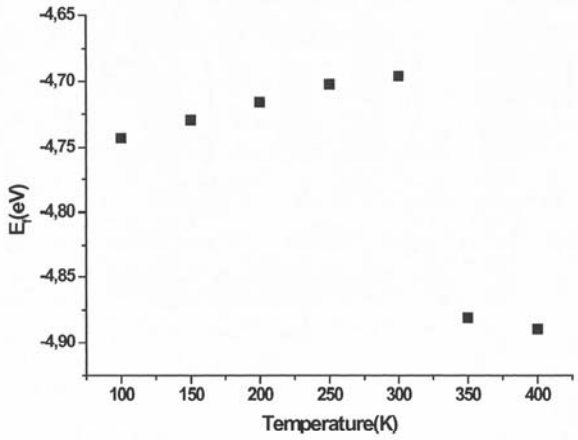


Figure 2: *Energy of formation as a function of temperature for CdS cluster.*

Probably a modified phase with larger amorphous contribution could appear between 300 K and 350 K, stabilizing at 400 K. This transformation could start from surface where the effects are stronger than inside the cluster: the effect is probably stronger for not perfectly spherical nanoclusters or in the absence of passivators.

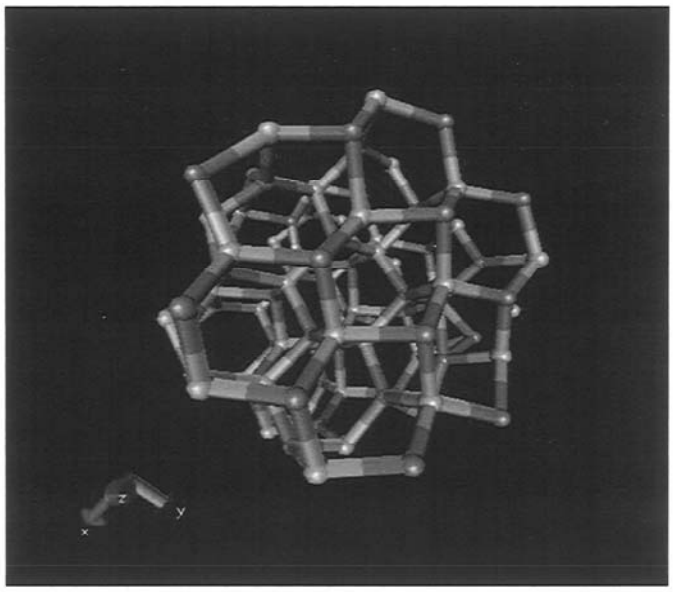


Figure 3: *CdS structure at 300 K.*

Finally in order to study in detail the structural behavior of the CdS nanostructures is notable to build a radial distribution function (RDF) of the atoms around a specific ones. In particular, in our case, the results from the RDF for initial CdS sample (with high structural order) and for RDF of the each structures at different temperature could be interpreted in terms of increase of the amorphized fraction.

REFERENCES

[1] ALIVISATOS A.P., "*Prospective on the Physical Chemistry of Semiconductor Nanocrystals*" The Journal of Physical Chemistry. **100** (1996) pp. 13226-13239.

[2] BRUS L.E. *"Electron-Electron and electron-hole interactions in small semiconductor crystallites: The size dependence of the lowest excited electronic state"* Journal of Chemical Physics **80** (1984) pp. 4403-4409.

[3] GAPONENKO S.V. "*Optical Properties of Semiconductor Nanocrystals*" Cambridge University Press, (1998) Cambridge UK

[4] ANTOLINI F., PENTIMALLI M., DI LUCCIO T., TERZI R., SCHIOPPA M., RE M., MIRENGHI L., TAPFER L . *"Structural characterization of CdS nanoparticles grown in polystyrene matrix by thermolytic synthesis"*. Materials Letters **59** (2005) pp. 3181-3187.

[5] HERRON N., WANG Y., AND ECKERT H. *"Synthesis and Characterization of Surface-Capped, Size-Quantized CdS"* Journal of the American Chemical Society **112** (1990) pp. 1322-1326.

[6] MURRAY C.B., NORRIS D.J., AND BAWENDI M.G. "*Synthesis and Characterization of Nearly Monodisperse CdE(E= S, Se, Te)"* Journal of the American Chemical Society **115** (1993) pp. 8706-8715.

[7] JUNKERMEIER C.E. AND LEWIS J.P. "*Amorphous nature of small CdS nanoparticles: Molecular dynamics simulations*" Physical Review B **79** (2009) DOI: 10.1103/PhysRevB.79.125323.

[8] KOTKATA M.F., MASOUD A.E., MOHAMED M.B., MAHMOUD E.A. "*Synthesis and structural characterization of CdS nanoparticles"* Physica E **41** (2009) pp. 1457-1465.

[9] MURAKOSHI K., HOSOKAWA H. AND YANAGIDA S. "*Controlling Microscopic Surface Structure, Crystalline Size and Crystallinity of CdS and Zns Nanocrystallites*" Japanese Journal of Applied Physics **38** (1999) pp. 522-527.

[10] RICCOLLEAU C., AUDINET L., GANDAIS M.,AND GACOIN T. "*Structural transformations in II-VI semiconductor nanocrystals"* European Physical Journal D **9** (1999) pp. 565-570.

[11] CAR R. AND PARRINELLO M. "*Unified Approach for Molecular Dynamics and Density-Functional Theory*" Physical Review Letters **55** (1985) pp.2471-2474.

[12] WEN B. AND MELNIK R.V.N. "*First principles molecular dynamics study of CdS nanostructure temperature-dependent phase stability*" Applied Physics Letters **92** (2008) pp. 261911 DOI: 10.1063/1.2952835.

DEVELOPMENT OF A MODEL FOR THE NUMERICAL SIMULATION OF MULTIPHASE REACTING FLOWS BY LARGE EDDY SIMULATION TECHNIQUE

Filippo Donato

ENEA-UTTEI/COMSO
C.R. Casaccia - Roma

Much attention has been paid to multiphase flows modeling in the last decades. Two concurrent reasons can justify such an interest in the research community. The first motivation is that multiphase flows can be found in a variety of industrial processes and components (e.g. fluidized beds, gas turbine burners, etc.) as well as in many everyday life devices (e.g. computer printers); the second reason can be addressed to the increasing computational capabilities, which is making Computational Fluid Dynamics (CFD) to be a more and more useful tool in the design and optimization process of such devices and components. In the CFD framework, Large Eddy Simulation (LES) approach is gaining in importance as a tool for simulating turbulent combustion processes. Such a technique is nowadays a standard as far as single phase phenomena are considered, while much work is being done to improve its performance in multiphase flow applications.

The present work is intended to be a contribution to the development of reliable models for the simulation of multiphase dispersed reacting flows within LES framework. The aim of this research project is to provide an improvement to the capability of existing particle transport models in predicting the dispersed phase evolution under dilute condition and for inertial particles. The main applications towards which this work is oriented are coal powder burners and spray combustors. The coupling of the LES accuracy in predicting gas phase turbulent combustion and an improved model for particle dispersion in the carrier gas could help in the design process of such components.

Within the present activity an Eulerian-Eulerian (EE) model for the description of a two-phase dispersed flow of inertial particles has been implemented in the HeaRT code. The HeaRT code is ENEA's CFD code for LES simulations of reacting flows. In the EE description of the heterogeneous gas-particles system, the dispersed phase is treated as a second fluid mixed to the continuous phase and conservation equations are solved for quantities assumed as representative of the all phase. The choice to adopt such a description instead of the more commonly used Eulerian-Lagrangian (EL) one, where single particles (or particle parcels) are tracked within the computational domain is due to the higher efficiency potential of the EE approach when massive parallelization is taken into account [1]. Details on the implementation, the numerical approach and possible improvement of the model can be found in [2].

The multiphase version of the HeaRT code has been validated on the Sommerfeld and Qiu [3] test case, as far as the transport model is considered. In the Sommerfeld and Qiu experiment a particle laden flow is obtained by injecting small (≃080 μm) glass particles in a confined swirled flow. The injection system is composed by a cylindrical duct and by an annular duct coaxial to the first one. Air and particles flow through the cylindrical duct to the test chamber

while from the annular duct a swirled air flow is introduced in the test section. The swirl number is 0.47. At the end of the test chamber an expansion chamber is present.

A sketch of the computational domain including a detail of the inlet zone is shown in *Figure 1*. Measures of both air and particle velocities are taken on 8 different radial planes, from 3 mm from the injection plane down to 315 mm from it. Data for both phases mean and RMS velocities and their radial distributions are given. Different classes of particle size are considered.

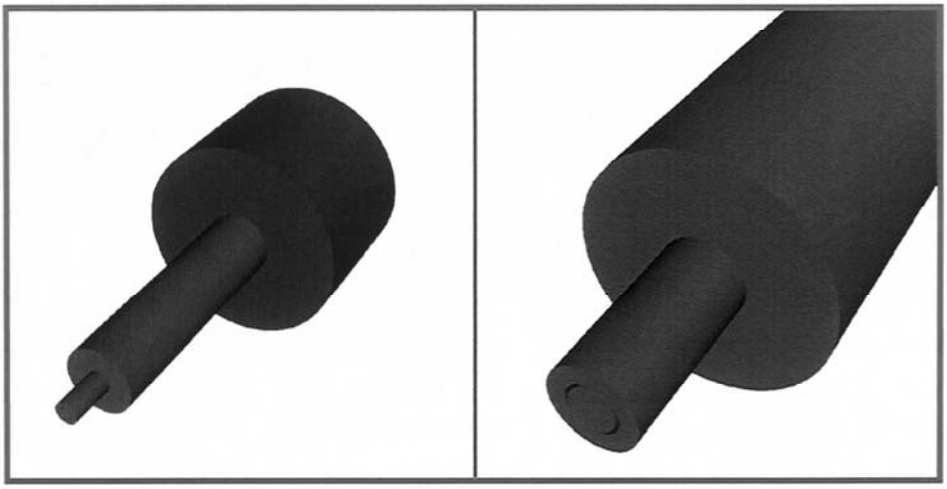


Figure 1: *Computational domain (left) and a detail of the inlet (right).*

There are different reasons that suggested the choice of this experiment for the validation of the models and numerics developed in this work:

a) there is a large variety of numerical publications based on it;

b) the extensive database of measures suits well to the validation objectives;

c) similar flow configurations can be found in pulverized coal burners and the particle size is within the range typically used for these applications, though the mean diameter is smaller (45 μm) here compared to the one (100 ÷ 200μm) typically adopted for power plant burners.

The test case conditions, that refer to standard atmosphere, are summarized in *Table1*. The test chamber is 0.703 m long and has a diameter D5=0.194 m, while the inlet duct length is 0.127 m. The expansion chamber is 0.464 m long and has a diameter D6 = 0.630 m.

The computational grid is structured and cylindrical. It is composed by 4 blocks: two for the inlet ducts, one for the test chamber and one for the expansion chamber. All these blocks present 64 azimuthal planes. The grid is thus made of 2775040 computational cells, whose smallest size is 0.0005 m.

Air flow	
Mass flow rate of the primary jet M f 1 (g/s)	9.9
Mass flow rate of the secondary jet M f 2 (g/s)	38.3
Inlet Reynolds number (with D3 = 64 mm)	52400
Swirl number	0.47
Particle phase	
Particle mass flow rate M p (g/s)	0.34
Particle loading in the primary jet	0.03
Particle properties	
Particle mean diameter (μm)	45
Particle material density (kg/m3)	2500

Table 1: *test conditions*

In order to give the best idea of the three-dimensional structure of the flow field, the module of the gas phase momentum over an axial and a radial plane normal to the axis is presented in *Figure 2*. It can be seen how the gas phase motion presents a wide central recirculation zone and the main stream is concentrated around it, close to the outer wall.

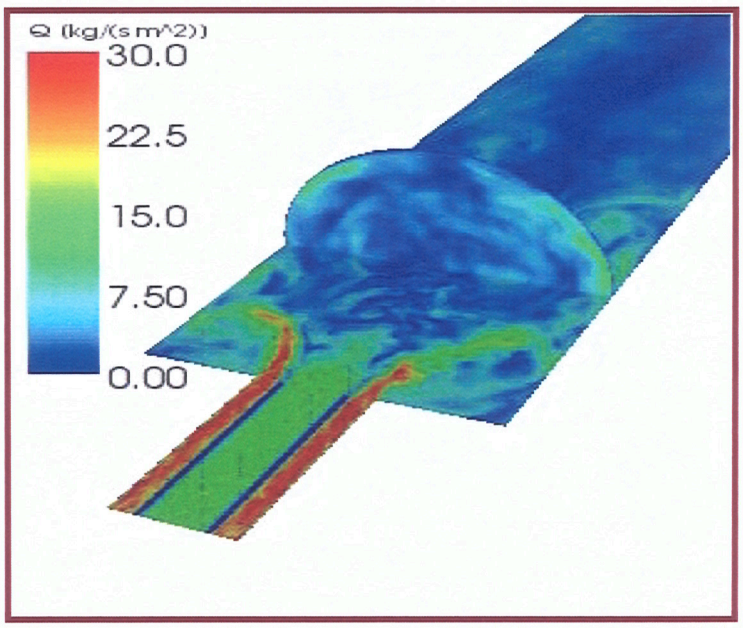


Figure 2: *Instantaneous field of the momentum module ρu. z normal plane at z = 150 mm.*

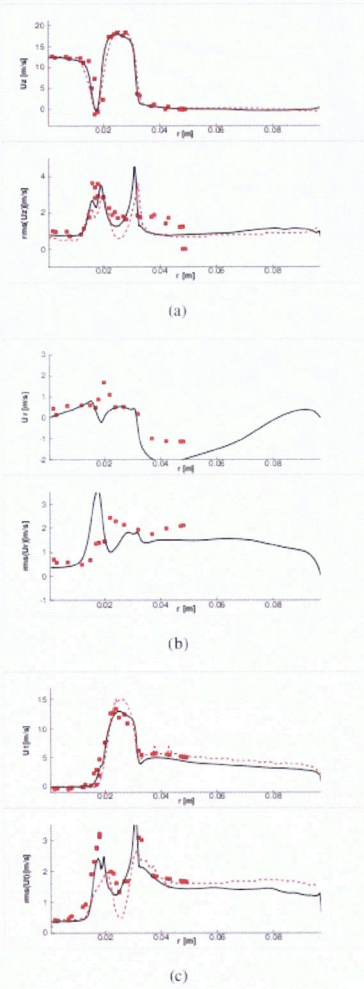
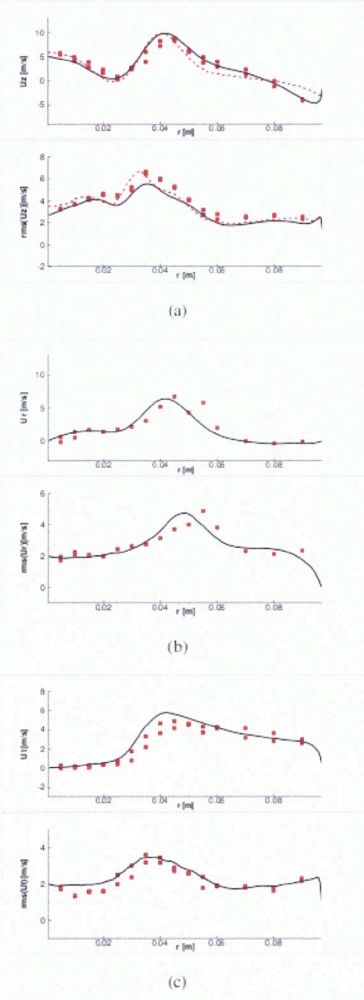


Figure 3: *Mean (upper part) and RMS (lower part) axial (a), radial (b) and tangential (c) velocities for z = 3 mm. Symbols are taken from experiments. Lines are computed.*



Figue 4: *Mean (upper part) and RMS (lower part) axial (a), radial (b) and tangential (c) velocities for z = 52 mm. Symbols are taken from experiments. Lines are computed.*

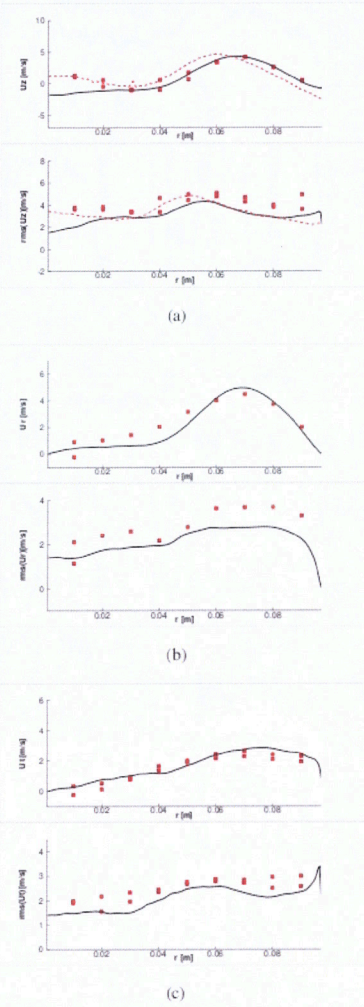


Figure 5: *Mean (upper part) and RMS (lower part) axial (a), radial (b) and tangential (c) velocities for z = 85 mm. Symbols are taken from experiments. Lines are computed.*

In *Figure 3* the experimental data and the computed solution for the gas velocity are presented for the plane at z = 3 mm, where z is the axial coordinate with its origin located at the test chamber inlet section. Solid lines refer to the present work while dashed lines, which can be assumed as a reference state of the art for the EE models, are taken from [4] and have been obtained by means of the CERFACS AVBP code with similar model and grid resolution. In *Figures 3 (a)* and *(c)* it can be recognized how both mean values (upper parts) and fluctuating velocities (lower parts) for the axial and tangential components of the gas velocity are well reproduced. The radial positions of local maximum are perfectly caught, the magnitudes are in very good agreement with the experiments. The same cannot be said for the radial component. Both the shape and the level of the RMS are not well predicted. This can be due to the poor

resolution adopted in front of the bluff body separating the inner and outer duct. In fact, since z = 3 mm, only 6 grid points are present before the plane where the measures are taken and it will not be possible to reconstruct smaller structures in that zone. In *Figure 4* the computed solution is compared to the experimental data for z = 52 mm. In *Figure 4(a)* it can be seen how both the mean and RMS axial velocities are very well reproduced by the HeaRT code. For both radial and tangential components (*Figure 4(b-c)*) the RMS velocities are very well predicted at all the radial positions with a slight overprediction for the tengential component. At z = 85 mm (*Figure 5*) the mean radial component is well predicted close to the center of the domain and in the outer part. The axial component presents an underprediction in the central part.

As to the dispersed phase, in *Figure 6* a 3D view of the particle dispersion process due to the interaction with the gas phase is presented. An isocontour of the particle volumetric fraction, colored by the dispersed phase local velocity can be seen. The condensed phase is dispersed in a large variety of structures and the effect of the swirl on particle distribution is clearly visible.

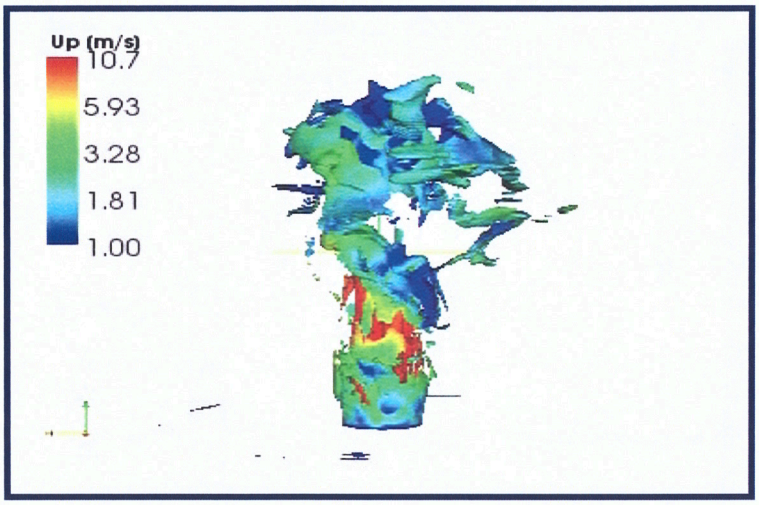
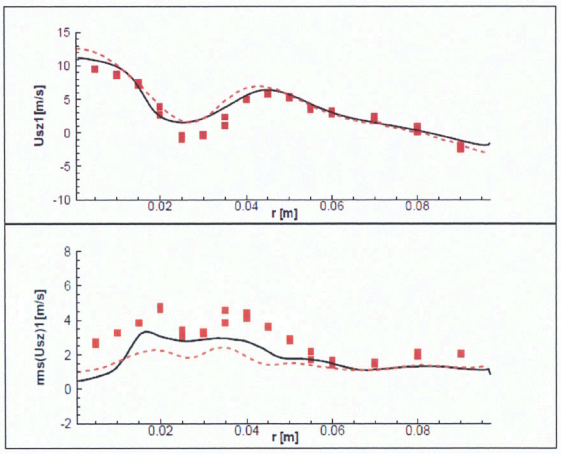


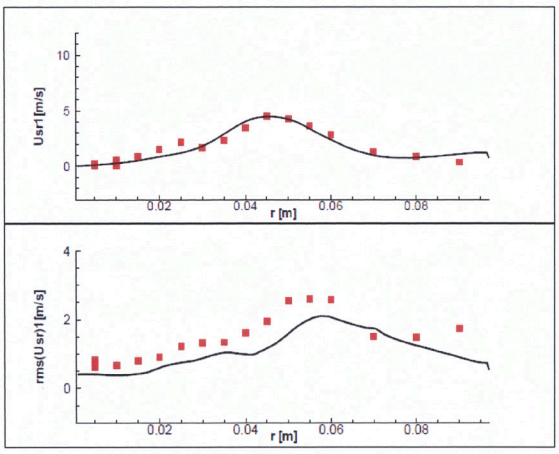
Figure 6: *Instantaneous isosurfaces of the dispersed phase volume fraction αp = 2.e-5.*

In *Figure 7* and *Figure 8* a comparison between predicted and experimental velocities for the dispersed phase is shown for z = 52 mm and z = 85 mm respectively. The mean velocities are well reproduced with the exception of the axial component for z = 85 mm (*Figure 7(a)*). The RMS levels are in an overall good agreement, though some underestimation is present at some planes.

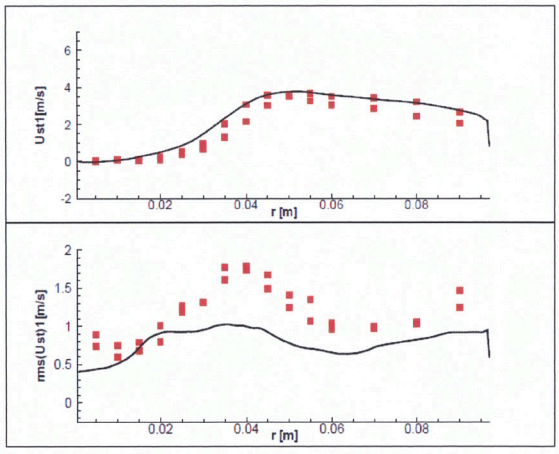
The validation of the implementation of the two-phase EE transport model in the HeaRT code can be considered satisfactory.



(a)

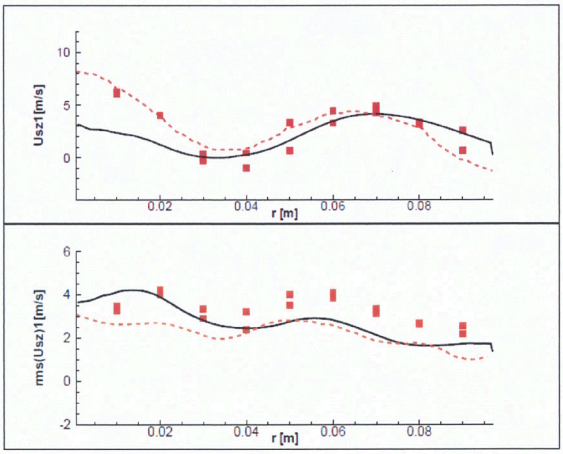


(b)

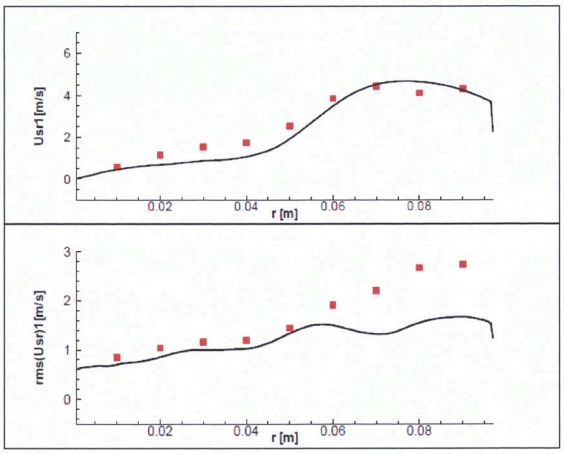


(c)

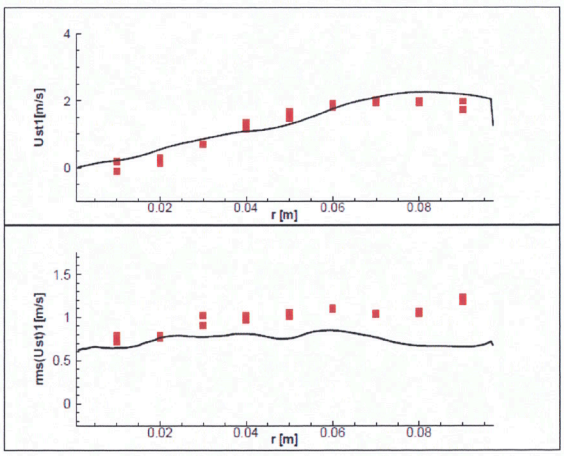
Figure 7: *Mean (upper part) and RMS (lower part) axial (a), radial (b) and tangential (c) velocities for z = 52 mm. Symbols are taken from experiments. Lines are computed*



(a)



(b)



(c)

Figure 8: *Mean (upper part) and RMS (lower part) axial (a), radial (b) and tangential (c) velocities for z = 85 mm. Symbols are taken from experiments. Lines are computed*

REFERENCES

[1] M. García E. Riber O. Simonin and T. Poinsot. Comparison between euler/euler and euler/lagrange LES approaches for confined bluff-body gas-solid flow. International Conference on Multiphase Flow, 2007.

[2] F. Donato. Numerical modeling of multiphase reacting flows. PhD thesis, "Sapienza" University of Rome, 2010.

[3] M. Sommerfeld and H.H. Qiu. Characterization of particle-laden, confined swirling flows by phase-doppler anemometry and numerical calculation. Int. J. Multiphase Flow, 19(6):1093–1127, 1993.

[4] M. Boileau, S. Pascaud, E. Riber, B. Cuenot, L.Y.M. Gicquel, and T. Poinsot. Investigation of two-fluid methods for Large Eddy Simulation of spray combustion in gas turbines. Technical report, 2007.

OPEN FOAM COMPUTATIONAL PERFORMANCES FOR LARGE PARALLEL DISTRIBUTED SIMULATIONS

Agostino Funel, Fiorenzo Ambrosino[1]

[1]ENEA UTICT-HPC C. R. Portici

Astract

We present some results concerning the computational performances of OpenFOAM [1], in terms of scalability and efficiency, on a High Performance Computing (HPC) system. This task has been accomplished by studying large, memory intensive simulations of external aerodynamics to evaluate the drag coefficient over the surface of a commercialized car. The flow field calculation was performed by using the simpleFoam solver with standard K-ε turbulence model. The geometry was decomposed for running in parallel. Simulations were carried out up to 2000 cores on CRESCO HPC cluster [2], a facility of ENEA [3], the Italian National Agency for New Technologies, Energy and Sustainable Economic Development. The hardware architecture consists of 256 blades IBM HS21 each supporting dual Xeon Quad-Core Clovertown E5345 processors (2.33 GHz, 8 MB L2) for total of 2048 cores. Two network configurations have been tested: InfiniBand (4XDDR) [4] and Giga Ethernet. In order to reduce the influence of I/O operations data were collected on the IBM General Parallel File System (GPFS) [5], which provides concurrent high-speed file access via InfiniBand. We also present results for the Intel Turbo Boost Nehalem E5530 processor (2.40 GHz, 8 MB L3) [6].

Itrodution

As the importance of High Performance Computing in CFD is increasing, industries are more and more interested in its applications. However, the cost of license of commercial CFD codes is proportional to the number of cores used and thus running large simulations in parallel on multi-core systems may be economically prohibitive unless an open source software like OpenFOAM is used. We have studied OpenFOAM computational performances in a case of industrial interest. We study the air flow over a car. Such a kind of simulations are used in automotive industry to model the shape of cars. Often one has to try many configurations before finding the best aerodynamics. However with very large cases the optimization process would require very long delays not compatible with industrial deadline commissions. It is essential then to sensibly reduce the computational time. OpenFOAM can accomplish this task by running in parallel on large HPC systems. The geometry of the problem is shown in *Figure 1*. The air enters perpendicularly to the inlet at a velocity **U**=39.0m/s. The operating pressure is P=1atm. The computational domain consists of about 30 millions tetrahedral cells.

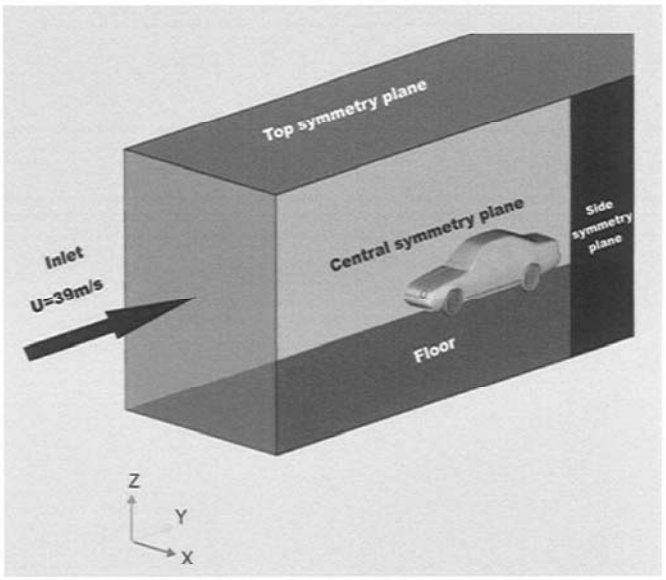


Figure 1: *The geometry.*

Methods

Hardware architecture - CRESCO HPC system is based on the leading multi-core x86_64 technology. The system consists of two main sections: (1) for **high memory requests** and moderate parallel scalability; (2) for limited memory and **high scalability** cases; and an experimental section (3) for special architectures (Cell, FPGA, CUDA). Simulations have been run on CRESCO section (2) which is composed by 256 blades IBM HS21 each supporting dual Intel Xeon Quad-Core Clovertown E5345 processors (2.33 GHz, 8 MB L2), 16 GB RAM for total of 2048 cores. An IBM General Parallel File System (GPFS) is shared via the InfiniBand (IB) 4XDDR network interconnect. The IB network is based on a CISCO SFS 7024 (288 ports), a CISCO SFS 7012 (144 ports) and 5 CISCO SFS 7000 (120 ports), as shown in *Figure 2*. In CRESCO section (2) have been added 28 blades each supporting dual Intel Xeon Quad-Core Turbo Boost Nehalem E5530 processors (2.40 GHz, 8 MB L3), 16 GB RAM for total of 224cores.

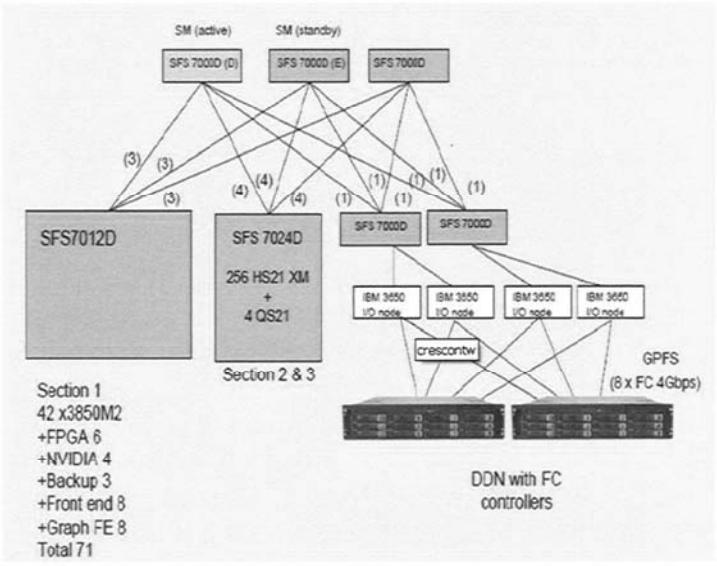


Figure 2: *Architecture of the InfiniBand network including the IBM/DDN 9550 storage system. The 4 I/O Nodes, directly Fibre Channel attached to the storage are the GPFS NSD servers.*

Software

CRESCO OS is Linux CentOS version 5.3. We have used OpenFOAM version 1.4.1 compiled from source with the GNU gcc-4.2.1 compiler. Compilation flags: C/C++ "-m64 –fPIC"; optimization: "-O3 –fno-gcse". Communication library: OpenMPI version 1.2.8 compiled with the Intel icc compiler version 11.0. Parallel decomposition method: metis.

Results

Results in terms of scalability and efficiency are shown in *Table 1*. OpenFOAM scales very well up to 2000 cores with an average efficiency of 87%. In *Figure 3* are shown results for Clovertown architecture.

cores	1	32	64	128	256	520	1024	1600	2000
speed-up	1	25.5	53.4	106.2	217.2	482.2	913.3	1491.4	1867.5
efficiency	1	0.80	0.83	0.83	0.85	0.93	0.89	0.93	0.93
time	62.7 days	2days 11h	1 day 4h	14.2 hours	7 hours	3h 7min	1h 38min	1 hour	48 min

Table 1: *In the table above are reported speed-up and efficiency values for simulations done with different numbers of cores. The "time" value is the total time for simulations to get convergence. Numerical schemes: upwind for laplacian term; Gauss linear for other terms.*

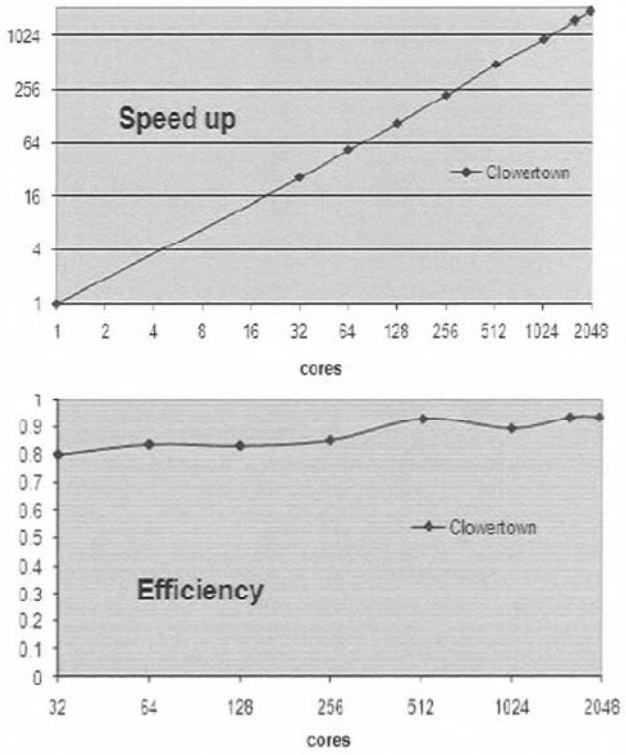


Figure 3: *OpenFOAM speed-up and efficiency up to 2000 cores with Dual Quad-Core Clovertown E5345 architecture. Data have been collected on GPFS via InfiniBand 4XDDR.*

Figure 4 shows a comparison of Clovertown to Nehalem processors. Measurements indicate that Nehalem is about two times faster than Clovertown up to 200 cores. *Figure 5* shows the final flow streamlines. We also have tested the Giga Ethernet network. In this case the total available bandwidth is reduced (1Gb/s compared to 20 Gb/s of IB). With 64 cores the obtained speed-up is 32.51, sensibly lower than 53.4, the value obtained with IB.

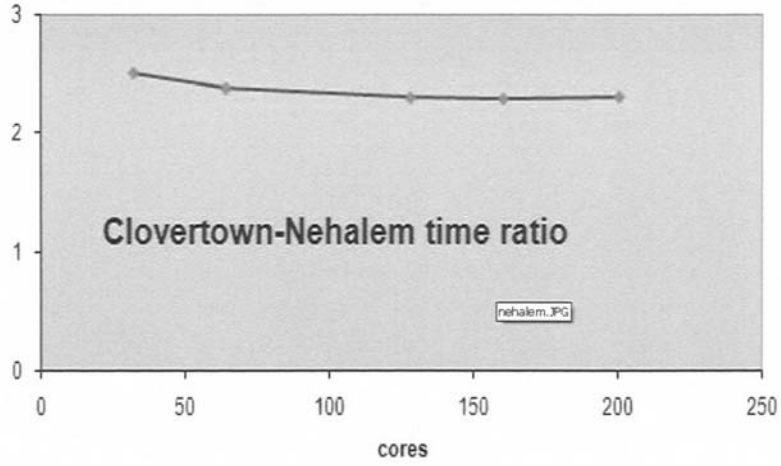


Figure 4: *OpenFOAM Clovertown-Nehalem comparison up to 200 cores. Nehalem processor is about two times faster than Clovertown. Data have been collected on GPFS via InfiniBand 4XDDR.*

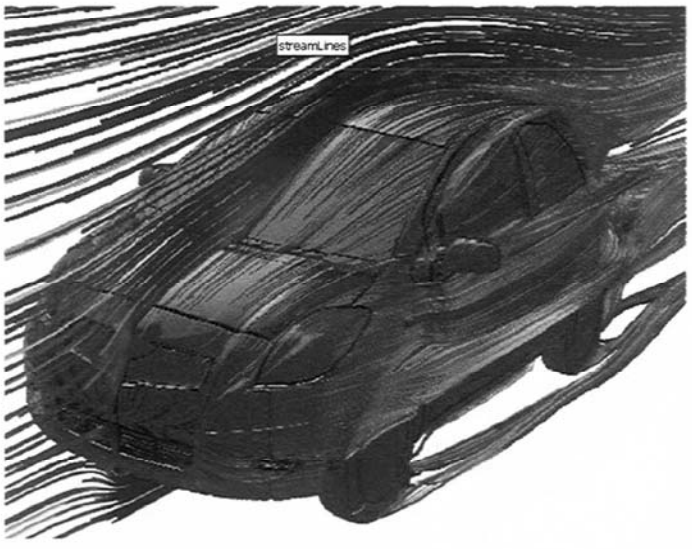


Figure 5: *Stream lines.*

REFERENCES

[1] http://www.openfoam.com/

[2] http://www.cresco.enea.it/

[3] http://www.enea.it/

[4] http://www.infinibandta.org/

[5] http://www-03.ibm.com/systems/software/gpfs/

[6] http://www.intel.com/

SYNGAS TURBULENT COMBUSTION MODELLING AND SIMULATION

E. Giacomazzi[1], D. Cecere[1], F. Donato[2], F.R. Picchia[1], N. Arcidiacono[1]

[1]ENEA UTTEI-COMSO C.R. Casaccia
[2]ENEA UTFISSM-SICSIS BO Martiri Monte Sole

General Aspects

It is well known that combustion in gas turbines may exhibit pressure oscillations due to poor flame stability. Flame anchoring and thermo-acoustic instabilities are, in fact, a major concern for modern and future combustors, as they tend to employ some form of premixing to reduce NO_X. Efforts are currently focused on fuel lean premixed combustors or partially premixed combustors with rapid mixing after fuel injection, with the tendence to operate close to the Lean BlowOut limit (LBO). However, these combustion strategies are less stable than conventional ones because lean mixtures imply weaker combustion processes and can therefore be easily perturbed, thus increasing the risk of flame blowout. This scenario goes against an acceptable operation of gas turbines.
Today the use of CCS (Carbon Capture and Sequestration) derived fuels is becoming more usual, especially in IGCC (Integrated Gasification Combined Cycle) power plants. The big difference between these fuels (syngases or "hydrogen blends" in general) and natural gas is that they contain hydrogen. In fact, hydrogen has particular molecular properties strongly affecting the features of fuels [1]. First, it tends to enhance flame stability in terms of increased flame speed, decreased ignition delay time, and enlarged flammability limits. This brings designers to use even more leaner mixtures to reduce NO_X. Second, "hydrogen blends", like syngases, have lower LHV (Lower Heating Value) with respect to natural gas. This brings designers to increase the fuel mass flow rate, and hence its injection velocity, to obtain a certain power. Besides these two issues, it has to be noticed that depending on the gasification process and which solid is gasified, substantial differences in the resulting syngas composition occur. Feeding lines contain fuels coming from different sites and inevitably gas turbines will have different thermal loads due to changes in fuel composition. All these design requirements and changes in composition enhance flame dynamics with respect to natural gas (in terms of thermo-acoustic instabilities, flashback, lean blowout, localized extinctions and reignitions) [2] whilst acceptable operation of gas turbines requires a "weak" dynamics.
Looking at the combustion scenario for gas turbine applications, combustion dynamics is expected to be "enhanced" in modern gas turbines operating with lean mixtures or with "hydrogen blends". This implies that unwanted and dangerous phenomena, such as combustion instabilities, flashback and lean blowout, are likely to occur not only in steady conditions, but also and more dangerously when rapid power changes are required. A fundamental question, for example, is to understand when and how pressure oscillations could make the flow dangerously unstable for engine operation.

Hence, unsteady characterization of combustion assumes a central role. If on one hand, the natural unsteadiness of involved physics increases the complexity of the analysis of processes, on the other hand, the increased number of degrees of freedom opens new possibilities to control them. With this end in view, focus on this issue can be reached both using numerical simulation and experimental diagnostics. In particular, this article deals with Large Eddy Simulation.

Description of the test case

The test case under study is the high pressure combustor rig at PSI, shown in *Figure 1*. The combustor, specifically designed to study turbulent, lean premixed flames, is capable of operations up to a pressure of 30 bars, with a maximum air flow rate of 750 mN3/h (0.3 kg/s) and adiabatic flame temperatures up to 2000 K. The combustion air can be electrically preheated up to 823 K. The combustor has a maximum thermal power of 400 kW. A complete optical access is provided by the cylindrical liner consisting of two coaxial quartz glass tubes (inner quartz glass tube diameter D = 75 mm) which are convectively air-cooled, and an axial window. The air/fuel pipe inlet diameter d is 25 mm. The flame is stabilized by the recirculation of hot flue gases due to the sudden expansion geometry. The fuel is injected coaxially into the air stream through five nozzles distributed in the combustor pipe cross section at 40 cm distance upstream of the geometrical expansion zone, where the flame anchors. As confirmed by analysis of the NO_x emissions, the mixing quality is believed to be perfect. Only for starting the rig, a hydrogen torch igniter is used to ignite the premixed fuel/air mixture. More detailed descriptions of the test facility can be found in [3-4].

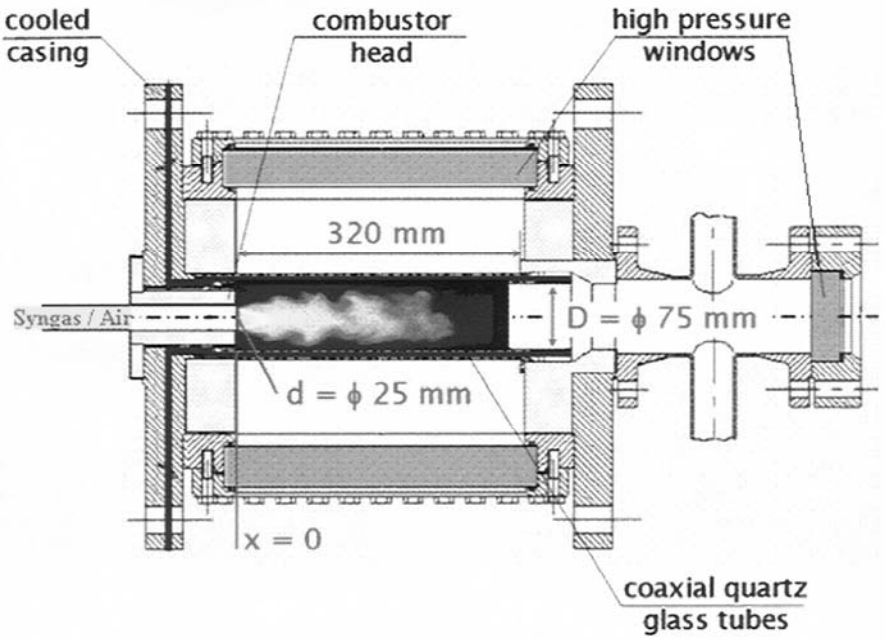


Figure 1: *Sketch of the experimental dump-combustor test rig.*

In particular, the experimental conditions simulated in this work refers to the combustor operated with a mixture of syngas and air (50/50 by volume) with λ= 2.2 (inverse of equivalence ratio). The nominal pressure is 5 bar. The inlet mass flow rate is 0.05 kg/s and the bulk inlet velocity is 40 m/s; the inlet temperature is 623 K; the associated Reynolds number is 82000. A cold (300 K) cross-flow of N_2 is injected in the inlet pipe 15 mm before the combustor dump-plane to reduce flashback risk, with a mass flow rate nearly equal to 1% (0.00046 kg/s) of the inlet mixture. Thermocouples within the dump-plate measured a nearly constant temperature of 807 K during operation.

For the time being, experimental analysis is only based on measurements of OH radicals by means of planar laser induced fluorescence (OH-PLIF) (see *Figures 2-3*). From these measurements, parameters such as turbulent burning velocity, instantaneous and most

probable flame position and flame brush thickness are derived. The setup for the OH-PLIF is described in details in [5].

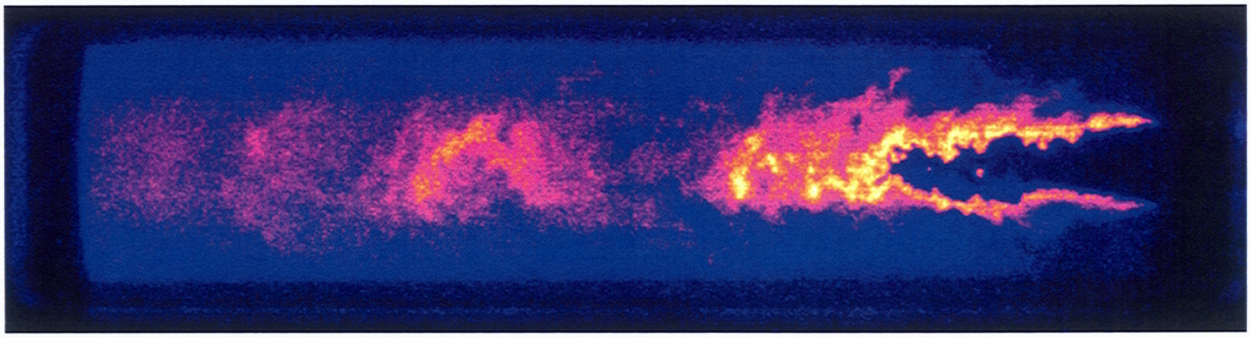


Figure 2: *Instantaneous OH field obtained by means of OH-PLIF.*

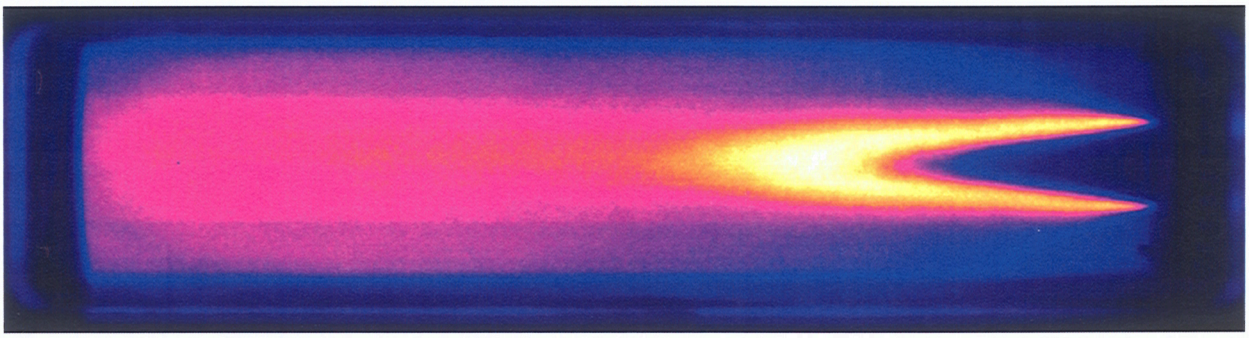


Figure 3: *Averaged OH field obtained by using 400 OH-PLIF samples.*

Numerical simulation set-up

The computational domain consists of two zones: the first is the inlet pipe (extending 0.1 m upstream of the dump-plane) and the second is the combustion chamber (extending up to 0.4 m downstream of the dump-plane). Two different computational grids are used. In the first one, the inlet zone is discretized by means of 276480 nodes distributed as 160x27x64 (axial, radial and azimuthal directions); the main zone is discretized by means of 2621440 distributed as 512x80x64; hence, in total there are 2897920 grid nodes. The maximum ratio between the axial grid size Δz and radial grid size Δr is 2.73, and the minimum is 0.67. The maximum ratio between the azimuthal arc length $\Delta\theta$ and the radial grid size Δr is 6.25, and the minimum is 0.1. The maximum $\Delta\theta/\Delta z$ is 3.70, and the minimum is 0.06. In the second grid, a new node is posed between two old nodes along the axial and radial directions; hence, there are 320x54x64 for the inlet zone, and 1024x160x64 for the main zone, thus having 11591680 grid nodes as total. The ratios $\Delta z/\Delta r$ are the same of previous grid. The maximum $\Delta\theta/\Delta r$ is 12.5, and the minimum is 0.1. The maximum $\Delta\theta/\Delta z$ is 12.34, and the minimum is 0.06.

Concerning chemistry, two chemical mechanisms are used. The first one is the reduced mechanism of Chen [6], consisting of 6 global reactions with 10 transported species (H_2, O_2, H, CO, CO_2, H_2O, OH, O, NO, N_2), based on 70 skeletal reactions with 5 species in partial equilibrium. The second one is the skeletal mechanism of Politecnico of Milan [7-9], referred to as POLIMI, consisting of 37 reactions and 12 transported species (H_2, O_2, H, CO, CO_2, H_2O, OH, O, HCO, N_2, HO_2, H_2O_2). *Figure 4* reports the one-dimensional flame related to the

POLIMI mechanism and to the PSI conditions chosen for the present simulation.

Simulation is firstly performed on the coarser grid with the first chemical mechanism. A spark ignition model [10] is used to ignite the mixture.

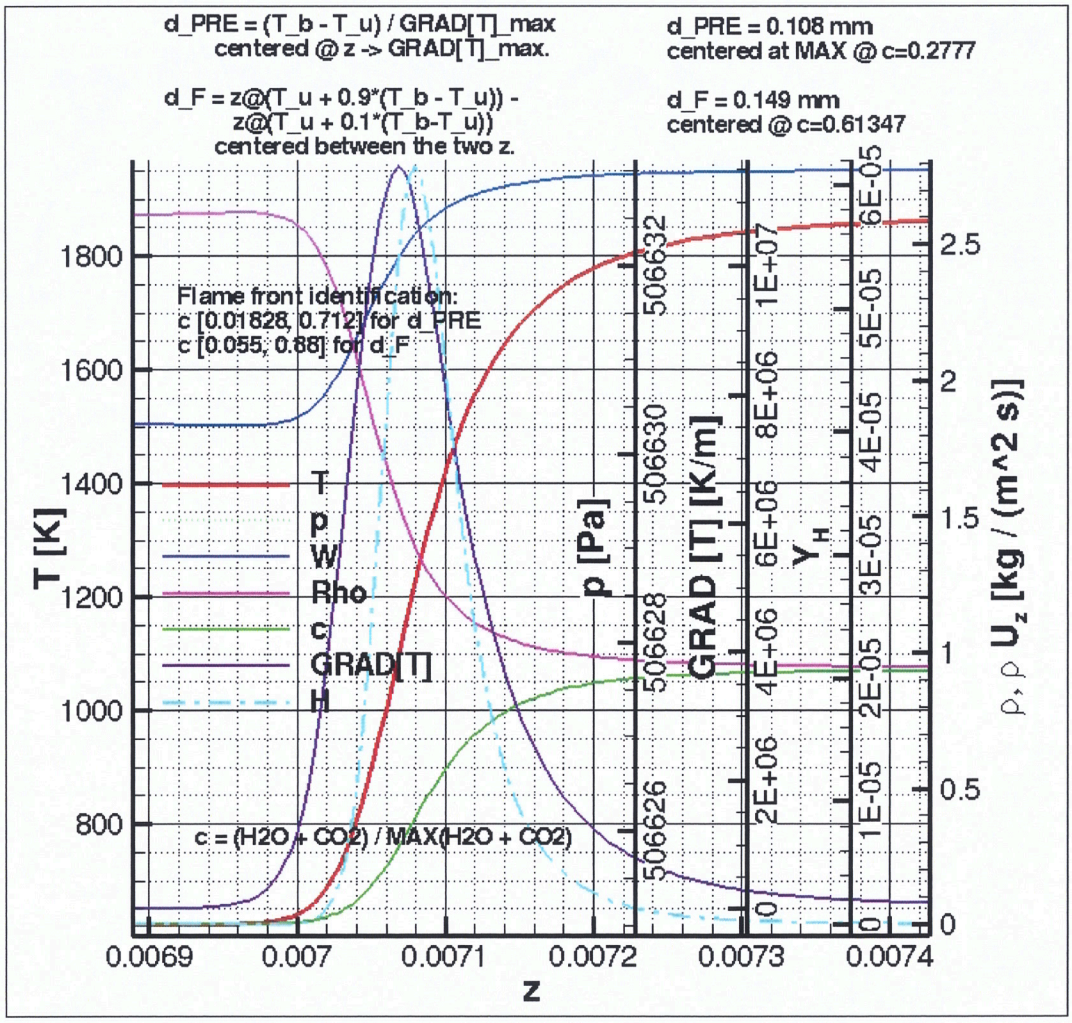


Figure 4: *One-dimensional flame predicted by the POLIMI mechanism.*

Characteristic Navier-Stokes boundary conditions are applied at inlet and outlet [11] with the extension to reactive flows [12]. In particular, the mass flow rate, temperature and composition are imposed at the inlet. The outlet is partially non reflecting, the pressure being relaxed towards an asymptotic value of 5 bar [13]. The outgoing waves at the exit are calculated considering also the contribution of heat release, as suggested in [14]; this procedure stabilizes the computation that otherwise shows an increase of pressure and a strong reverse flow condition at the exit when the flame crosses the outlet thus becoming unstable.

Walls are modelled as adiabatic and viscous, but the dump-plane wall is at fixed temperature as observed in experiments (T=807 K). In the future the wall of the combustion chamber will be cooled as in the experiments, modelling the heat flux subtracted by the cooling air surrounding the combustor (the cooling air is in counterflow with respect to the hot flue gases, enters at 300 K and at the end of the combustor wall is at 770 K).

Results

The flame topology predicted by using the Chen chemical mechanism is that of a flat flame. This topology is completely different from the one observed in the experiment (see *Figures 2-3*) and it is due to an over-predicted laminar flame speed. When switching to the POLIMI chemical mechanism, the flame changes its structure very quickly approaching the actual configuration revealed by experiments.

Looking at the experimental OH fields of *Figure 2* and *3*, it is clear the presence of a non-reacting potential core of mixture with its average vertex located at 0.072 m from the dump-plane. Furthermore, the flame front appears corrugated by small spatial scales. Measurements on the same combustor at the same velocity and pressure conditions but burning natural gas, show that the axial turbulence intensity (with respect to the inlet bulk velocity of the mixture), $u_z'/U_{z\ Bulk}$, is of the order of 15-20% in the shear layer at 0.025 m downstream of the dump-plane. The effect of inlet turbulence intensity was analyzed: increasing turbulence intensity results in a shorter distance before the Kelvin-Helmotz instability start-point and in more turbulent structures developed in the shear layer. The increase in turbulence intensity seems to affect weakly the configuration of the non reacting potential core and of the flame front.

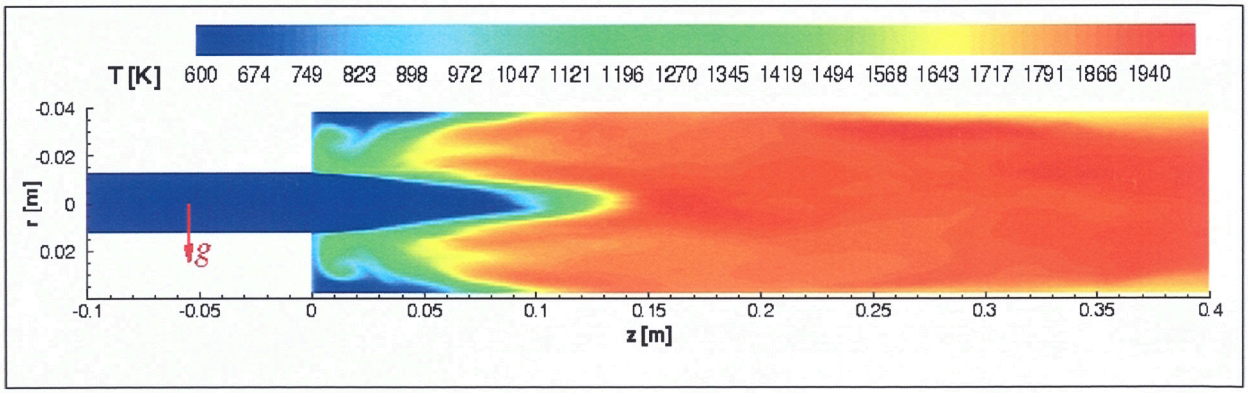


Figure 5: *Average temperature field obtained with the POLIMI chemical mechanism.*

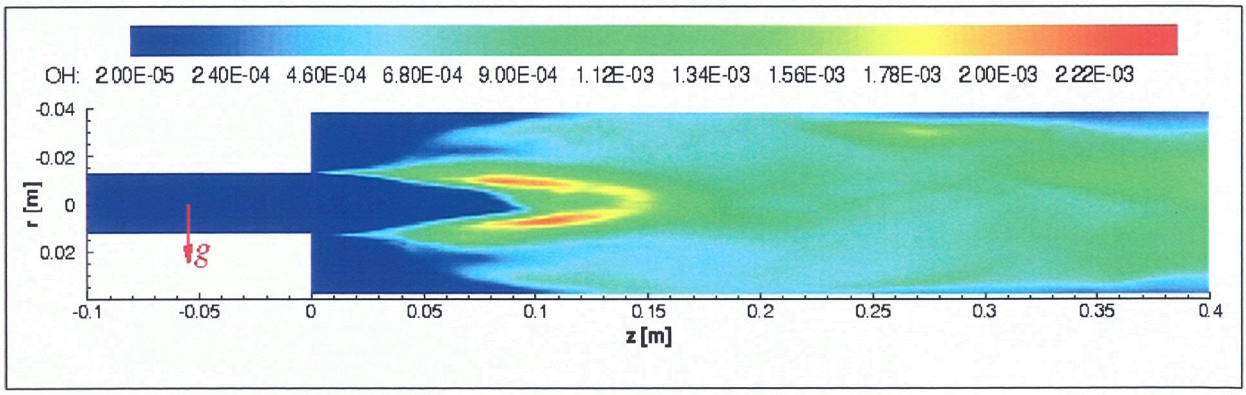


Figure 6: *Average OH mass fraction field obtained with the POLIMI chemical mechanism.*

From a first analysis (see *Figures 5-6*) the predicted flame length is nearly 0.01 m greater than the experimental one. The corrugation of the flame front at small scales is confirmed by numerical simulations. In the corrugated and high curvature regions, preferential diffusion effects are proved to be important. Some modeling improvements have been developed to

better predict flame characteristics. Furthermore, a new approach based on flame progress variable and mixture fraction with tabulated chemistry is currently under study to reduce computational time.

REFERENCES

[1] Lieuwen T., McDonell V., Petersen E., and Santavicca D.: *ASME paper*, 90770 (2006).

[2] Tuncer O., Acharya S., and Uhm J.H. Dynamics: *Int. J. Hydr. En.*, **34**:496–506 (2009).

[3] Boschek, E., Griebel, P., Jansohn, P.: *ASME Turbo Expo*, 27496, Montreal, Can. (2007).

[4] Gricbcl, P., Boschek, E., Jansohn, P.: *J. of Engineering for Gas Turbines and Power-Transactions of the Asme*, **129**(2):404-410 (2007).

[5] Griebel, P., Bombach, R., Inauen, A., Kreutner, W., Schären, R.: *6th European Conference on Industrial Furnaces and Boilers*, p. 45-54 (2002).

[6] Chen J.-Y., *www.ca.sandia.gov/TNF/chemistry.html* (2002).

[7] Frassoldati A., Faravelli T., E. Ranzi E.,: *Int. J. Hydr. Energy*, **31**(15):2310-2328 (2005).

[8] Frassoldati A., Faravelli T., Ranzi E.: *Int. J. Hydr. Energy*, **32**:3471-3485 (2007).

[9] Cuoci A., Frassoldati A., Buzzi Ferraris G., Faravelli T., Ranzi E.: *Int. J. Hydr. Energy*, **32**:3486-3500 (2007).

[10] Lacaze G., Richardson E., Poinsot T.: *Combustion and Flame*, **156**:1993-2009 (2009).

[11] Poinsot T.J., Lele S.K.: *J. Computational Physics*, **101**:104-129 (1992).

[12] Baum M., Poinsot T., Thevenin D.: *J. Computational Physics*, **116**:247-261 (1994).

[13] Polifke W., Wall C.: *Proceedings of Summer Prog. at Cent. Turb. Res.*, Stanford (2002).

[14] Sutherland J.C., Kennedy C.A.,: *J. Computational Physics*, **191**:502-524 (2003).

MATERIALS FOR HYDROGEN STORAGE: NUMERICAL SIMULATION OF HYDROGEN DYNAMICS AT A MG-MgH_2 INTERFACE USING CPMD CODE ON CRESCO HPC

Simone Giusepponi, Massimo Celino

ENEA UTTMAT-DIAG C.R. Casaccia

Hydrogen desorption from hydride matrix is still an open field of research. Extensive *ab-initio* molecular dynamics simulations are performed to characterize the desorption process at the interface MgH_2-Mg. The numerical model successfully reproduces the experimental desorption temperature for the hydride with and without Fe catalyst. Formation energies and work of adhesion are computed and linked to the desorption mechanism. Moreover a detailed analysis of the structural data reveals the role played by the catalysts in the lowering the desorption temperature.

Introduction

The development of a viable hydrogen storage system is becoming increasingly important for promoting the "hydrogen economy". Several materials are good candidate for storing hydrogen, but, among them, magnesium hydride MgH_2 is one of the most promising candidates as hydrogen storage media in automotive industry, due to its very high capacity (7.7 wt%), light weight and low cost of production. However, further research is needed since Mg has a high operation temperature and slow absorption kinetics that prevent for the moment the use in practical applications. Experimentally, great efforts have been made to improve the H_2 absorption and desorption kinetics by mechanically milling MgH_2 with various additives [1] [2] [3]. The main effects of ball-milling are increased surface area, formation of micro/nanostructures and creation of defects on the surface and in the interior of the material. The induced lattice defects aid the diffusion of hydrogen in materials by providing many sites with low activation energy and internal interfaces for hydrogen diffusion [4] [5] [6]. To further enhance the hydrogen diffusion, metal additives, which act as catalysts, are used to destabilize the hydride. These additives are incorporated by ball milling and can be chosen among many possibilities [7] [8].

However the atomic level mechanism by which hydrogen flows in and out of the hydride along the MgH_2-Mg interfaces still remains unclear [9]. In order to have a better understanding of this mechanism, extensive *ab-initio* Molecular Dynamics (MD) simulations have been performed to gain detailed information of the atomic level dynamics during desorption. Electronic structure calculations are used to determine the equilibrium properties and the behavior of the surfaces in terms of total energy considerations. Furthermore, the interface is studied at several values of the temperature, thereby characterizing the hydrogen atomic displacement. The hydrogen desorption happens at the same temperatures measured experimentally [10].

Moreover, to gain more insights on the role played by catalytic metals in MgH_2, the H dynamics at the interface is characterized with and without Fe catalyst. Accordingly with experiments, the temperature of H desorption is lowered significantly in presence of Fe atoms [11] [12].

CPMD code

Car-Parrinello Molecular Dynamics (CPMD) code [13] [14], is an *ab-initio* electronic structure and MD program using a plane wave/pseudopotential implementation of density functional theory. It is mainly targeted at Car-Parrinello MD simulations, but also supports geometry optimizations, Born-Oppenheimer MD, path integral MD, response functions, excited states and calculation of some electronic properties. In *ab-initio* molecular dynamics simulation, the forces acting on atoms are calculated from an electronic structure calculation repeated every time step ("on the fly"). Thanks to electronic structure calculation by using Density Functional methods, simulations of large systems with hundreds of atoms are now standard. Originally developed by Roberto Car and Michele Parrinello for applications in solid state physics and material science, this method has also been used with great success for the study of molecular systems. Applications of *ab-initio* Car-Parrinello Molecular Dynamics simulations range from the thermodynamics of solids and liquids to the study of chemical reactions in solution and on metal and oxide surfaces. CPMD code runs on many different computer architectures and allows good scalability till a large number of processors depending on the system size. We test the CPMD code by using the system representing a MgH_2-Mg interface constituted by 132 atoms of magnesium and 120 atoms of hydrogen. We use a CPMD code compiled with Intel Fortran Compiler, MKL (Math Kernel Library) and MPI (Message Passing Interface) parallelization on the high performance computer CRESCO [15]. The benchmark results demonstrate the high performance of the CPMD parallelization on CRESCO architecture and the good scalability up to hundreds of cores [16] [17].

Computational details

We employed for all the calculations the CPMD code with Goedecker-Teter-Hutter pseudopotentials for magnesium and hydrogen, together with Pade approximant LDA exchange-correlation potentials [18]. The electronic wave functions are expanded in a plane-wave basis set with a kinetic energy cut-off equal to 80 Ry. The latter value was optimized by preliminary calculations both on simple molecules (Mg_2, MgH and H_2) and on the crystalline structures of metallic Mg and magnesium hydride. All the calculations are performed in the supercell approximation, in view of the large-scale molecular dynamics simulation of the interface, with periodic boundary conditions meant to mimic an infinitely extended system.

A supercell approach is used to simulate the atomic system with the interface. As shown in Fig. 1, the MgH_2-Mg interface is built by putting nearby two free surfaces obtained cutting both the Mg and MgH_2 crystals. The main constraint in building the interface is in the selection of two commensurate surfaces (one for Mg and one for MgH_2) fulfilling a simulation cell with periodic boundary conditions. Such a constraint narrows the possibilities of finding two suitable free surfaces for a proper interface. The interface considered is the one composed by facing the (010) of Mg with the (110) of MgH_2. These systems can interact through their internal free surfaces forming an interface as shown in *Figure 1*. Each free surface, selected according to periodic boundary conditions, is structurally optimized. Both internal surfaces are separately studied versus temperature to asses their stability: till T= 900 K there is not surface reconstruction nor hydrogen diffusion. On both sides of the system, a void region is used to suppress the interaction, due to periodic boundary conditions, between the external surfaces of Mg and MgH_2.

Zero-temperature total energy calculations are used to evaluate interface stability. Moreover, formation energy and work of adhesion are computed to characterize the stability of the interface. The latter quantity is defined as the bonding energy per unit area needed to reversibly

separate an interface into two free surfaces, neglecting plastic and diffusional degrees of freedom. Subsequently, MD simulations at constant volume and constant temperature are performed by using Nose-Hoover thermostats [19].

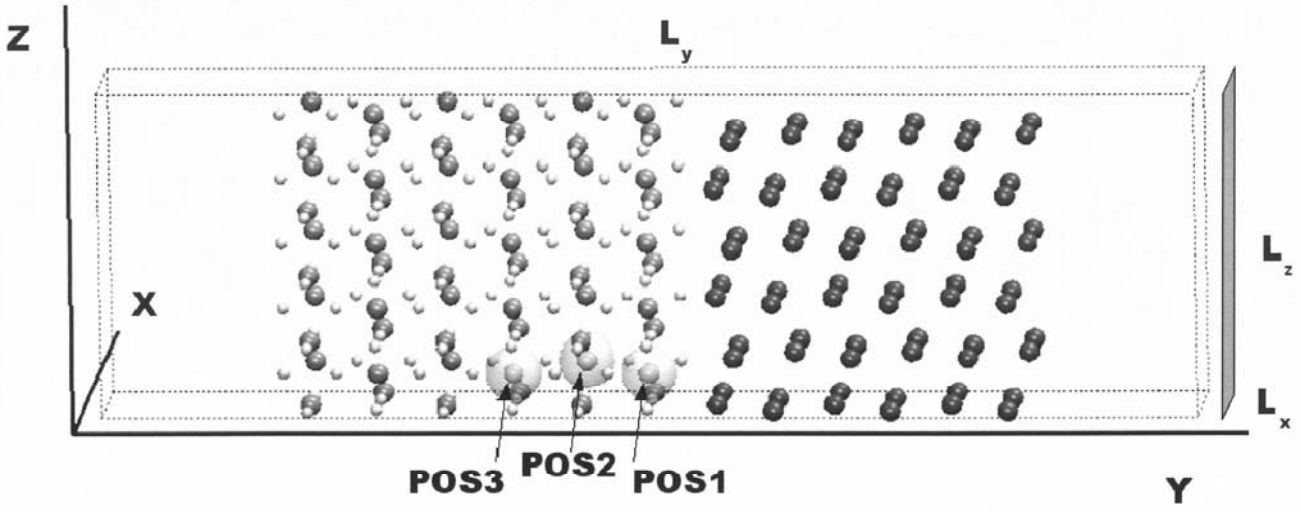


Figure 1: *Schematic of the MgH_2-Mg interface in the simulation box. The structure is infinitely extended in both x and z directions. Free surfaces and MgH_2-Mg interface are perpendicular to the xy plane. H atoms are in white, Mg atoms are in grey (light grey for Mg in the hydride side and dark grey for magnesium crystalline side). POSx indicates those Mg atoms that are substituted by a Fe atom to study the effect of the catalyst at different distances from the interface.*

The MgH2-Mg interface

To fulfill a simulation cell with periodic boundary conditions, commensurable surfaces are needed: both Mg and MgH_2 lattice parameters are scaled of 1% and 2%, respectively. The total length of the system is thus L_x= 6.21 Å, L_y= 50.30 Å and L_z= 15.09 Å. On both sides of the system, a void region of length $2L_x$ is imposed to suppress the interaction, due to periodic boundary conditions, between the external surfaces of Mg and MgH_2.

The Mg system is composed by 72 atoms in 6 layers exposing on both sides the (001) surface (see *Figure 1*): the external surface (on the right) is the same of the internal one (on the left) but shifted by a half lattice parameter in the z direction. The MgH_2 crystalline system is composed by 6 layers: 60 Mg atoms and 120 H atoms (see *Figure 1*). The internal MgH_2 surface faced to the Mg surface is on the right. The external MgH_2 surface (on the left) is the same of the internal one but shifted by a half lattice parameter in the x direction. To find the interface configuration corresponding to the lowest total energy, we perform a geometry optimization moving rigidly by hand in all directions the Mg part keeping fixed the MgH_2 one. Mg atoms of the internal surface of MgH_2 prefer sites that continue the hexagonal sequence of the magnesium hcp bulk across the interface in good agreement with Ref. [20]. Then we perform ionic relaxation to optimize locally the atomic configuration by keeping fixed two layers of Mg atoms on both external surfaces to mimic the bulk behavior. The optimized configuration is characterized by a contraction of the Mg side of about 2%. This displacement goes rapidly to zero involving, on both sides, only Mg rows near the interface. MgH_2 side follows the Mg side adjustments adapting its internal surface to the hcp symmetry of the internal Mg surface. The hydrogen atoms at the interface follow these adjustments losing their linear symmetry parallel to the hydride surface.

To understand the atomic level dynamics of hydrogen diffusion at the interface, MD simulations are performed at constant volume and constant temperature. Starting from temperature T = 300 K the system with interface is warmed up to T = 900 K. Hydrogen diffusion starts at T = 700 K in very good agreement with experimental observations for unmilled MgH_2 without catalyst (T = 670 K). This result makes us confident on the reliability of the numerical system. At T = 700 K the diffusion involves only the H atoms at the interface. In the range from T = 700 K to T = 800 K also internal hydrogen atoms diffuse but only by jumps from one H site to the nearest. Above T = 800 K, also the internal hydrogen atoms diffuse toward the interface causing disorder in the hydride lattice. As expected, H atoms does not diffuse into the Mg lattice. When a stationary configuration is reached, hydrogen atoms in the interface are eliminated and the MD simulation is restarted. Mg atoms of the hydride near the interface adapt themselves to continue the hcp symmetry of the Mg crystalline part freeing behind them another row of hydrogens in the new interface.

The catalytic role of Fe

The catalytic role of Fe in the desorption of H atoms is investigated substituting near the MgH_2-Mg interface an Mg atom with a Fe atom in three different positions. We consider three equivalent Mg atom positions in the hydride side, from the interface: POS1 (first Mg layer), POS2 (second Mg layer) and POS3 (third Mg layer), as shown in *Figure 1*. To determine the influence of these positions on the destabilization of the system, both reduction ΔH in the formation enthalpy and increase ΔW in the work of adhesion are computed for all POSx positions and reported in *Table 1*. The former is an index of chemical destabilization and the latter represents the excess energy that surface atoms have with respect to bulk atoms due to unsatisfied electronic bonds. Formation enthalpy has a maximum when the Fe atom is in position POS2. On the contrary the work of adhesion increases moving the iron atom toward the interface.

	Fe in POS1	Fe in POS2	Fe in POS3
ΔH (eV)	4.42	4.46	4.40
$\Delta W(mJ/m^2)$	170.80	8.88	0.86

***Table 1**: Reduction ΔH in the formation energy and increase ΔW in the work of adhesion respect to an interface without catalyst.*

Geometry optimization reveals structural destabilization caused by the insertion of a iron atom. In fact, significant atomic level distortions of the MgH_2 side in the lattice structure are measured and reported in *Table 2* and *Table 3*. Looking at bonding distances in the MgH_2 bulk structure, we define shells of coordination and we count how many atoms (H or Mg) are inside each shell (H-shell and Mg-shell respectively). An iron atom, independently from its position POSx, shortens the bonding distances with the nearest atoms (both Mg and H) causing changes in the nearest shells of coordination. In particular H atoms in both first and second H-shell lower their bonding distance with Fe atom (from 1.93 Å, and 1.89 Å to 1.61 Å). In the starting configuration there are two H atoms in the 1st H-shell (one in the case of POS1 near the interface) and four H atoms in the 2nd H-shell. In the system with Fe, after geometry optimization all these H atoms are inside the first H-shell at a mean distances of 1.61 Å. Thus the 2nd H-shell is empty and the other H atoms remain at distances greater than 3.30 Å. This H displacement creates free space that is partially taken by Mg atoms in the first and second Mg-shell.

H shell	MgH_2 bulk	POS1 in MgH_2-Mg			POS2 in MgH_2-Mg			POS3 in MgH_2-Mg		
		SC	Mg	Fe	SC	Mg	Fe	SC	Mg	Fe
1st	2 (1.89)	1	0	5 (1.59)	2	0	6 (1.61)	2	0	6 (1.61)
2nd	4 (1.93)	4	5 (1.99)	0	4	6 (1.93)	0	4	6 (1.94)	0
3rd	4 (3.34)	4	4 (3.36)	3 (3.37)	4	4 (3.36)	4 (3.37)	4	4 (3.34)	4 (3.35)
4th	4 (3.56)	2	2 (3.72)	4 (3.70)	4	4 (3.59)	4 (3.58)	4	4 (3.60)	4 (3.59)
5th	8 (3.94)	4	0	4 (3.96)	8	6 (3.99)	7 (3.98)	8	8 (4.02)	8 (3.96)

***Table 2**: Hydrogen coordination of an atom in POS1, POS2 and POS3 (see Fig. 1). For each POSx three columns report the coordination at the starting configuration (SC), after geometry optimization with an Mg in POSx (Mg) and with a Fe in POSx (Fe), respectively. Shells are determined on the base of distances in the MgH_2 bulk. Average distance inside each shell is reported in parenthesis (Å).*

Mg shell	MgH_2 bulk	POS1 in MgH_2-Mg			POS2 in MgH_2-Mg			POS3 in MgH_2-Mg		
		SC	Mg	Fe	SC	Mg	Fe	SC	Mg	Fe
1st	2 (3.02)	2	3 (3.12)	6 (2.79)	2	2 (3.02)	3 (2.92)	2	2 (3.01)	2 (2.81)
2nd	8 (3.45)	9	8 (3.47)	4 (3.59)	8	8 (3.48)	7 (3.38)	8	8 (3.50)	8 (3.37)
3rd	4 (4.39)	6	6 (4.53)	4 (4.36)	4	4 (4.53)	4 (4.48)	4	4 (4.45)	4 (4.40)

***Table 3**: Mg coordination of an atom in POS1, POS2 and POS3 (see Fig. 1). For each POSx three columns report the coordination at the starting configuration (SC), after geometry optimization with in Mg in POSx (Mg) and a Fe in POSx (Fe), respectively. Shells are determined on the base of distances in the MgH_2 bulk. Average distance inside each shell is reported in parenthesis (Å).*

The three optimized configurations with the Fe atom in three different positions in the hydride, are the starting configurations for extensive constant temperature MD simulations. MD simulations are performed with the temperature in the range from T = 300 K up to T = 700 K. Hydrogen displacement already starts at the temperature T = 500 K indicating a lowering of the desorption temperature, in agreement with experimental results [21]. During the simulations the number of both Mg and H atoms in the first two coordination shells of the Fe atom are unchanged respect to starting configurations. As Shown in *Figure 2*, Fe atom has an higher

coordination than Mg atoms in the hydride lattice and this configuration seems stable for all the simulation run. However this distortion creates voids in the second H-shell, partially filled by other Mg atoms.

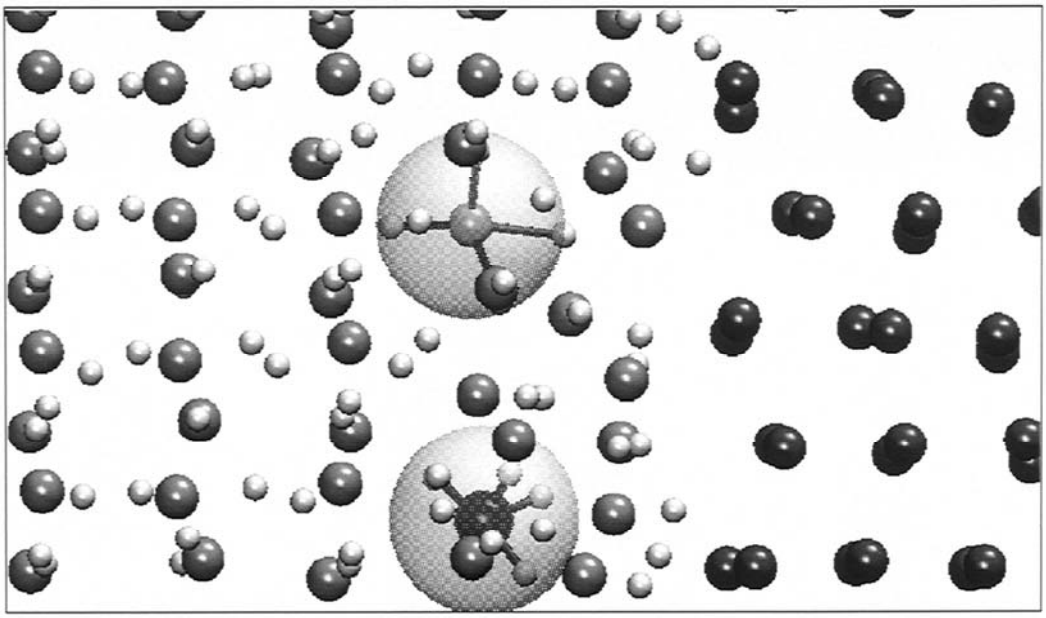


***Figure 2**: Snapshot of the MgH_2-Mg interface with Fe in POS2 at T= 500 K. H atoms are in white, Mg atoms are in light (MgH_2 side) and dark (Mg side) grey and Fe atom is in black. Large transparent circles are used to indicate the first H-shell of an Mg atom (up circle) and of the Fe atom (bottom circle). These circles enlighten the different first-shell coordination of the two atoms.*

The main result of this atomic level rearrangement is evidenced by the MD simulations where H atoms in the third shell of Fe coordination have the possibility to leave their lattice positions toward the interface. This mechanism can be very effective for a Fe atom that is an intermediate situation between interface and bulk environment. As a matter of fact, H atoms in the system with the Fe atom in POS2 are able to start diffusion in shorter times than the other two systems.

Summary

In summary a reliable numerical model to understand the H desorption from MgH_2 is proposed. Desorption temperature is estimated in agreement with experimental results in both cases, with and without catalysts. Moreover the role of Fe catalyst and its interplay with interfaces is clarified in agreement with recent experimental results [21]. The higher coordination of Fe respect to Mg in the hydride raises the probability that H atoms in the third shell of coordination can start diffusion toward the interface. This behavior is more pronounced for a catalyst in an intermediate atomic configuration between bulk and interface. This confirm the idea that a accurate computation of the formation energy can be more predictive in desorption processes than the work of adhesion.

Acknowledgment

We acknowledge ENEA-HPC team for supporting our computational activities on the ENEA-GRID infrastructure. Financial support from Italian Ministry for University and Research under project FISR-TEPSI is acknowledged. We thank also M. Vittori Antisari, A. Montone, A. Aurora and F. Cleri for valuable discussions and suggestions.

REFERENCES

[1] Montone A., Grbovic J., Stamenkovic L., Fiorini A. L., Pasquini L., Bonetti E., Vittori Antisari M., *Mat. Sci. Forum*, **494** (2005) p. 137.

[2] Bassetti A., Bonetti E., Pasquini L., Montone A., Grbovic Y., Vittori Antisari M., *Eur. J. Phys. B*, **43** (2005) p. 19.

[3] Cleri F., Celino M., Montone A., Bonetti E., Pasquini L., *Mat. Sci. Forum*, **555** (2007) p. 349.

[4] Huot J., Liang G., Boily S., Van Neste A., Schulz R., *J. Alloy Compd.*, **293-295** (1999) p. 495.

[5] Hanada N., Ichikawa T., Orimo S.I., Fujii H., *J. Alloys Compd.*, **366** (2004) p. 269.

[6] Vittori A.M., Aurora A., Mirabile Gattia D., Montone A., *Scripta Mat.*, **61** (2009) p.1064.

[7] Sakintuma B., Lamary-D. F., Hirscher M., *Int. J. of Hydrogen Energy,* **32** (2007) p.1121.

[8] Shang C. X., Bououdina M., Song Y., Guo Z. X., *Int. J. of Hydrogen Energy,* **29** ('07) p. 73.

[9] Schober T., *Metall. Trans. A*, **12** (1981) p. 951.

[10] Giusepponi S., Celino M., Cleri F., Montone A., *Solid State Phenomena*, **139** (2008) p. 23.

[11] Celino M., Montone A., Cleri F., Aurora A., Mirabile Gattia D., Giusepponi S., Vittori Antisari M., *Defect and Diffusion Forum*, **297** (2010) p 263.

[12] Giusepponi S., Celino M., *Advances in Science and Technology*, **72** (2010) p. 205.

[13] Car R., Parrinello M., *Phys. Rev. Lett.*, **55** (1985) p. 2471.

[14] CPMD, http://www.cpmd.org/, Copyright IBM Corp. 1990–2008.

[15] Please visit the website www.cresco.enea.it for more information.

[16] Giusepponi S., Celino M., Cleri F., Montone A., *Il Nuovo Cimento C*, **32** (2009) p. 139.

[17] Celino M., Giusepponi S., *Proceedings of the Final Workshop of the Grid Projects of the Italian National Operational Programme 2000-2006 Call 1575*, (2009) p 480.

[18] Goedecker S., Teter M., Hutter J., *Phys. Rev. B*, **54** (1996) p. 1703.

[19] Nose S., *J. Chem. Phys.*, **81** (1984) p. 511; *Mol. Phys.,* **52** (1984) p. 255. Hoover W. G., *Phys. Rev. A,* **31** (1985) p. 1965.

[20] Kelekar R., Giffard H., Kelly S. T., Clemens B. M., *J. Appl. Phys.*, **101** (2007) p. 114311.

[21] Vittori Antisari M., Montone A., Aurora A., Mancini M. R., Mirabile Gattia D., Pilloni L., *Intermetallics*, **17** (2009) p. 596.

CFD SIMULATION OF A LEAN-BURN AERO-ENGINE COMBUSTOR WITH LOW POLLUTANT EMISSION

R. La Gala[1] , F. Ambrosino[2] , A. Funel[2], S. Migliori[3], P. Di Martino[4], S. Colantuoni[4], A. D'Anna[1]

[1]Dipartimento di Ingegneria Chimica - Università degli Studi di Napoli Federico II, Italy
[2]ENEA UTICT-HPC C. R. Portici (NA), Italy
[3]ENEA UTICT Roma Sede (Italy)
[4]AVIO S. p. A. Pomigliano D'Arco (NA), Italy

Abstract

The results of the CFD simulation of a tubular combustor for aero-engine applications developed by AVIO are presented. Simulations have been performed by using OpenFOAM. A Large-Eddy-Simulation approach is used to model turbulence; the compressible Smagorinsky model is used to describe Sub Grid Scale (SGS) motions and their interactions with the solved large-scale structures. A multi-step reaction mechanism and variable thermophysical properties are used to model the oxidation process. The turbulence model has been also coupled with a flamelet approach for the modelling of the detailed chemistry and multiphase processes. The model has been tested for different operating conditions, for which experimental data are available. For these test cases, predicted temperatures, local fuel-air ratios and CO2 concentrations are compared with experimental data. The code has been run on CRESCO, the ENEA high power computing cluster.

Introduction

Reduction in pollutant emissions has always been a target in the modern combustion device concept. Combustors for aero-engine applications are traditionally diffusion-flame combustors because of their performance and stability characteristics. This type of combustor usually produces high levels of pollutants (e.g. thermal NOx). The actual strict legislation for pollutant emissions has induced the aero-engine constructors to develop combustors that can respect the established limits for emissions. In the past, the most used techniques to reduce pollutant emission in aero-engine combustors were based on the injection of fuel diluents, as water and steam, employed to reduce the flame temperature. In spite of the effectiveness of these techniques, this approach determines the reduction of thermodynamic cycle efficiency, and undesirable effects such as quenching of CO burnout and mechanical corrosion of turbine blades due to impurities in the water. For this reason new concepts for combustion technology for the gas turbine industry have been introduced, particularly LPP (Lean Premixed Prevaporised) combustion, when liquid fuels are employed. This technology seems to be the most promising one for practical systems at the present time [1]. Fuel injection and mixing are critical phases to achieve efficiency in combustion in modern gas turbine engines, when they are powered by gaseous or liquid fuels. For gaseous fuels, the major concern is to obtain an optimal level of mixing among the air, fuel, and combustion products in the combustion zone. When liquid fuels are employed, they must be atomized into small droplets and then distributed in an air stream before entering the combustion zone. In addition to its primary function of preparing a combustible mixture, the fuel injector acts as a sensitive element capable of modulating the flow-

field and combustion processes in a combustor. Annular, both co-rotating both counter-rotating swirlers premixed injectors are adopted [2]. A short annular liner is applied in the premixed design to minimize the amount of air required for convective cooling. The combustor volume, however, is approximately twice that of a conventional diffusion combustor, to increase the residence time for complete reactions of CO and unburned hydrocarbons.

LES approach of the turbulence problem

The study of turbulent flow inside the mixing duct of a high-pressure combustion chamber for aero-engine use needs the use of CFD (Computational Fluid Dynamics) software. The fluid flow results in turbulent motions and gradients and requires analysis by CFD to determine trends in the fields of velocity and pressure inside. A Large-Eddy-Simulation approach is used to model turbulence. The decision to directly simulate the turbulent structures belonging to the inertial range is based on the characteristics of the turbulent structures, in fact, the large scale turbulent structures have a dynamic behaviour that depends very much on the current and the direction of observation, on the contrary, small turbulent structure dynamics has an isotropic character. The Large Eddy Simulation, which provides the velocity field, averaged in space, and turbulent structures belonging to the inertial range turbulent spectrum are directly simulated, using a suitably dense grid, while the smaller grid scales of motion (or SGS i.e. Sub Grid Scale), are parameterized with a specific model. The compressible Smagorinsky model is used to describe the SGS motions and their interactions with the solved large-scale structures. In LES the "cut" between what is simulated and what is modelled depends not only on the geometry of the problem but is linked with the position of the inertial range. The closure of LES equations represents the problem of closure and is obtained with the LES turbulence models in sub-grid which should reproduce the effects of unresolved scales on those issues are resolved directly. A multi-step reaction mechanism and variable thermo-physical properties are used to model the oxidation process. The turbulence model has been also coupled with a flamelet approach for the modelling of the detailed chemistry and multiphase processes. This physical problem requires the use of commercial software, and open source availability of OpenFOAM CFD software (release 1.6), makes it preferable to other characterized by considerable licensing fees.

AVIO LPP Injector

The LPP injection system developed by AVIO consists of two co-rotating radial swirlers separated by a venturi profiled wall and a converging pre-vaporization tube. Two principal inlets for primary and secondary air and additional holes are present for dilution air on the converging tube to reduce the risks of flash-back and to promote a secondary atomization. At the centre of the two swirlers there is an atomizer that carburets the flow that is introduced in the mixing duct. The fuel and the air are mixed in the premixing duct before entering into the combustion chamber. The injection system is welded on the water cooling tubular combustor bottom. The tubular combustor has an internal diameter of 100 mm and is 230 mm long. *Figure 1* shows a scheme of the used injector.

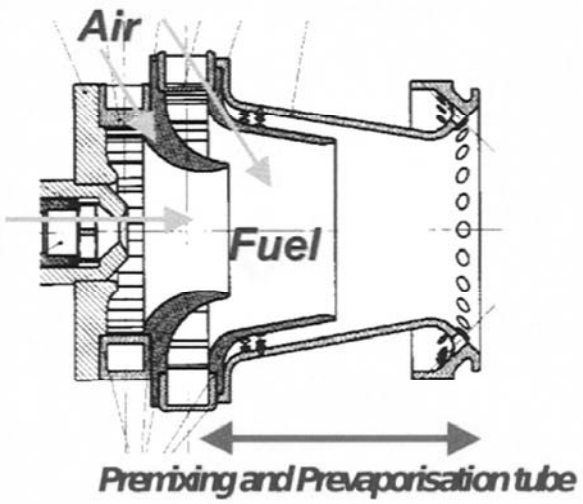


Figure 1: *AVIO LPP injection system.*

Mesh generation and case settings

OpenFOAM has solvers which can solve for different chemical gas reaction and combustion. The utilised solver models compressible reacting flow. The Lagrangian particle tracking approach is utilised to follow the evaporating liquid particles in the three-dimensional domain. The solver utilised for this purpose is named XiFOAM. The case study examined here concerns the flow field in a tubular combustor for aero-engine applications equipped with a LPP injector developed by AVIO. The swirlers have different vane angles respectively of 16.4 and 36 degrees (Swirl Number = 0.29, 0.63). At the end of the first swirler the fuel injector is localized. The injector type is a hollow cone injector. The utilized fuel is n-heptane (C7H16) the whose mass flow rate is equal to 0.01832 kg/s. The case was produced by OpenFoam-1.6 using and modifying standard applications present in it. The geometry and grid were generated by pre-processing tools, equipped in OpenFOAM, run by tutorials present in the software, post-processed and visualized by OpenFOAM integrated visualization application, Paraview [3]. The geometry is outlined in *Figure 2*, in which it is shown the generated grid that includes 1,072,000 cells.

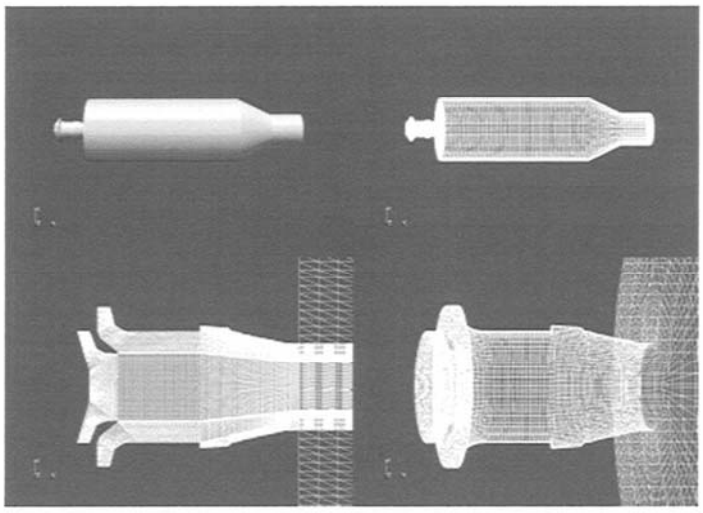


Figure 2: *Mesh generated for the case (up) and inlet detail (down) for mesh.*

The case has been run using XiFOAM tutorial. It is a compressible premixed/partially premixed combustion solver with turbulence modelling [4]. XiFOAM is based on the Weller flamelet combustion model for RANS turbulence models modified properly for LES approach. The equations of motion integrated in XiFOAM tutorials are the classical Navier-Stokes equations:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \rho \vec{U} = 0$$

$$\frac{\partial \rho \vec{U}}{\partial t} + \nabla \cdot \rho \vec{U} \times \vec{U} = -\nabla p + \nabla \cdot \vec{S}$$

$$\frac{\partial \rho e}{\partial t} + \nabla \cdot \rho e \vec{U} = -p \nabla \cdot \vec{U} + \vec{S} \cdot \vec{D} + \nabla \cdot k \nabla e$$

where:

$$S = \lambda \nabla \cdot \vec{U} I + 2\mu \vec{D}$$

$$\vec{D} = \frac{1}{2}(\nabla \vec{U} + \nabla \vec{U}^T)$$

U is the velocity, ρ the density, p the pressure and μ the dynamic viscosity. Turbulence is modelled by LES approach so the equations take the usual Favre-averaged form utilised for LES approach. To obtain the LES equations the Navier–Stokes equations are spatially filtered using an appropriate filter function. The characteristic filter width used in this operation is the smallest resolvable eddy size (which is also the characteristic length scale for the unresolved eddies). In variable density flows, density weighted filtering of variables (called Favre filtering) is often used and is adopted here. Thus, the filtering operation on a general flow variable φ results in the filtered variable is $\widetilde{\Phi} \equiv \rho\overline{\Phi} / \overline{\rho}$.

The case requires initial and boundary conditions settings for all the fields [5]. Boundary conditions in OpenFOAM are attributed splitting boundary faces in a set of patches. These patches were specified using appropriate keywords that define the patch 'type'. Moreover, for each involved field a file exists where initial numerical conditions have to be set. In the present case air flow rate of 0.4196 kg/s is introduced through three split inlets. The percentage in which the air is divided are: primary air = 40%, secondary air = 48.2%, dilution air through the holes: 11.8%. In this case the air comes in radially into the three inlets with an initial value of the velocity respectively of 76, 108, 20.5m/s. The initial pressure field is 18 atm for all the inlet boundaries. Air inlet temperature is equal to 765 K. The coefficients used for the turbulence model are Cμ = 0.09, C1 = 1.44, C2 = 1.92. The value of the turbulent kinetic energy is k = 13.86 m^2 s^{-2} and 28 m^2 s^{-2} at the swirlers inlet, and 1.01 m^2 s^{-2} for the dilution holes. The value of the energy dissipation rate is ε = 4.08E+04 m^2 s^{-3}. We adopted as numerical scheme the standard finite volume discretization of Gaussian integration. The initial boundary conditions are clearly shown in the *Figure 3*.

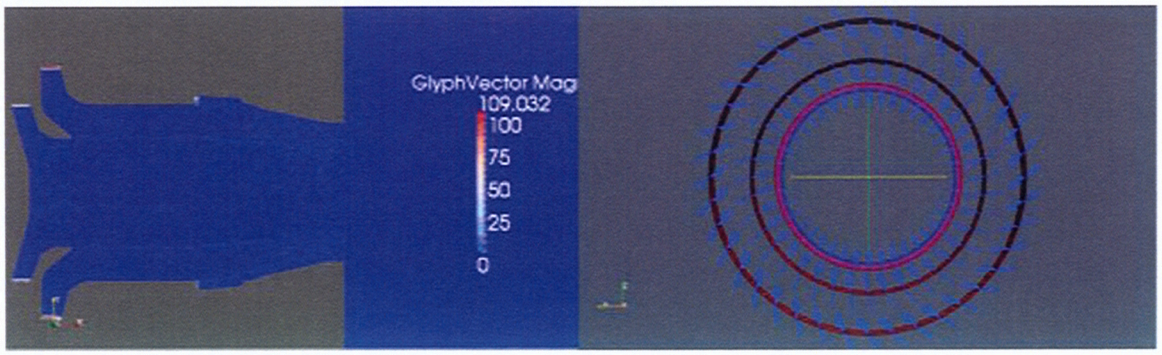


Figure 3: *Defined initial conditions for the velocity.*

Results

Two different analysis are achieved for the case in exam. Primarily, a cold flow (without fuel) analysis in the mixing duct and in the combustion chamber is performed. The study will focus on the cold-flow motion that occurs under non-reactive conditions. In summary, in these conditions the flow will be inviscid, compressible and non-reactive. Under these hypothesis, simplified equations of motion may be written in standard vector and scalar notations. The purpose of this study is to show the air flow distribution in the chamber. This aspect is investigated because in a gas turbine combustor, the cold loss is the most part of dynamic head, while the fundamental loss (additional pressure loss due to combustion) is only about one tenth of the inlet dynamic head. This delineates the importance of the cold flow study in the combustor. For these reasons a lot of studies have been done in the simplified cold flow conditions. This analysis is achieved in the mixing duct to show how the mixing process is occurring. Visualizing the velocity plots (*Figure 4*, left side) of the cold flow we may have an overall view regarding the flow pattern through the whole combustor, instead a more detailed observation can be obtained showing distributions of these quantities in more restricted area of the combustor. Particularly the swirled flow through the two inlets and the radial one through the dilution holes mix themselves in the premixing duct and join together with the fuel injected. After the duct the mixed flow enters into the combustion chamber and undergoes an abrupt expansion due to the difference of diameter between the duct and combustion chamber. A central recirculation zone is visible that is consequence of the swirling flow in the mixing duct. In a second moment the fuel flow is introduced to show how the chemistry and combustion influence the fields of pressure, temperature and velocity in combustion chamber. This second step concerns the reactive flow analysis thanks to a flamelet approach for the modelling of the detailed chemistry and multiphase processes (that concerns the liquid fuel mixing with the air). A multi-step reaction mechanism and variable thermo-physical properties are used to model the oxidation process. Analyzing the reactive flow case (*Figure 4*, right side) the mixture ignites in an asymmetric zone respect to the axis depending on injector position and on boundary conditions. This means that the ignition zone depends both on the injector position and both on the inlet air velocity values (for this reason the importance of cold flow analysis). It is only a preliminary study of the combustion phenomena in the combustor that will be furtherly studied in future simulations.

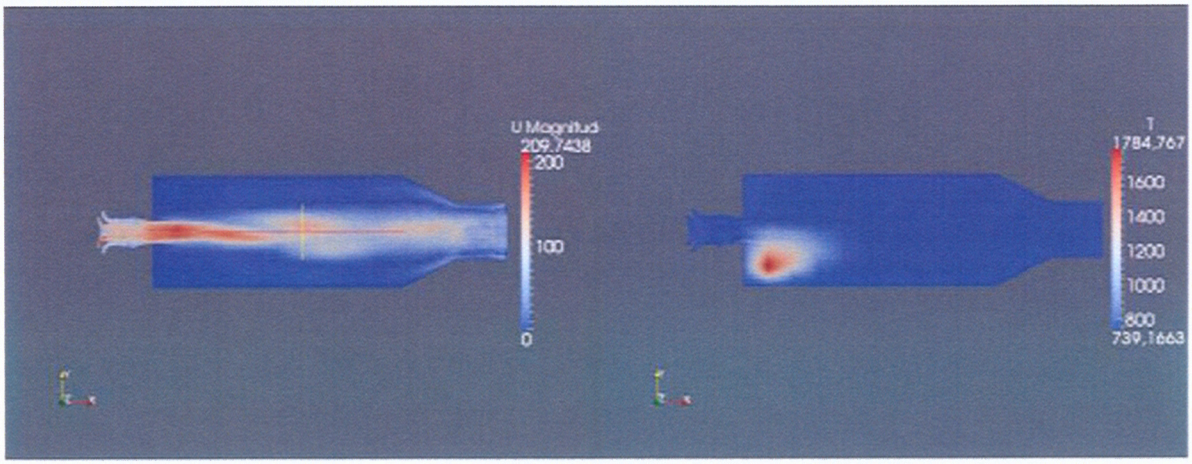


Figure 4: *Results of simulation: cold flow (left) and reactive flow (right).*

A detailed observation of the mixing duct exit may show the flow condition that occur in the expansion zone from mixing pipe to dump combustion chamber. Another important aspect is to investigate ,in the reactive case, where the ignition and the combustion occur, to prevent possible flash-back phenomena. At the end the exam of the temperature profile in the overall combustor, gives a prediction for the possible NOx formation zone.

Conclusions

In summary we have shown how the combustion can occur in a new conception combustion chamber for aero-engine applications. The turbulent mixing flow in the case in exam shows a good degree of mixing mainly thanks to the presence of the two swirlers, and after it ignites in an optimal position in the combustion chamber giving no problems for flash-back phenomena. Also the reached temperature in the chamber are not so high connected with problems of NOx formations. Further investigations, mainly on chemical and combustion aspects will be realized in future simulations.

REFERENCES

[1] Huang Y.,Yang V., Progress in Energy and Combustion Science, **35,:** 293-364 (2009).

[2] Lu X., Wang S. Hsieh S., Yang V. , J. Fluid Mech., **527,** 171-195 (2005).

[3] http://www.opencfd.co.uk/openfoam/doc/user.html

[4] http://www.opencfd.co.uk/openfoam/doc/programmer.html

[5] Page M., Nilsson H. Petit O., Bounous O. ,Third OpenFOAM Workshop Milano Italy, 10-11 July 2008.

CRYSTALLINE AND LIQUID Si_3N_4 CHARACTERIZATION BY FIRST-PRINCIPLES MOLECULAR DYNAMICS SIMULATIONS

Aurelio Mauri

MICRON-NUMONYX, Via Olivetti 2, Agrate Brianza (MI), Italy

Massimo Celino

ENEA, C.R. Casaccia, via Anguillarese 301, Rome, Italy

Silicon nitride (Si_3N_4) has a wide range of engineering applications where its mechanical and electronic properties can be effectively exploited. In particular, in the microelectronics field, the amorphous silicon nitride films are used as charge storage layer in metal-alumina-nitride-oxide nonvolatile memory devices. Atomic structure of amorphous silicon nitride is characterized by high concentration of traps that control the electric behavior of the final device by the trapping-de-trapping mechanism of the electrical charge occurring in its traps. In order to have a deep understanding of the material properties and, in particular, the nature of the electrical active traps a detailed numerical characterization of the crystalline and liquid phases is mandatory. For these reasons first-principles molecular dynamics simulations are extensively employed to simulate the crystalline Si_3N_4 in its crystalline and liquid phases. Good agreement with experimental results is obtained in terms of density and formation enthalpy. Detailed characterization of crystalline Si_3N_4 electronic properties is performed in terms of band structure and band gap. Extensive first-principles molecular dynamics simulations are performed to obtain a reliable liquid sample. Detailed characterization of the atomic structure is achieved in terms of radial distribution functions and total structure factor.

Introduction

In the recent years new non volatile memory devices have been introduced to overcome the emerging scaling limitation in the standard floating gate approach due for example to the statistical charge injection, cell to cell interference and technology complexity. One of the most promising advanced memory for replacing the NVM-NAND architecture is the so called Charge Trapping Devices [1] based on the discrete trapping of electronic charge in the material defects: in general an amorphous Si_3N_4 layer of few nanometer has been employed in the actual memory device preparation. Typically silicon nitride layer is deposited at high temperature (> 700 °C) by the help of Chemical Vapour Deposition technique with stoichiometric or Si-rich ratio obtained by changing reactant gas flow rate. Due to the recent technology approach, i.e. 3D geometry shape [2], material intrinsic defects become more relevant than the one produced by implantation damage or induced by spurious element such as Boron or Phosphorus. To explore the formation of a such defects we use a different approach with respect to Ref. [3] based on cluster models taking into account the dynamical evolution at the atomic level of defects and their environment. Our approach is similar to that one reported in Ref. [4] where first-principles molecular dynamics simulations are used to produce a reliable amorphous sample of Si_3N_4. Our approach allows the study of a larger system on longer time scales for a full characterization of the Si_3N_4 system.

Si_3N_4 crystalline α phase

Si_3N_4 presents two crystalline hexagonal phases: α (shown in *Figure 1*) and β belonging to point group c1 and c6h, respectively. We compare our results with the experimental ones reported in Ref. [5] and in Ref. [6]. We consider a simulation cell composed by 28 atoms (12 Si and 16 N).

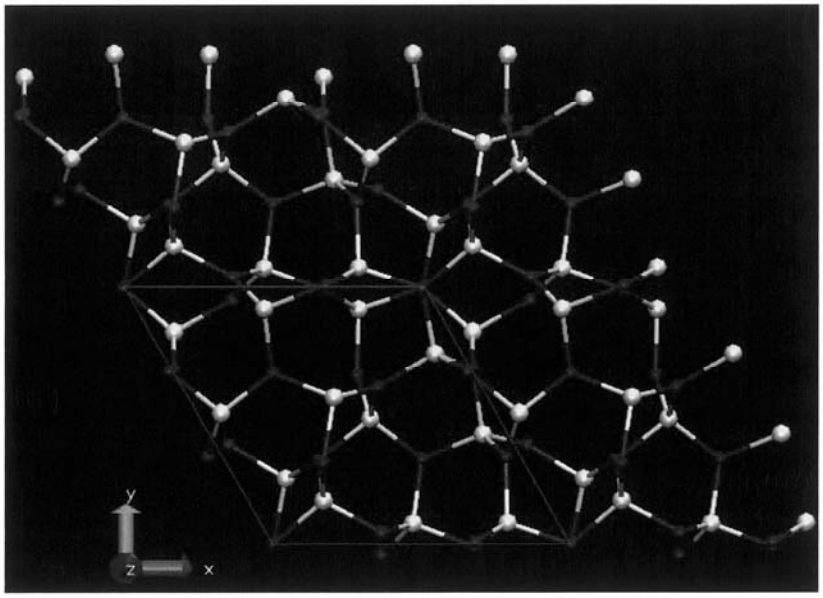


Figure 1: *Atomic structure of Si_3N_4 material in its crystalline α phase. Spheres are the atoms and segments indicates the nearest neighbors atoms of each atom.*

At the experimental values of the lattice parameters (α phase: a= 7.753Å, c= 5.618 Å; β phase: a= 7.706 Å, c= 2.909 Å) we compute the electronic band structures in very good agreement with the experimental results [6]. In *Figure 2* the electronic density of states (EDOS) of the crystalline α phase is reported. The band gap is formed between Γ and K points in good agreement with Ref. [6]. In *Figure 3* we report the calculated total density of electronic states where an HUMO-LUMO band gap of 4.63 eV is found. This value is in good agreement, within 8 % of error, with the experimental one (4.9-5.1 eV) [6].

Computational details

To produce a reliable liquid configuration, extensive first-principles Molecular Dynamic simulations have been performed by using the standard Car-Pirandello method [7]. All calculations are performed at constant volume on a system of 223 atoms: Si-N mixture with a ratio of r= [N]/[Si]= 1.347 in order to increase the probability to find point defects. The periodically cubic cell has a size of 13.91 Å corresponding to the experimental density of amorphous Si_3N_4 phase (3.1 g/cm) at room temperature. The starting configuration has been chosen by randomly generating the coordinates of each atom discarding positions with distance less than 1.74 Å from every other atom. This procedure assures a faster and more accurate thermalization process. Norm conserving pseudopotentials are used in conjunction with a generalized gradient approximation (GGA) for the exchange and correlation part of the total energy. The wave functions are expanded in plane waves at the Γ point of the super cell. A cutoff

of 140 Ry yields converged properties for simple molecules (SiSi, NN and SiN dimers) and reproduce accurately electronic properties of the crystalline phases, as shown in the previous section. A temperature control is implemented for both ionic and electronic degrees of freedom by using the Nosé-Hoover thermostats. A simulation time of about 5 ps is used to thermalize the disordered system at high temperature (T= 3200 K) to assure that the system has lost memory of its initial conditions. Physical quantities are then calculated and averaged on a trajectory lasting about 7 ps at the temperature T= 1800 K.

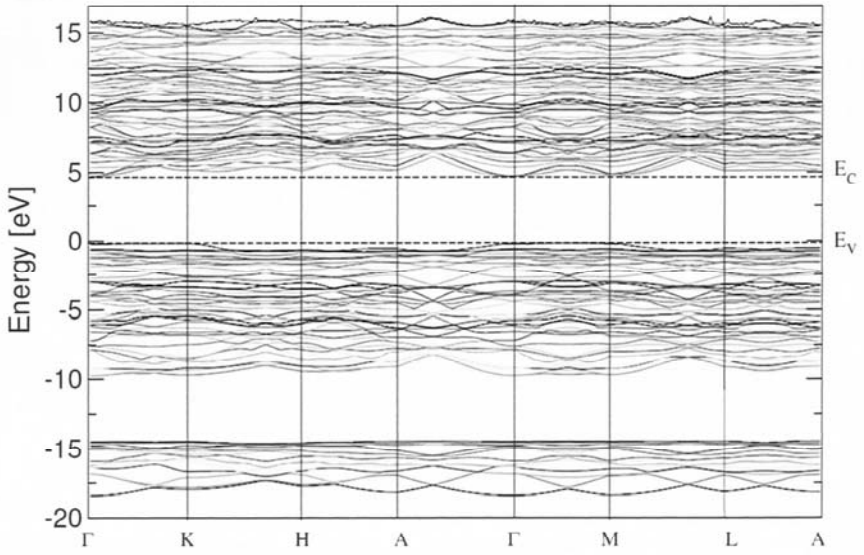


***Figure 2**: Electronic density of states of α Si_3N_4 material*

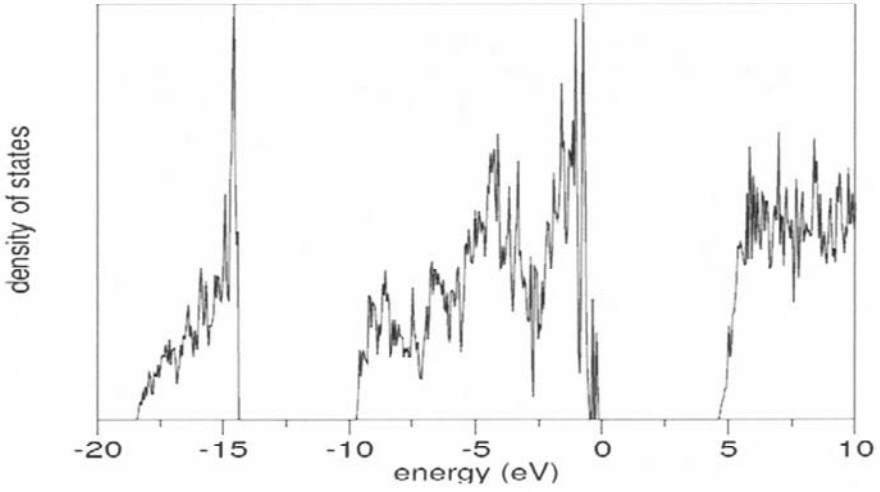


Figure 3: Total electronic density of states of α Si_3N_4 material.

Si_3N_4 liquid phase characterization

To analyze the short-range order of the obtained liquid phase we have calculated the total radial distribution function, reported in *Figure 4*. The comparison with the experimental one reported in Ref. [8] confirms the reliability of our numerical model. It is interesting to note that our model well reproduces the very first peak at 1.15 Å. Partial radial distribution functions reported in *Figure 4* reveal the origin of this first peak in the total one. This peak is not related to the presence of hydrogen in the sample as explained in Ref. 4 but to the presence of small but not negligible quantity of N-N bonds. The distance distribution for S-N bond is peaked at 1.80 Å, for Si-Si at 2.38 Å while in the case of N-N bond at 1.15 Å and 2.73 Å.

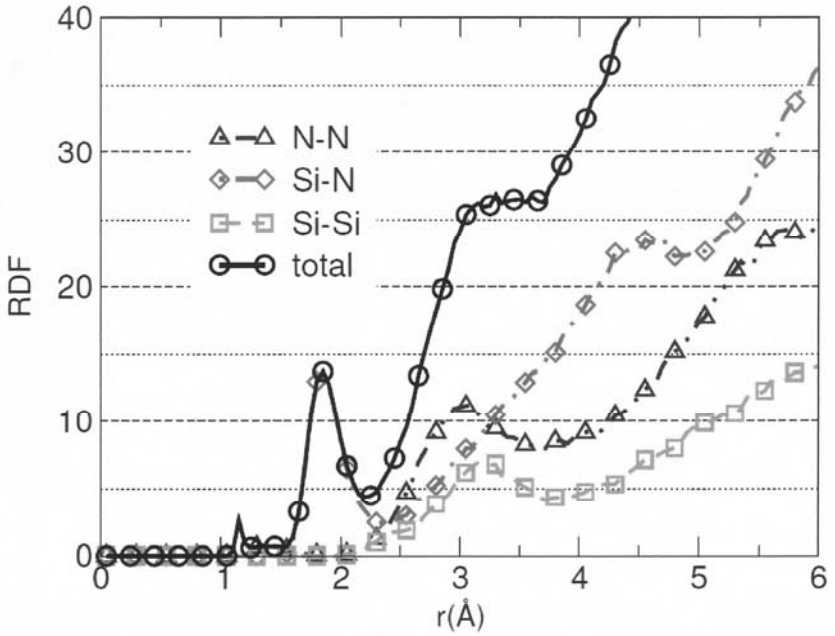


***Figure 4**: Radial distribution function of liquid Si_3N_4*

The angular distribution is shown in *Figure 5*. The N-Si-N distribution has a peak at 106.5° very close to the peculiar value of regular tetrahedra. The Si-N-Si has instead a peak at 115° indicating the presence of planar NSi_3 units. Both *Figures 4* and *5* confirm the picture of a liquid system composed essentially by Si atoms fourfold coordinated and only few threefold and fivefold coordinated. On the contrary, almost all the nitrogen atoms are bound to three silicon atoms. Homopolar bonds are essentially due to nitrogen atoms. Thus the liquid Si_3N_4 system can not be considered completely chemically ordered. This picture explains the great interest of this system for applications in microelectronics.

To verify the reliability of our approach also on long length atomic scales, we have computed the structure factor. The agreement is very good over the entire k range with experimental results reported in Refs. [4,8], the position of maxima and minima being accurately reproduced. As expected, the shape of the structure factor of our liquid sample is less structured than the experimental one obtained on an amorphous system. In particular the width of the peaks is more pronounced and their height significantly lower.

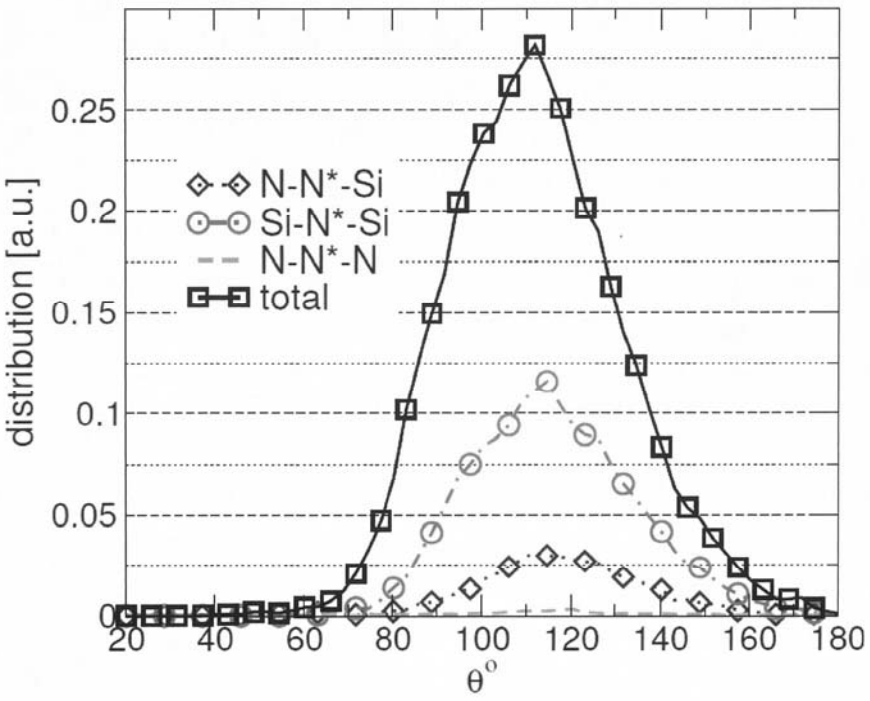


***Figure 5**: Angular distribution of liquid Si_3N_4*

Conclusions and further work

In this paper our first goal has been to achieve a precise characterization of the liquid atomic structure of the Si_3N_4 system. This liquid sample, that we have demonstrated a reliable representation of the real liquid phase of the Si_3N_4 system, can be considered a good starting point for a quenching process toward the amorphous phase. We have shown that our theoretical approach is able to model accurately all the main features of the liquid system both in the short and long range length scales. Moreover we are confident that this procedure can be applied successfully to produce and characterize the amorphous phase. This approach will allow to study accurately the relation between atomic scale intrinsic defects and the macroscopic structural and electronic properties of the Si_3N_4 system.

Acknowledgment

We acknowledge ENEA-HPC team for supporting our computational activities on the ENEA-GRID infrastructure.

REFERENCES

[1] C. Monzio Compagnoni, A. Mauri, S. M. Amoroso, A. Maconi, A. S. Spinelli, IEEE Trans Electron Devices, **56** (2009) pp. 2008-2015.

[2] H. T. Lue, proc. of IEEE IEDM, (2009) pp. 839.

[3] G. Pacchioni, D. Erbetta, Phys. Rev. B, **61** (2000) pp. 15005.

[4] L. Giacomazzi, P. Umari, Phys. Rev. B, **80** (2009) pp. 144201.

[5] N. Ruddlesen, P. Popper, Acta Crystal., **11** (1958) pp. 465-468.

[6] B. Shaposhnikov, Phys. Solid States, **49** (2007) pp. 1629.

[7] R. Car, M. Parrinello, Phys. Rev. Lett., **55** (1985) pp. 2471.

[8] M. Misawa, T. Fukunaga, K. Niihara, T. Hirai, K. Suzuki, J. Non-Cryst. Solids **34** (1979) pp. 313-321.

FULLERENE DERIVATIVES AS ELECTRON ACCEPTORS IN POLYMER SOLAR CELLS: A QUANTUM CHEMICAL STUDY

Pasquale Morvillo

ENEA UTTP-NANO C.R. Portici

Introduction

The conversion of solar energy into electricity is considered a promising solution to solve the energy crisis the world is facing today. In the last years, polymer solar cells (PSC) have attracted a considerable interest as a low cost alternative to current silicon solar cells [1, 2]. The most efficient and promising device structure (and also the most studied) is based on a donor (semiconducting polymer) / acceptor (fullerene derivative) bulk heterojunction (BHJ) in which the active layer is sandwiched between two electrodes with different work functions [3, 4] (see *Figure 1*).

The most studied donor polymer is the regioregular poly(3-hexylthiophene) (P3HT); PSC with a power conversion efficiency (PCE) above 5% have been realized using a blend of this polymer and the [6,6]-phenyl-C_{61}-butyric acid methyl ester ([60]PCBM) [5]. Actually, the best PCE reported up to now (7.9%) has been obtained using other materials combinations [6].

Anyway, the commercial application of these solar cells requires a further increase in the PCE (although other factors, like the stability and the large-area processing, are also important) [7, 8].

The power conversion efficiency for a solar cell is directly proportional to the short -circuit current density (J_{sc}) and to the open-circuit voltage (V_{oc}).

The J_{sc} is strictly related to the absorption of the blend film. Since [70]PCBM has a stronger and broader absorption compared to [60]PCBM, it has been proven that this acceptor can improve the current output of the corresponding PSC [9, 10].

Moreover, in devices with ohmic contacts, the V_{oc} is proportional to the energy difference between the highest occupied molecular orbital (HOMO) of the electron donor and the lowest unoccupied molecular orbital (LUMO) of the electron acceptor [11, 12, 13] In addition, for an efficient exciton dissociation, also the relative position of donor LUMO and acceptor LUMO is crucial.

The aim of this work is to use quantum chemical methods to predict the acceptor strength of different C60 derivatives in order to understand the effect of substituents on the reduction potential of the fullerene moiety and to foresee the Voc in polymer-fullerene solar cells.

The electronic structure of the fullerene derivatives already used in the fabrication of polymer-fullerene solar cells are investigated using the density functional theory (DFT). The theoretical LUMO levels are correlated to the experimentally determined Voc of the corresponding BHJ solar cells with poly(2-methoxy-5-(3',7'-dimethyloctyloxy)-1,4-phenylenevinylene) (MDMO-PPV) with similar device structure.

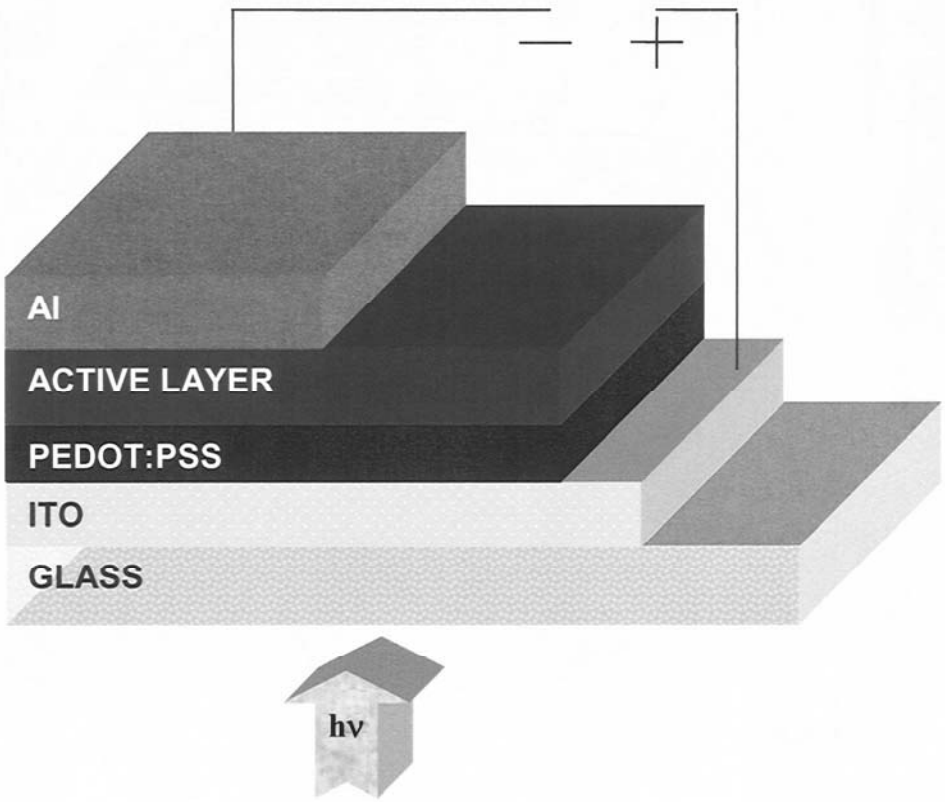


***Figure 1**: Schematic device structure of a polymer/fullerene bulk heterojunction solar cell. The active layer is sandwiched between two contacts: an indium-tin-oxide electrode coated with a hole transport layer PEDOT:PSS and an aluminium top electrode.*

Computational Method

All calculations were carried out using the Firefly package [14], which is partially based on the GAMESS (US) [15] source code.

Firefly (previously known as the PC GAMESS) is a freely available ab initio and DFT computational chemistry program developed to offer high performance on Intel-compatible x86, AMD64, and EM64T processors. The project coordinator and leading developer is Dr. Alex A. Granovsky. Firefly has been constantly being developed to incorporate new functionality, improve performance, and extend existing features. For example, Firefly uses real time data compression/decompression, efficient modern algorithms of 2-e integral evaluation for direct calculation methods, very efficient MP2 energy and energy gradient modules, very fast RHF MP3/MP4 energy code, and state-of-the-art DFT, TDDFT, MCSCF, MCQDPT, and XMCQDPT implementation. Firefly runs parallel on SMP systems, clusters of computers, or both; with special attention paid to good scalability even on large clusters and many-core systems [14].

The geometric structures of the investigated fullerene derivatives were optimized using DFT method with 3-21G* basis set [16, 17] and Becke's three-parameter functional (which defines the exchange functional as the linear combination of Hartree-Fock, local and gradient-corrected exchange terms) in combination with the correlation functionals of Lee, Yang and Parr (B3LYP) [18, 19]. Since these supramolecules contain over 60 atoms, a smaller basis set is chosen to save computation time. Indeed, this method has proved to be very adapted to study macromolecular systems with fullerene units [20]. Afterwards, the localization of the frontier orbitals was investigated using a single point calculation at the same level of theory. The prediction of the

geometries of fullerenes using DFT B3LYP/3-21G* is fairly good although this agreement is accredited to fortuitous error cancellation from exchange-correlation functional and basis set.

Results and Discussion

The geometries of C_{60} and [60]PCBM were optimized using the DFT (B3LYP/3-21G*) method. The obtained structures (reported in *Figure 2*) are consistent with the literature and the prediction of bond lengths and angles is in excellent agreement with the experimental values [12].

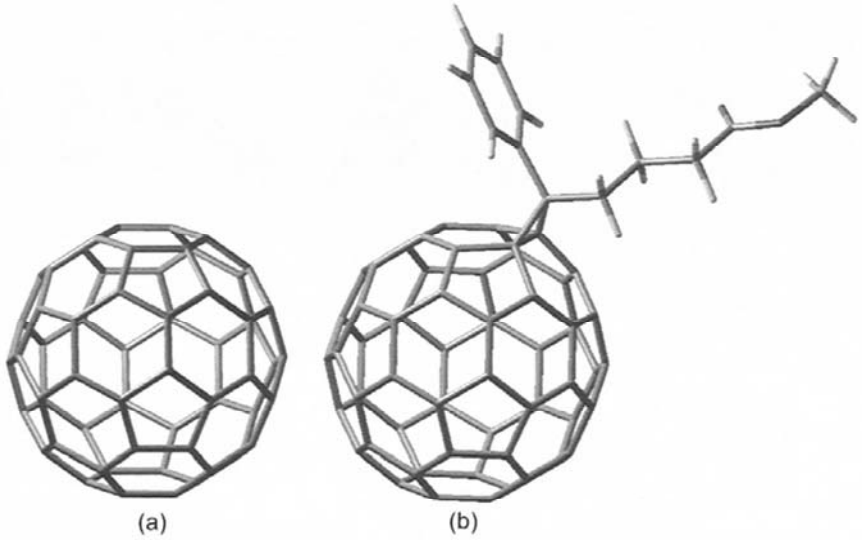


***Figure 2**: C_{60} (a) and [60]PCBM (b) structure optimized using DFT B3LYP/3-21G* method.*

The influence of the addend on the geometry of the C_{60} is restricted to the neighbouring pentagons and hexagons. The lengths for the bonds of the other rings are very close to the values reported for the C_{60} molecule (1.458 Å for the bond fusing five- and six-member rings and 1.401Å for the bond connecting six-member rings).

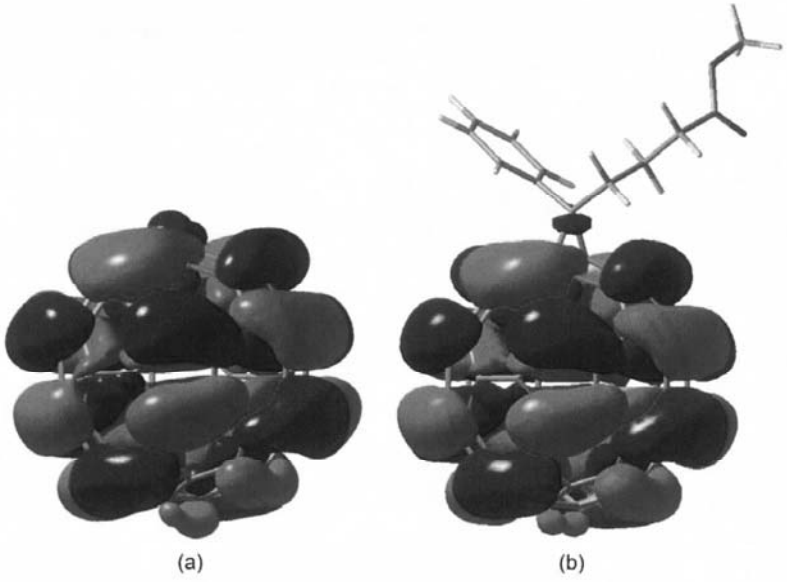


***Figure 3**: Isosurface of the LUMO level of a) C_{60}, b) [60]PCBM obtained using DFT B3LYP/3-21G* method.*

In *Figure 3*, the Kohn-Sham LUMO levels of C_{60} and [60]PCBM are shown. Also in [60]PCBM, it is fully located on the fullerene cage (i.e. there is no intramolecular photoinduced electron transfer to the addend). By the way the value of the LUMO level of [60]PCBM is significantly raised compared to C_{60}, minimizing the energy losses in the electron transfer from the LUMO level of the donor polymer and increasing the gap between the HOMO of the donor and the LUMO of the acceptor (resulting in an increase of the open circuit voltage of the corresponding device).

Recently, a new class of [60]PCBM derivatives have been proposed with the aim to influence the reduction potential of these species by placing substituents on the phenyl ring [21]. Electron-donating (methoxy groups) substituents have been successfully used to raise the LUMO levels of these species. The studied molecules were tested as electron acceptors in blend with MDMO-PPV in the bulk heterojunction solar cells and it was possible to increase the V_{oc} of the corresponding device.

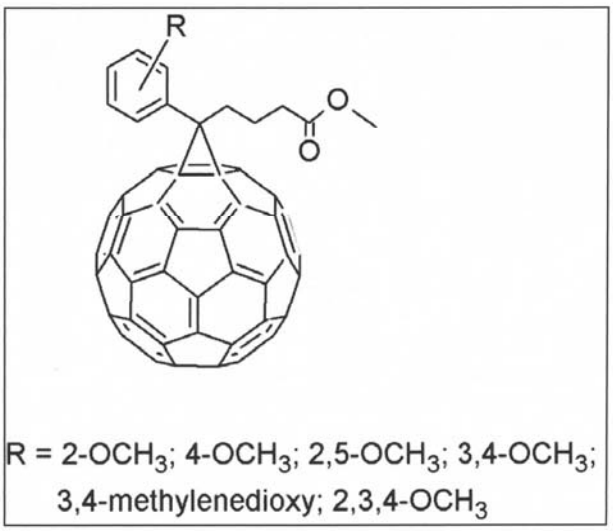


***Figure 4**: Chemical structure of the investigated fullerene derivatives.*

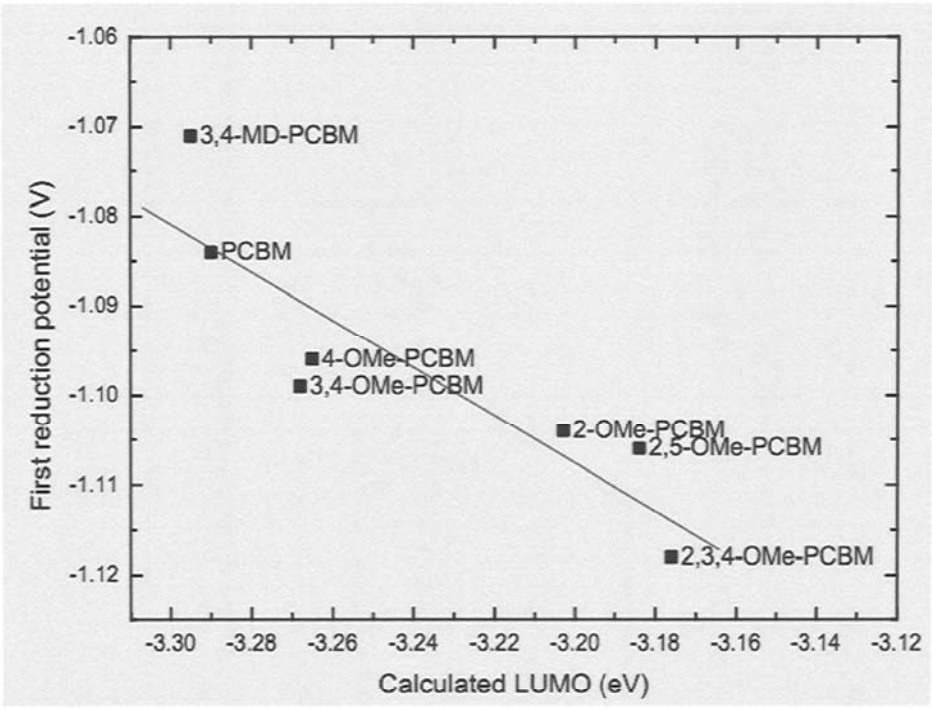


***Figure 5**: First reduction potentials of fullerene derivatives [21] versus LUMO levels calculated with DFT B3LYP/3-21G* level of theory.*

The LUMO level of the fullerene derivatives reported in *Figure 4* was calculated using the DFT (B3LYP/3-21G*) method. In *Figure 5*, the first reduction potential (V vs. Fc/Fc+), obtained from cyclic voltammetry data [21], is reported versus the calculated LUMO level [22, 23]. The correlation is very nice although the LUMO levels were calculated for isolated molecules in gas phase, while the electrochemical measurements were made in solution.

This correlation can be useful to foresee the first reduction potential of unknown fullerene derivatives once the LUMO level has been calculated. This relationship was also proven for fullerene derivatives that differ not only by a side chain: the calculated LUMO, LUMO+1 and LUMO+2 levels of [60]PCBM, N-3-(2-ethylhexyloxy)benzyl azafulleroid, C_{60} and N-3-(2-ethylhexyloxy)-benzyl ketolactam correlate well with first, second and third reduction potential of these species in solution.

The effect of methoxy group on the LUMO level of [60]PCBM decreases going from the ortho to the para position. So it is important that the electron donor substituent is closer to the fullerene cage. Furthermore, a significant influence on the LUMO level is observed with increasing the number of alkoxy substituents.

In *Figure 6* and *7*, the V_{oc} of the photovoltaic devices realized using these acceptors and MDMO-PPV, as donor polymer, are reported as a function of the calculated LUMO levels.

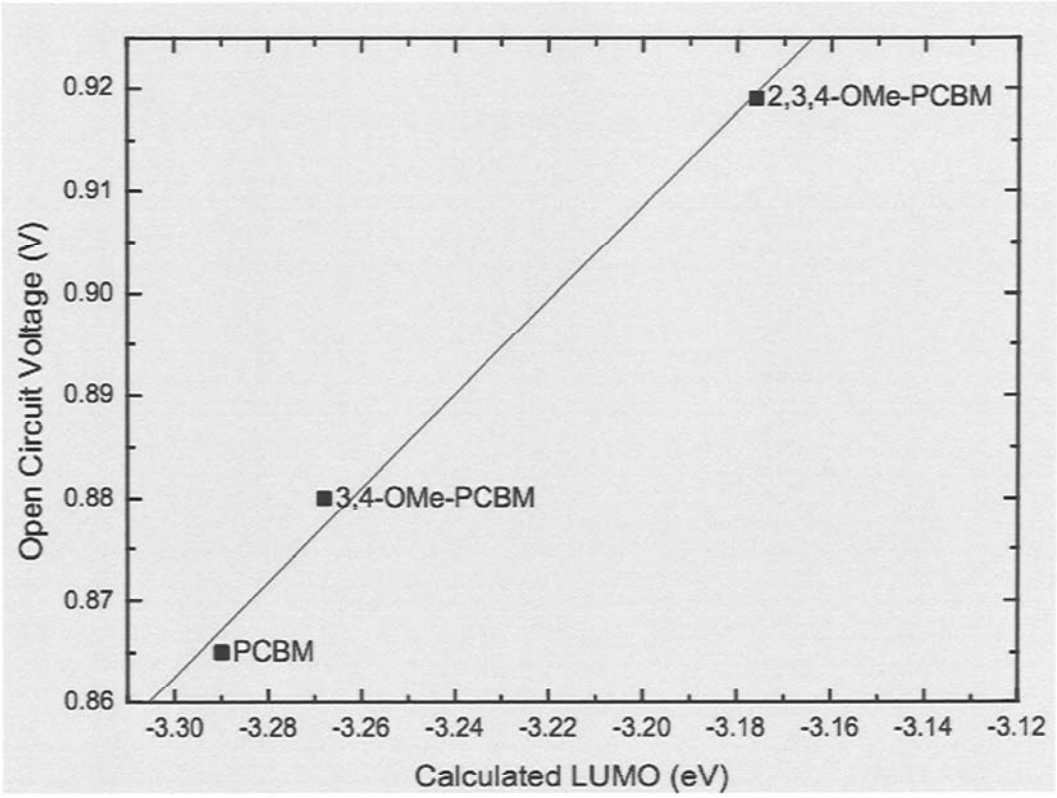


***Figure 6**: Open circuit voltage of MDMO-PPV:fullerene devices [21] versus theoretical LUMO levels (B3LYP/3-21G*) of acceptors for blend film dissolved in chlorobenzene.*

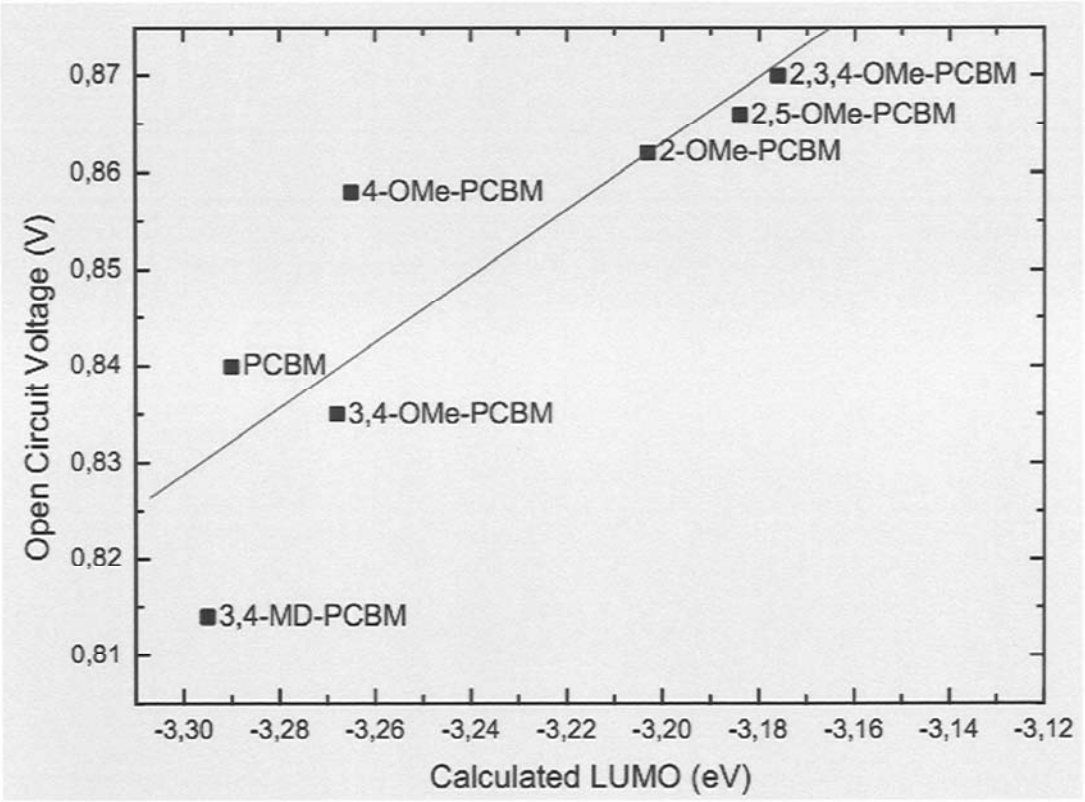


***Figure** 7: Open circuit voltage of MDMO-PPV:fullerene devices [21] versus theoretical LUMO levels (B3LYP/3-21G*) of acceptors for blend film dissolved in o-dichlorobenzene.*

Also in this case, there is a good linear correlation, although the V_{oc} was not obtained from fully optimized devices. Not all the acceptors were soluble in chlorobenzene; for this reason, two set of data are shown, related to devices based on blend dissolved in chlorobenzene or o-dichlorobenzene. Since the V_{oc} of polymer–fullerene solar cells is affected by the morphology of the device, it is not possible to compare these two set of data.

The same theoretical approach has been also applied to C_{70} derivatives with interesting results [23-25].

Conclusions

In this work the LUMO levels of different fullerenes derivatives are studied using DFT/ B3LYP/3-21G* method. Since the efficiency of the solar cells is directly proportional to the open circuit voltage that, for these devices, is related to the difference between the HOMO of the donor and the LUMO of the acceptor, we have proposed a theoretical approach to design fullerene derivatives with a proper LUMO level. The LUMO level was calculated for fullerene derivatives successfully used as electron acceptor for polymer solar cells. The obtained value is well correlated with the first reduction potential of these species and with the V_{oc} of the corresponding polymer-fullerene device. This approach can be useful to design new fullerene derivatives in order to maximize the open circuit voltage and improve the performance of such solar cells.

REFERENCES

[1] BRABEC C.J., DYAKONOV V., PARISI J., SARICIFTCI N. S. *"Organic photovoltaics: concepts and realization"* **60** (2003) Springer: Berlin.

[2] GÜNES, S., NEUGEBAUER, H., AND SARICIFTCI, N.S. *"Conjugated polymer-based organic solar cells"* Chemical Review **107** (2007) pp. 1324-1338.

[3] YU G., HEEGER A.J. *"Charge separation and photovoltaic conversion in polymer composites with internal donor/acceptor heterojunctions"* Joural of Applied Physics **78** (1995) pp. 4510-4515.

[4] YU G., GAO J., HUMMELEN J.C., WUDL F., HEEGER A.J. *"Polymer photovoltaic cells: Enhanced efficiencies via a network of internal donor-acceptor heterojunctions"* Science **270** (1995) 1789.

[5] MA W., YANG C., GONG X., LEE K., HEEGER A.J. *"Thermally stable, efficient polymer solar cells with nanoscale control of the interpenetrating network morphology"* Advanced Functional Materials **15** (2005) pp. 1617-1622.

[6] Solarmer Energy Inc.

[7] KREBS F.C., SPANGGAARD H. *"Significant improvement of polymer solar cell stability"* Chemistry of Materials **17** (2005) pp. 5235-5237.

[8] KREBS F.C. *"Fabrication and processing of polymer solar cells: A review of printing and coating techniques"* Solar Energy Materials and Solar Cells **93** (2009) pp. 394-412.

[9] WIENK M.M., KROON J.M., VERHEES W.J.H., KNOL J., HUMMELEN J.C., VAN HAL P.A., JANSSEN R.A.J. *"Efficient methano[70]fullerene/MDMO-PPV bulk heterojunction photovoltaic cells"* Angewandte Chemie, International Edition **42** (2003) pp. 3371-3375.

[10] YAMANARI T., TAIMA T., SAKAI J., SAITO, K. *"Highly efficient organic thin-film solar cells based on poly(3-hexylthiophene) and soluble C_{70} fullerene derivative."* Japanese Journal of Applied Physics **47** (2008) pp. 1230.

[11] SCHARBER M.C., MÜHLBACHER D., KOPPE M., DENK P., WALDAUF C., HEEGER A.J., BRABEC C.J. *"Design rules for donors in bulk-heterojunction solar cells - Towards 10 % energy-conversion efficiency"* Advanced Materials **18** (2006) pp. 789.

[12] MORVILLO P., BOBEICO E. *"Tuning the LUMO level of the acceptor to increase the open circuit voltage of polymer-fullerene solar cells: a quantum chemical study"* Solar Energy Materials and Solar Cells **92** (2008) pp 1192-1198.

[13] MORVILLO P., BOBEICO E. *"A fullerene silirane derivative to improve the open circuit voltage in a polymer-fullerene solar cell: a theoretical study"* Physica Status Solidi – Rapid Research Letters **2** (2008) pp. 260-262.

[14] GRANOVSKY A.A., PC GAMESS/ Firefly version 7.1.E, www http://classic.chem.msu.su/gran/gamess/index.html.

[15] SCHMIDT M.W., BALDRIDGE K.K., BOATZ J.A., ELBERT S.T., GORDON M.S., JENSEN J.H., KOSEKI S., MATSUNAGA N., NGUYEN K.A., SU S., WINDUS T.L., DUPUIS M., MONTGOMERY J.A. *"General atomic and molecular electronic structure system"* Journal of Computational Chemistry **14** (1993) 1347.

[16] DOBBS K.D., HEHRE W.J. *"Molecular orbital theory of the properties of inorganic and organometallic compounds 4. Extended basis sets for third-and fourth-row, main-group elements"* Journal of Computational Chemistry **7** (1986) pp. 359-378.

[17] PIETRO W.J., FRANCL M.M., HEHRE W.J., DE FREES D.J., POPLE J.A., BINKLEY J.S. *"Self-consistent molecular orbital methods. 24. Supplemented small split-valence basis sets for second-row elements"* Journal of the American Chemical Society **104** (1982) pp. 5039-5048.

[18] BECKE A.D. *"Density-functional thermochemistry. III. The role of exact exchange"* Journal of Chemical Physics 98 (1993) pp. 5648-5652.

[19] LEE C., YANG W., PARR R.G. *"Development of the colle-salvetti correlation-energy formula into a functional of the electron density"* Physical Review B **37** (1988) pp. 785-789.

[20] ZANDLER M.E., D'SOUZA F. *"The remarkable ability of B3LYP/3-21G(*) calculations to describe geometry, spectral and electrochemical properties of molecular and supramolecular porphyrin–fullerene conjugates"* Comptes Rendus Chimie **9** (2006) pp. 960-981.

[21] KOOISTRA F.B., KNOL J., KASTENBERG F., POPESCU L.M., VERHEES W.J.H., KROON J.M., HUMMELEN J.C. *"Increasing the Open Circuit Voltage of Bulk-Heterojunction Solar Cells by Raising the LUMO Level of the Acceptor"* Organic Letters **9** (2007) pp. 551-554.

[22] MORVILLO P., BOBEICO E. *"A theoretical approach to design electron acceptors for bulk heterojunction solar cells"* International Journal of Energy, Environment and Economics **17** (2009) pp. 211-228.

[23] MORVILLO P., BOBEICO E. *"C70 derivatives as electron acceptors in polymer-fullerene solar cells"* Proceedings of "24th European Photovoltaic Solar Energy Conference" (2009) pp. 617-620.

[24] MORVILLO P. *"Higher fullerenes as electron acceptors for polymer solar cells: a quantum chemical study"* Solar Energy Materials and Solar Cells **93** (2009) pp. 1827-1832.

[25] MORVILLO P., BOBEICO E. *"Bisadducts of C70 as electron acceptors for bulk heterojunction solar cells: a theoretical study"*, Fullerenes, Nanotubes and Carbon Nanostructures (2010), accepted.

NUMERICAL CFD ACTIVITIES PERFORMED BY AVIO

Fabio Paglia[1]

Avio S.p.A. Divisione Spazio - Progettazione Meccanica Lanciatori
Via Ariana Km 5.2 - 00034 Colleferro (Roma)

The CFD activities performed by AVIO on the GRID ENEA (CRESCO) have been focused on the following topics:

- TRANSCRITICAL COMBUSTION
- FLUID STRUCTURE INTERACTION
- AEROTHERMAL LOADS AT HIGH ANGLE OF ATTACK
- PRESSURE OSCILLATIONS
- HYPERSONIC FLOWS

TRANSCRITICAL COMBUSTION

The activities performed in the frame of the transcritical combustion topic has the scope to simulate the fluid dynamic field inside the combustion chamber of a liquid rocket engine (LM10-MIRA Demonstrator) in trans-critical conditions.

The LM10-MIRA Demonstrator is a liquid propellants rocket motor, designed for a set of test activities. The employed propellants are the methane and the liquid oxygen (LOX); these propellants are pushed in the combustion chamber through 60 coaxial injectors.

The methane is supplied to the injector head manifold, than through the casing orifices and to the cavity between the face plate and deflector. The methane flows in a radial direction from the injector head periphery to its centre, along the face plate, and ensures its cooling; then it comes to the cavity between the deflector and the casing, where it is distributed through the injectors (see *Figure 1*). The methane comes to the combustion chamber through the annular channels of the injectors. The Oxidizer comes to the oxygen cavity of the injector head, and further through the central channels of the injectors to the combustion chamber (see *Figure 1).*

The single injector element is designed in order to provide the mixing of the two propellants and at the same time to limit the pressure drop within itself.

CFD analyses have been performed on a computational grid of about 2 ML of computational cells.

---

[1] mailto: fabio.paglia@aviogroup.com

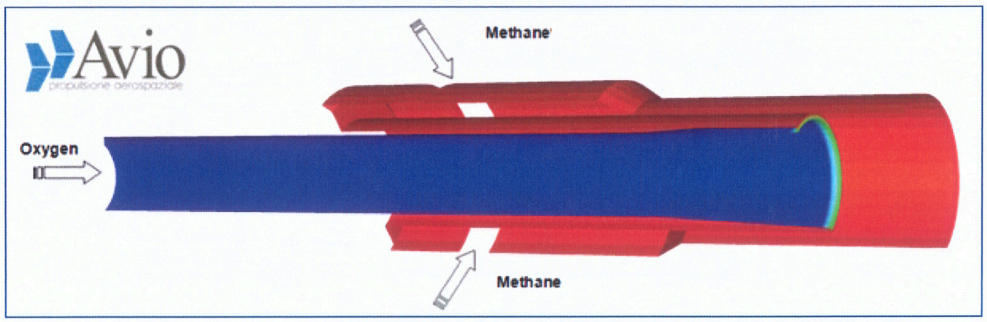


Figure 1 : *Injector geometrical configuration.*

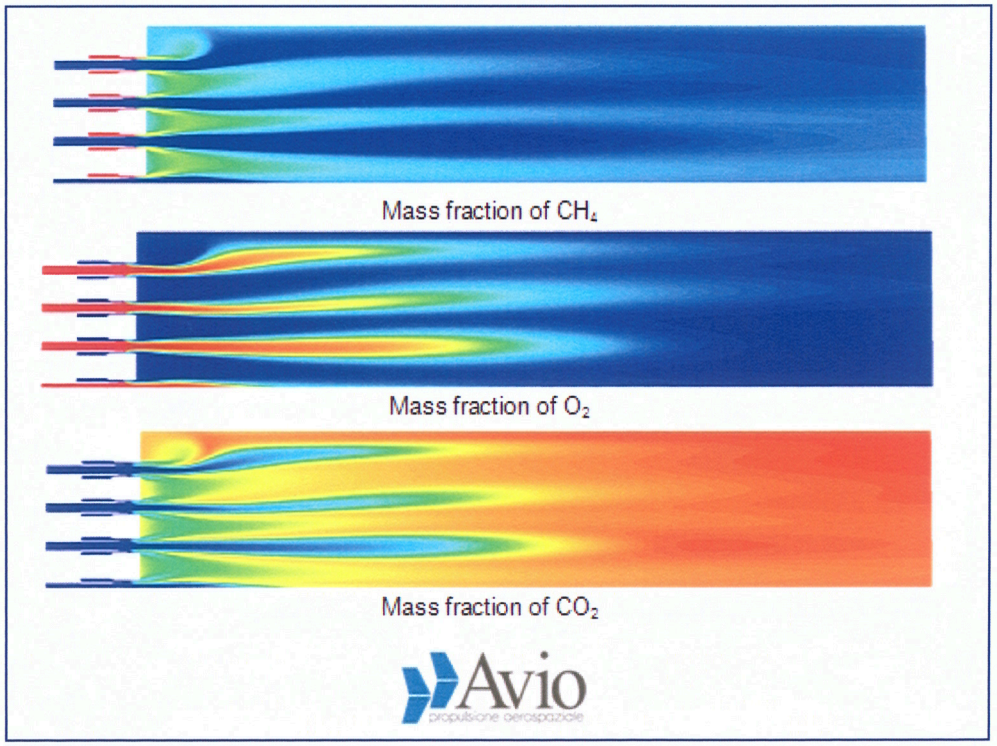


Figure 2 : *Mass fractions.*

FLUID STRUCTURE INTERACTION

In large segmented solid rocket motors, the thermal protection (PTF) are often placed on propellant frontal surfaces to prevent the combustion of the frontal face of propellant grain segments.

During combustion the high temperature flow that develops in the SRM hits the PTF, inducing erosion and determining its geometrical evolution.

Since heat flux distribution along the PTF is highly influenced by its deformation, a proper prediction of PTF erosion requires a coupled calculation of flow field in SRM and PTF deformation. The different positions assumed by the PTF, in fact, determine different qualitative heat flux distributions on the structure and, as consequence, different rates of regression of PTF.

In order to properly follow and predict the evolution in time of PTF an erosive-ablative module has be added to a Fluid Structure Interaction (FSI) methodology [1], [2].

Proposed method is based on the coupled use of three different computational modules for fluid dynamics, solid deformation and PTF ablation. In the numerical implementation of the procedure only the first two are coupled into an iterative loop, while the ablative module is used separately at the end of the FSI cycle (see *Figure 3*).
The iterative procedure of the FSI is explained by the following flowchart and it can be synthesized in five different steps:

CFD analyses have been performed on a computational grid of about 1.5 ML of computational cells.



Figure 3 : *FSI Flow chart and PTF deformations.*

1. The CFD solver calculates the fluid-dynamic field;
2. The pressure is transferred on the structural domain;
3. The CSD solver calculates the structural deformation field;
4. The deformations are transferred on the fluid-dynamic domain;
5. The loop restarts from the step 1 until the convergence is reached.

It must be noted that at each iteration, since the structure deformation changes the CFD boundary, the domain has to be re-meshed; then, a mesh generation module is also included immediately before the CFD solver.

AEROTHERMAL LOADS AT HIGH ANGLE OF ATTACK

The scope of such numerical activities is to evaluate the required inputs (i.e. pressure coefficients) to evaluate the aero-thermal loads at high angle of attack, in fact during the VEGA second stage powered phase a guided trajectory is foreseen and in particular this manoeuvre leads the launcher to high angles of attack.

Therefore in this activity the heat flux has been calculated at 10° , 15° and 20° of angle of attack employing an AVIO dedicated software fed by the pressure coefficient distribution calculated with the CFD analyses (*Figure 5*).

CFD analyses have been performed on a computational grid of about 3 ML of computational cells (see *Figure 4)*

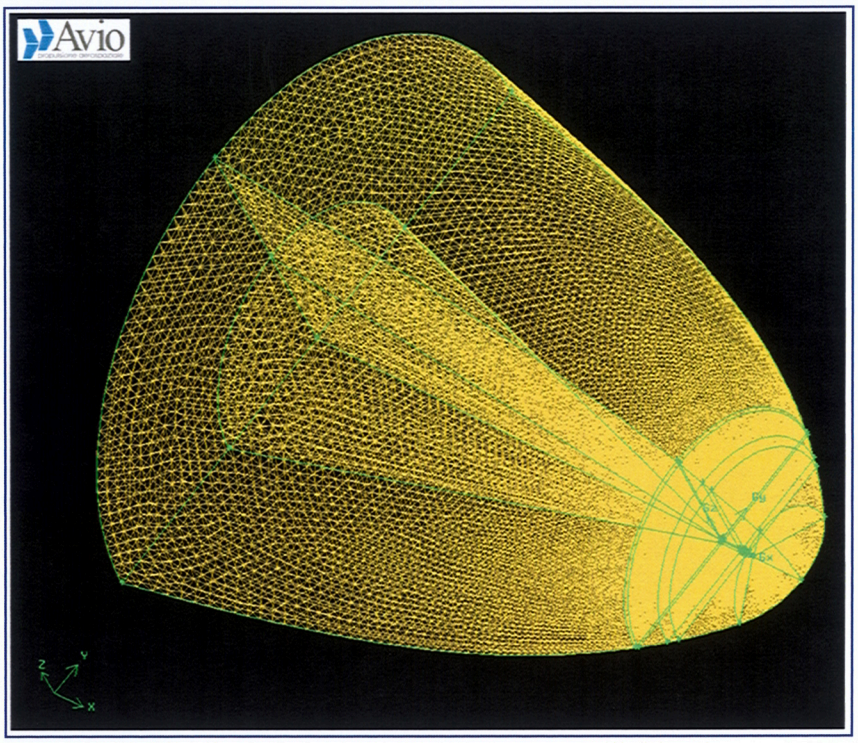


Figure 4 : *Mesh adopted*

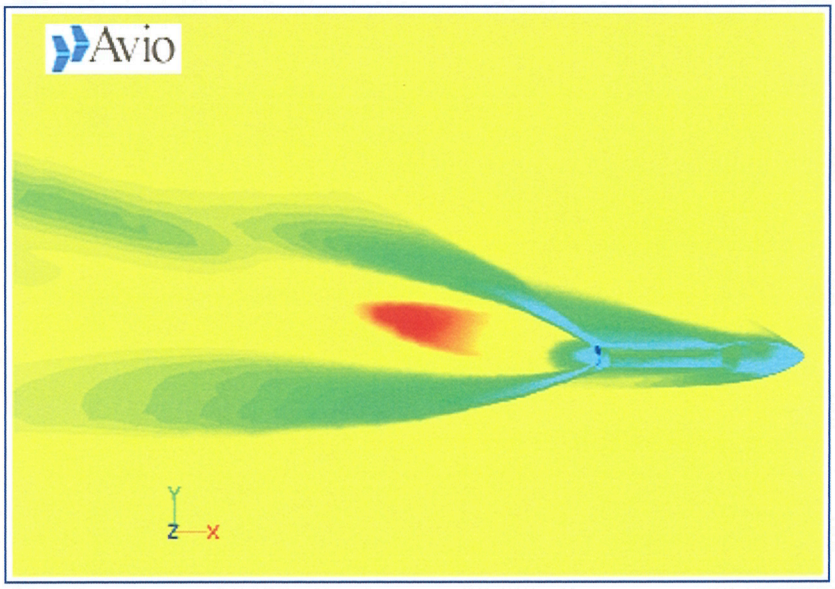


Figure 5 : *Mach number contour plot.*

PRESSURE OSCILLATIONS

The MPS P230 of Ariane 5, like the other large solid propellant motors, has to be designed as segmented for technological reasons. Such motor experience pressure oscillations leading to thrust fluctuations, at frequencies close to its first longitudinal modes. This behaviour induces some penalties on the overall performance of the launcher, so the study of pressure oscillations has attracted the interest of many researchers during last few decades.

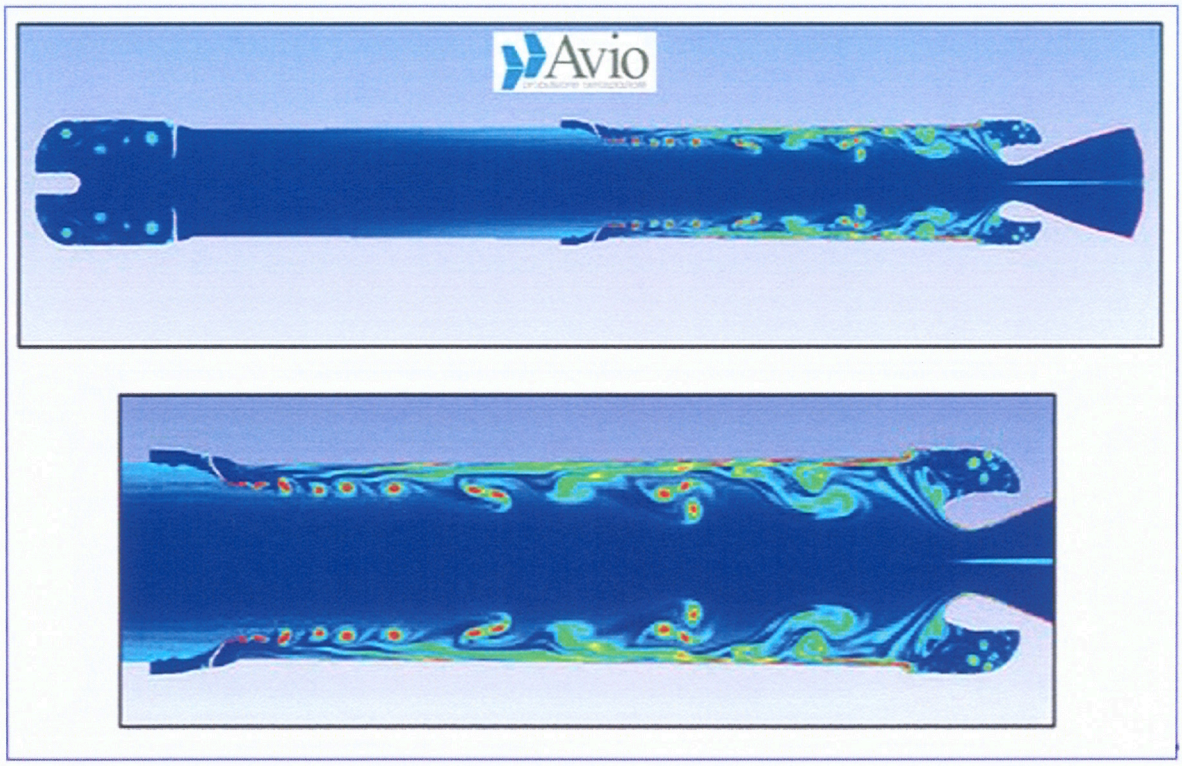


Figure 6 : Vorticity contour plot.

It is well known that numerical simulation of pressure oscillations in SRM is a very challenging task in terms of accurate modelling and capability of numerical simulation for highly intensive computing problems. Main reason of this complexity is the richness of the physical phenomenon under study (see *Figure 6*). This process can be summarized as follow:

1. A strong shear layer is created by the presence of the PTF in the core flow;
2. The shear layer becomes unstable and is the source of vortex structures;
3. Vortices are rolled – up and transported downstream;
4. Follows the impingement of the vortices on the surface of the nozzle head that generates an acoustic perturbation;
5. The acoustic perturbation propagates from the downstream source;
6. The acoustic perturbation reaches the PTF zone and works as trigger for the generation of a new vortex.

In order to keep into account all these different physical phenomena, with different length and time scales, numerical simulation requires a high level of accuracy and is clearly a key point for the success of the project.

An exacting work has been realized on MPS P230 configuration in order to acquire the capability of reproducing pressure oscillation behaviour during combustion.

CFD analyses have been performed on a computational grid of about 1.5 ML of computational cells [3].

HYPERSONIC FLOWS

During the re-entry phase due to the extremely high velocity (Mach = 20) of the vehicle it is in contact with high-energy fluid. Under these conditions, it is well known that chemical reactions may happen in the fluid and in particular, that oxygen starts dissociating creating ionized particles.

The physics of all this process is quite complex therefore mathematically model the problem is not a trivial task.

In the present analysis, to study the vehicle re-entry flows, a simplified approach based on the introduction in the CFD code of modified values for CP, Thermal conductivity and Viscosity has been adopted; such modified values have been derived from a NASA report and have been implemented as a polynomial function of the temperature .

CFD analyses have been performed on a computational grid of about 5 ML of computational cells.

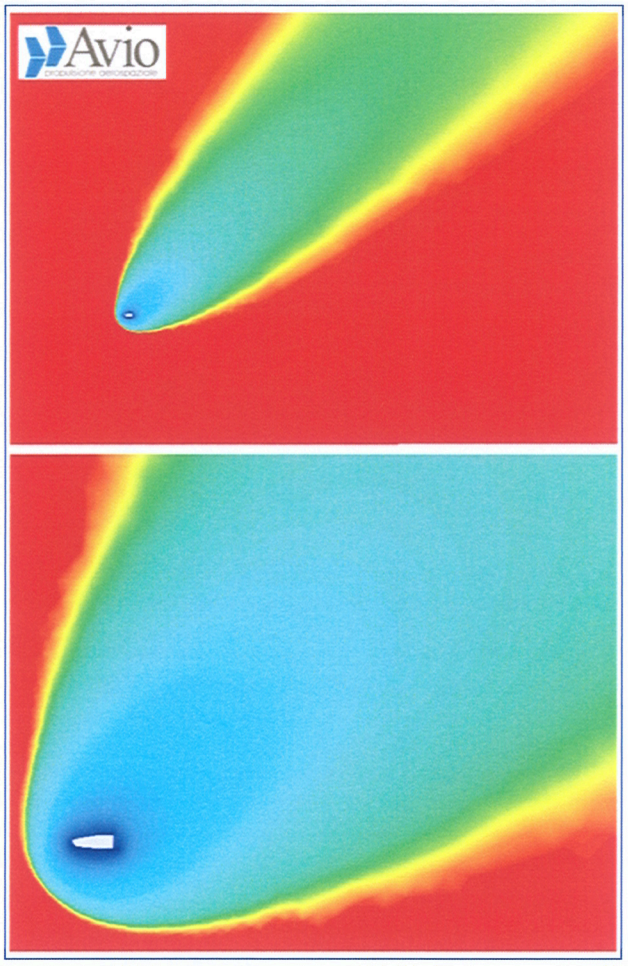


Figure 7 : *Mach number contour plot.*

REFERENCES

[1] BUCCHIGNAMI E., STELLA F., PAGLIA F. *"A partition method for the solution of a coupled liquid-structure interaction problem"* Applied Numerical Mathematics **51** (2004) pp. 463-475.

[2] STELLA F., GIANGI M., PAGLIA F., CASATA A, SIMONE D., GAUDENZI P. *"A numerical simulation of Fluid Structure Interaction in internal flows"* Numerical Heat Transfer, Part B **47** (2005) pp. 403-418.

[3] STELLA F., PAGLIA F. *"Pressure oscillation in solid rocket motors: numerical study"* Aerospace Science and Technology **15** (2011) pp. 53-59.

ARK3D – L'INFRASTRUTTURA ENEA-GRID PER IL 3D REMOTO

D.Abate, R.Ciavarella, G.Guarnieri, G.Furini, S.Migliori, S.Pierattini

ENEA-UTICT C.R. Bologna, Portici

{dante.abate, roberto.ciavarella, guido.guarnieri, graziano.furini, silvio.migliori, samuele.pierattini}@enea.it

Introduzione

La visualizzazione tridimensionale è divenuta oggi uno strumento fondamentale in molteplici ambiti di ricerca che vanno dalla produzione industriale, alla medicina e, non ultima, la fruizione, e valorizzazione dei Beni Culturali.

Realizzare un modello tridimensionale che descriva il proprio progetto significa poter presentare in modo efficiente ed efficace le proprie teorie; poter analizzare i risultati scientifici in maniera interattiva e condivisa; ricostruire lo *status quo ante* laddove i dati originali siano andati perduti; ed infine conservare la memoria digitale dell'oggetto e/o progetto che potrebbe andar perduta nella giungla dei diversi formati digitali.

La possibilità di poter accedere interattivamente ad oggetti 3D ad alta risoluzione, con tutte le funzionalità che offre un ambiente di visualizzazione tridimensionale è un aspetto fondamentale per arrivare alla piena condivisione e fruizione dei modelli virtuali. Essere in grado di visualizzare i risultati di una simulazione, senza dover spostare i dati, (sempre più voluminosi in termini di dimensioni), dal sito di supercalcolo, in cui sono stati generati, al mondo WEB o consentire a più ricercatori di condividere i risultati, sta diventando una necessità improcrastinabile.

Ad oggi distribuire sulla rete un modello 3D ad alta definizione significa cedere il modello stesso all'utente, che a sua volta deve disporre: di una connessione di rete veloce, di un hardware sufficientemente potente, e se disponibile, installare l'eventuale applicazione per la visualizzazione del modello scaricato.

Obiettivi

Lo studio condotto da ENEA, nell'ambito del progetto CRESCO [1], si basa sull'implementazione di un'architettura hardware e software che permette l'accesso remoto a modelli tridimensionali, ad alta risoluzione, multi disciplinari, e visualizzabili da WEB, installando una piccola applet dedicata sui computer client. La tecnologia sviluppata per la visualizzazione, libera l'utente finale dalla necessità di disporre di risorse hardware e software specifiche, nonché protegge i diritti d'autore collegati al modello 3D, di cui non verrà effettuato alcun download in locale. L'utente interagisce con l'applicazione utilizzando le risorse hardware e software remote (Remote 3D). L'infrastruttura è in grado di bilanciare il carico su cluster di macchine grafiche ed ottimizzare l'utilizzo di risorse (hardware e di rete) in base all'applicazione richiesta. Poiché l' hardware del computer dell'utente finale non viene caricato eccessivamente, le applicazioni disponibili potranno essere eseguite anche su notebook, netbook o addirittura su iPad! il solo requisito necessario è un collegamento affidabile a Internet (*Figura 1*).

Infrastruttura

Il progetto utilizza l'infrastrutture ICT di ENEA-GRID [2] ed in particolare il cluster grafico realizzato nell'ambito del progetto CRESCO, disponibile presso il centro ENEA di Portici (NA). L'infrastruttura HPC CRESCO è dotata di un cluster suddiviso in sezioni, distinte in base alle caratteristiche hardware (cpu, ram, schede grafiche), e complessivamente dispone di 28 Tflops di picco. La sezione di grafica è composta da 12 workstation o nodi per un totale di 68 core. Le schede grafiche utilizzate sono: NVidia Quadro FX. Il sistema di storage prevede l'accesso diretto sia al file system geografico AFS [3] che al GPFS (General Parallel File Systems - High speed storage 2 GByte e 160 TByte) [4] e dispone di un sistema centralizzato di backup. La connessione verso la rete internet avviene tramite la rete GARR (The Italian Academic & Research Network), con due linee da 1 Gbit.

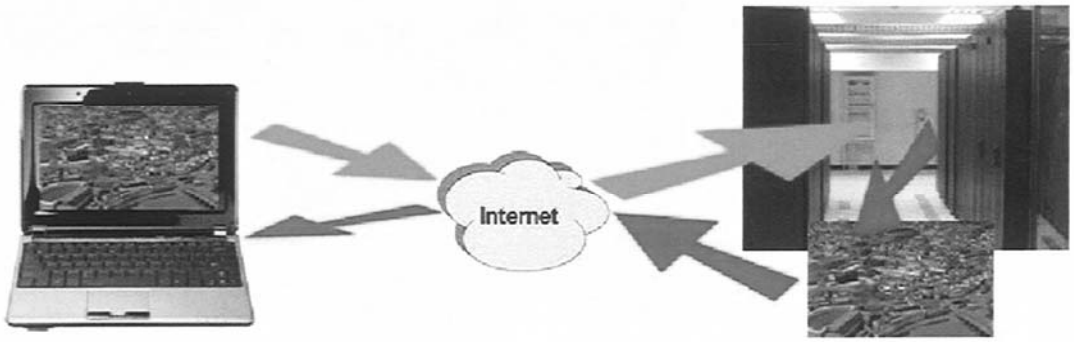


Figura 1: *3D Remoto*

Architettura

All'utente, che tramite browser si connette alla pagina WEB dedicata, viene richiesto di eseguire un applet java dedicata, necessaria per l'accesso all'applicazione grafica, che installa e configura in modo automatico il client. Non è essenziale da parte dell'utente conoscere parametri di configurazione o installare altre applicazioni. L'applicazione viene eseguita sulla macchina remota (server), utilizzando le risorse hardware compresa la scheda grafica per il rendering (Remote Rendering). La visualizzazione può avvenire tramite qualsiasi applicazione grafica che utilizzi le librerie OpenGL [5] su Sistema Operativo Linux, sia proprietarie che Open Source. Attualmente il cluster grafico dedicato al progetto è composto da workstation con 4 processori AMD dual-core, da 16GB di RAM e scheda grafica NVidia Quadro FX (*Figura 2*).

Tra client e server il flusso di dati è esclusivamente di tipo streaming, di immagini compresse, generate remotamente dall'applicazione di visualizzazione, oltre alle funzioni di interazione della tastiera e del mouse. Le operazioni demandate al client sono limitate alla decompressione dell'immagine ed alla gestione degli input da tastiera e mouse. Grazie a questa tecnologia, per accedere alla piattaforma è possibile utilizzare strumenti hardware come netbook o palmari nonostante la limitata banda di connessione. E' possibile inoltre trasmettere anche l'audio generato dall'applicazione remota.

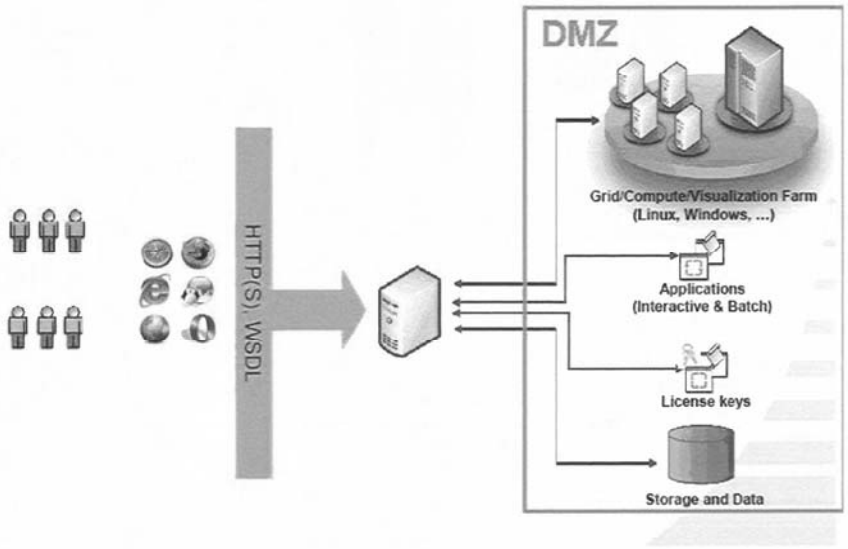


Figura 2: *Immagine schematica dell'architettura*

Esempi

E' possibile accedere ai progetti e alle applicazioni pubblicati sull'infrastruttura ARK3D, tramite il sito web: https://www.ark3d.enea.it - area "Projects". 3DWS-Project [6] è un repository, indicizzato come un Data Base di modelli 3D caricati da utenti registrati e visualizzabili in alta risoluzione (*Figura 3*).

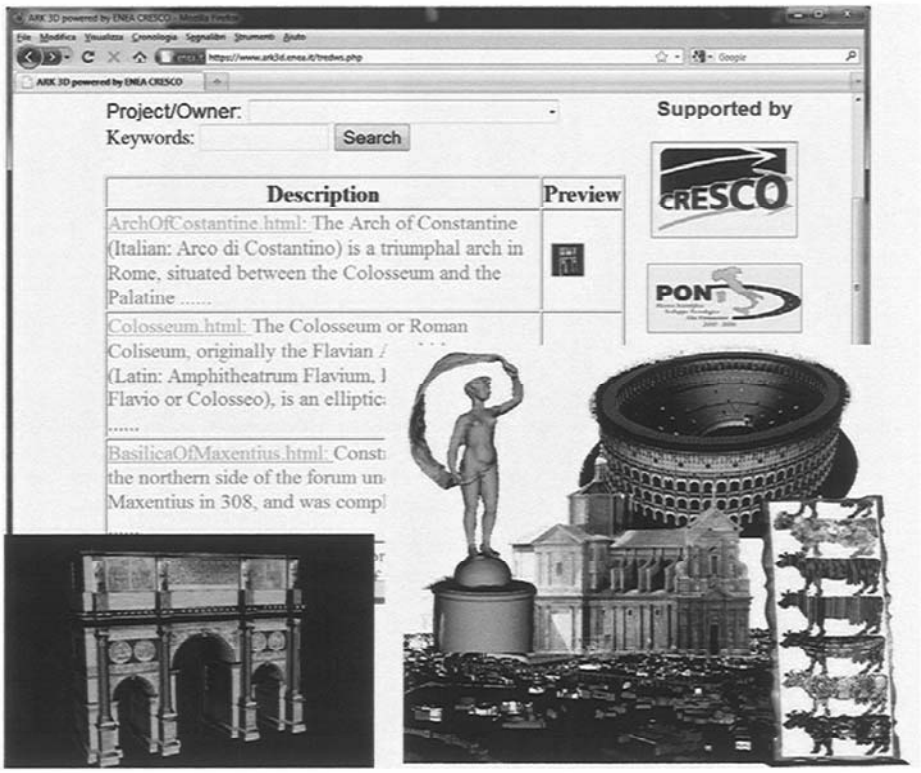


Figura 3: *Ricerca nel DataBase dei modelli 3D e visualizzazione — 3DWS Projects*

Clima è un esempio di accesso ai dati, in questo caso di simulazioni di modelli climatici, risultato di svariate ore di calcolo sull'infrastruttura ENEA-GRID che generano output di grosse dimensioni e visualizzabili tramite applicazioni personalizzate (*Figura 4*).

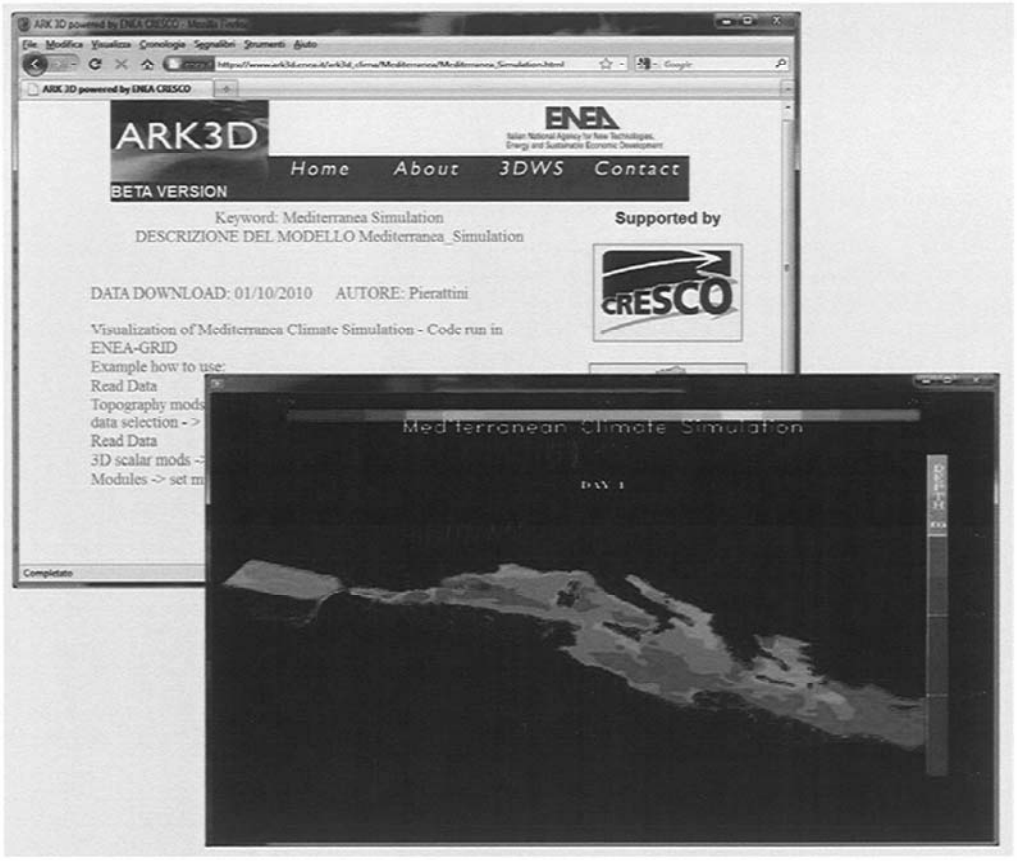


Figura 4: *Esecuzione remota di applicazione 3D*

RIFERIMENTI

[1] Progetto CRESCO – http://www.cresco.enea.it

[2] ENEA-GRID - http://www.eneagrid.enea.it/

[3] OpenAFS - http://www.openafs.org/

[4] GPFS (IBM) - http://www-03.ibm.com/systems/software/gpfs/index.html

[5] OpenGL - http://www.opengl.org/

[6] 3DWS - "3D Web Service Project" - https://www.ark3d.enea.it/tredws.php

IMPROVING FREE ENERGY PERTURBATION METHODS THROUGH MULTIPLE STATE SAMPLING: THE HISTIDINE TAUTOMERISM CASE

Simone Marsili and Piero Procacci[1]

Dipartimento di Chimica, Università di Firenze
Via della Lastruccia 3, I-50019 Sesto Fiorentino, Italy

Centro per lo Studio delle Dinamiche Complesse(CSDC)
Via Sansone 1, I-50019 Sesto Fiorentino, Italy

Abstract

We propose a new strategy for computing free energy difference for alchemical reactions based on a combination of free energy perturbation techniques and Hamiltonian replica exchange. The technique has been tested for the determination of the relative thermodynamic stability in normal condition of the tautomers of the histidine side chain.

Introduction

Free energy perturbation (FEP) methods in statistical thermodynamics are based on the Zwanzig equation [1] for the calculation of the free energy difference between two state A and B characterized by the Hamiltonians H_A and H_B :

$$e^{-(F_B-F_A)/k_BT} = \left\langle e^{-(H_B-H_A)/k_BT} \right\rangle_A \qquad (1)$$

where $FB = -kbT \ln Z_B$, $FA = kbT \ln Z_A$ are the free energies of state A and B and $< .. >_A$ denotes a canonical average of the system in the state A. From a computational standpoint, FEP is normally applied to systems that differ in the potential energy (e.g. alchemical reactions)[2] and is implemented by simulating the system in the state A and averaging the exponential of the potential energy difference V_B - V_A. The Zwanzig equation Eq. (1) can be viewed as a special case of the Jarzynsky identity [3] $e^{-\beta\Delta F} = < e^{-\beta W} >$, with the non equilibrium transformation from state A and B involving the work W done at infinite speed. Calculation of the free energy difference $\Delta F=F_B-F_A$ via the Jarzynsky identity is notoriously affected by a large statistical error [4, 5, 6, 7], mainly because the exponential average depends crucially on a small fraction of realizations that transiently violate the second law of thermodynamics (*i.e.*, realizations that yield a negative dissipated work $W-\Delta F$). This is of course especially true for FEP that involves *instantaneous* highly dissipative transformations whereby the probability of obtaining an energy

[1]To whom correspondence should be addressed; E-mail:procacci@unifi.it

close to the free energy ΔF - hence yielding a significant contribution to the exponential average of Eq. (1) - is vanishingly small.

The problem of the pathological error in the Jarzynsky exponential averages can be circumvented by resorting to non equilibrium *bidirectional* transformations[8, 9, 10, 11, 12] related by a time reversal protocol and obeying the Crooks theorem. The FEP variant of the Crooks theorem reads

$$\frac{P(\Delta E)}{\tilde{P}(-\Delta E)} = e^{(\Delta E - \Delta F)/k_B T} \qquad (2)$$

where $P(\Delta E)$ is the probability of obtaining the energy difference $V_B - V_A$ for states canonically sampled according to the Hamiltonian H_A, and $P(-\Delta E)$ is the probability of obtaining the same energy difference for states canonically sampled with Hamiltonian H_B. The two distributions $P(\Delta E)$ and $P(-\Delta E)$ cross each other at $\Delta E=\Delta F$. Such crossing point can be effectively recovered using using the so-called Bennett acceptance ratio (BAR).[13] Again the error in the BAR is large when the overlap between the two distributions is small and this problem is particularly acute for bidirectional FEP. For example, when evaluating the free energy difference between two solvated isomers involving polar moieties, the difference V_B-V_A is usually very large since the surrounding solvent is canonically distributed for the sampled state of only one isomer (whether A or B). If the sampled thermodynamic state is A then V_B-V_A is large and positive and if sampled state is B then V_B-V_A is large and negative and the overlap between the distributions $P(\Delta E)$ and $\tilde{P}(-\Delta E)$ is practically zero. The difference becomes smaller (and the overlap larger) with increasing temperature since the solvation shell of the solute becomes more and more structureless. On parallel platforms, one can exploits this fact by combining the replica exchange method (REM)[14, 15, 16] with the bidirectional FEP. As the temperature is raised along the progression of replicas, the structure of the solvent surrounding the solute becomes gradually insensitive to the isomeric state. In these conditions $\langle V_B\text{-}V_A\rangle_A$ decreases and $\langle V_B\text{-}V_A\rangle_B$ becomes less negative: the two distributions get closer, hence increasing their overlap. These overlapping distributions for the hot system can be projected onto the target temperature by using recently developed reweighting techniques such as the multiple Bennett acceptance ratio (MBAR).[17] thereby filling the gap between the forward and reverse distribution at normal temperature. In order to keep the needed number of replicas at minimum, one can use the Hamiltonian replica exchange variant (HREM) by selectively scaling ("heating") the solute-solvent interactions that are directly involved in the isomeric equilibrium.

In this paper we illustrate such a synergic approach involving bidirectional FEP, HREM in a paradigmatic example of alchemical transformations involving isomers, namely the determination of the solvation free energy difference between the tautomer δ and ε of histidine side chain.

Methods

Replica exchange simulation is ideal for parallel platforms as the amount of communications among processors is kept at minimum (basically only the scaling factors are exchanged among the processes). The simulated system contains the compound Ace-His-Nme solvated in 504 TIP3[18] water molecules. The force field for the Ace-His-Nme is based on the Amber03 (ffamber03)[19] parametrization. The two REM simulations referring to the δ or ε His tautomers

were performed using the program ORAC[20, 21] for 12 ns in the NPT ensemble at T = 300 K and P = 1 atm. Replica exhanges were attempted every 0.360 ps. Each *k* replica differs in the electrostatic interaction potential between the subset *S* of atoms that characterize the tautomers and all the rest of the system, namely

$$V_K = V_0 + \sum_{i \in S} \sum_{j \notin S} (c_k - 1) \frac{q_i q_j}{r_{ij}} \qquad (3)$$

where V_0 is the unperturbed potential of the target system, q_i, q_j are partial atomic charge and c_k is the replica scaling factor. In the case of the δ tautomer, the three atoms $N\delta_\delta$, Hδ and $N\varepsilon_\varepsilon$ form the subset *S*, while for the ε tautomer the hydrogen H_ε takes the place of H_δ (see *Figure* 1). With this minimal choice for the replica potential, one can afford a large scaling range and eight replicas of the peptide Ace-His-Nme were sufficient for obtaining a mean replica exchange rate beyond 50%. The c_k factors are 1.00000, 0.76260, 0.58156, 0.44350, 0.33822, 0.25793, 0.19669, 0.15000. The first replica corresponds to the unperturbed solvated peptide. The other replicas progressively lose their capability to form hydrogen bonds with the surrounding, such that the solvation of the two tautomers in the 8-*th* replica is similar, and the simulations with the highest replica index sample states that are common to the tautomers, bridging the solvated states typical to the δ and the ε tautomers.

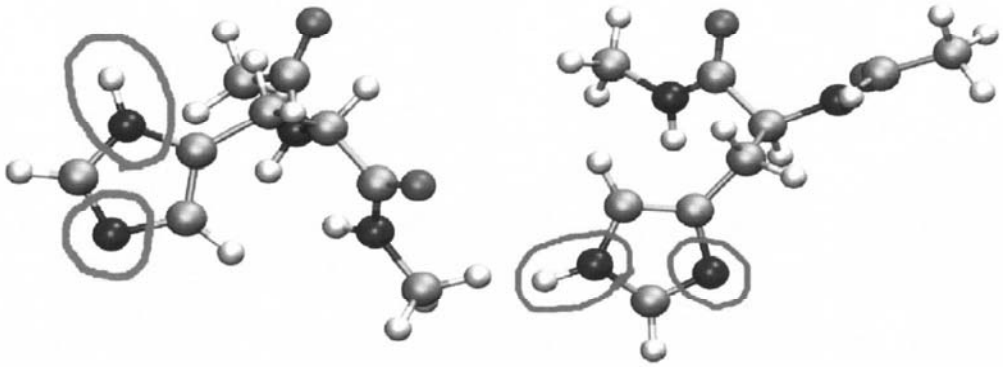


***Figure 1**: Ace-His δ-Nme (left) and Ace-His ε-Nme (right) tautomers. The atoms in the circle are the subset of atoms whose electrostatic interactions with the rest of the system are scaled in the HREM simulation.*

Results

Combining FEP with HREM/MBAR, we shall determine the *solvation* free energy difference due to the tautomeric exchange in the his residue, represented by the reaction

$$(His_\delta)_{vacuo} + (His_\varepsilon)_{H_2O} \rightarrow (His_\varepsilon)_{vacuo} + (His_\delta)_{H_2O} \qquad (4)$$

To this end, for each microstate *x* sampled in the 8 replicas simulation of the δ tautomer, the energy difference

$$\Delta E^{solv}_{\delta \rightarrow \varepsilon}(x) = E_\varepsilon(x) - E_\delta(x)$$

is computed, where $E_\delta(x)$ is the energy of the microstate and $E_\varepsilon(x)$ is the energy of the corresponding configuration of the ε tautomer (obtained removing the δ hydrogen and inserting the ε hydrogen on the imidazole ring). According to Eq. (), this energy is corrected subtracting the energy difference for the corresponding transition *in vacuo*, $\Delta E_{\delta->\varepsilon}$, obtaining

$$\Delta E_{\delta\to\varepsilon} = \Delta E^{solv}_{\delta\to\varepsilon}(x) - \Delta E^{vacuo}_{\delta\to\varepsilon}(x).$$

The same procedure is followed for the δ–>ε transition. Finally, 2×k histograms, namely

$$P^{(k)}_{\delta\to\varepsilon}(\Delta E) = \left\langle \delta\left(\Delta E - \Delta E^{solv}_{\delta\to\varepsilon}(x) + \Delta E^{vacuo}_{\delta\to\varepsilon}(x)\right)\right\rangle^{(\delta)}_{k}$$
$$P^{(k)}_{\varepsilon\to\delta}(\Delta E) = \left\langle \delta\left(\Delta E - \Delta E^{solv}_{\varepsilon\to\delta}(x) + \Delta E^{vacuo}_{\varepsilon\to\delta}(x)\right)\right\rangle^{(\varepsilon)}_{k} \qquad (5)$$
$$k = 1,2,..,8$$

where the notation, $<...>_k^\delta$ e.g., means that the canonical average is acquired using the configurations canonically sampled according k-th scaled potential of Eq. (3) for the tautomer in the δ state. For k=1 (i.e. for the target replica), such histograms intersects at the value of the free energy difference ΔG we are looking for. In *Figure 2* (left panel), the two distributions obtained for the target system (c_1=*1*) are shown.

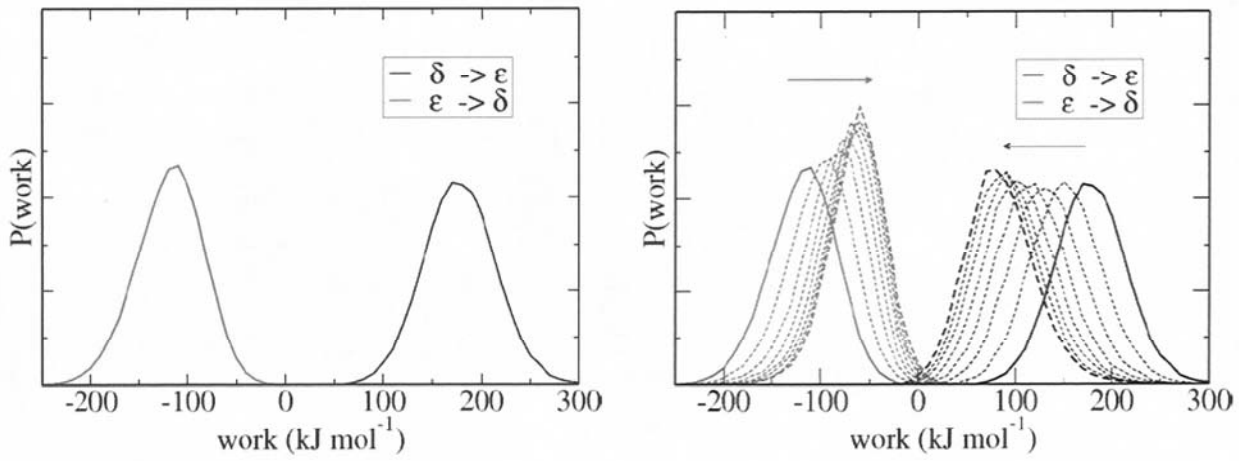


***Figure 2**: Probability distribution of energy difference for the tautomeric exchange for states sampled from the target (unperturbed) replica (left panel). In the right panel we show such distributions for all the eight replicas. Solid thick lines refer to the distribution of the target (c_1=1) replica and dashed thick lines to the distributions of the hottest replica (c_8=0.15).*

The tails of the distributions are separated by almost 100 kJ mol and this separation translates in a huge error in the ΔG estimate if one straightforwardly applies the BAR formula to Eq. (2). In the right panel of *Figure 2* the two distributions are shown for all the eight replicas; the distributions δ–>ε and ε–>δ corresponding to the hottest and unscaled replicas are drawn with a solid line. As stated in the introduction, the potential scaling results in closer distributions, since both the potential hydrogen bond acceptor (the N atom) and the hydrogen bond donor (the N-H group) of the imidazole ring lose their capability to interact favourably with their

environment, and the work spent in switching an hydrogen between the N_δ and the N_ε decreases. However, statistics collected from replicas with perturbed potential V_k has to be reweighted for the proper sampling with the original potential V_0. Data from multiple equilibrium simulations at different thermodynamic conditions can be optimally analyzed via the Multiple Bennett Acceptance Ratio method (MBAR)[17], computing for every configuration x sampled in the generalized ensemble of the simulation a weight[22] $w(x)$ proper to the unperturbed system.[17] Then, new histograms

$$P_{\delta\to\varepsilon}(\Delta E)=\frac{\int w(x)\delta\left(\Delta E-\Delta E^{solv}_{\delta\to\varepsilon}(x)+\Delta E^{vacuo}_{\delta\to\varepsilon}(x)\right)_\delta dx}{\int w(x)dx}$$
$$P_{\varepsilon\to\delta}(\Delta E)=\frac{\int w(x)\delta\left(\Delta E-\Delta E^{solv}_{\varepsilon\to\delta}(x)+\Delta E^{vacuo}_{\varepsilon\to\delta}(x)\right)_\varepsilon dx}{\int w(x)dx} \qquad (6)$$

can be accumulated using the statistics from all the eight replicas, as shown in *Figure 3*. Comparison with *Figure 2*, where the histograms obtained from the unperturbed replica alone are shown, points out the gain in sampling accuracy in the tails of the two distributions, as emphasized from the logarithmic plot in the inset. Using the histograms accumulated from all configurations in the generalized ensemble, the free energy difference ΔG (corresponding to the ΔE for which the histograms have the same value) can now be accurately determined through the Bennett formula, modified inserting the weights computed through the MBAR algorithm. The resulting value is ΔG =13.3 kJ mol, i.e. the most stable His side chain tautomer, as long as the solvation free energy of Ace-His-Nme compound is concerned, is the δ form.

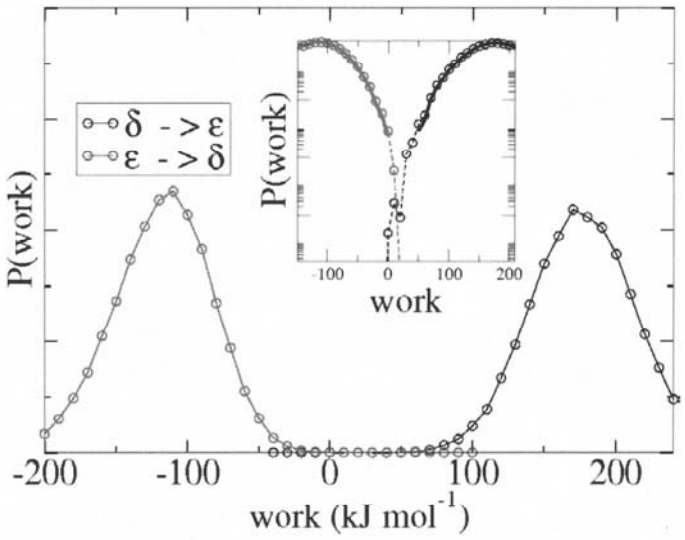


Figure 3: *Probability distribution of energy difference for the tautomeric exchange for states sampled from the multiple state system using MBAR. In the inset, in a logarithmic scale, it is shown the reconstruction of the two distributions at small overlap (thin solid lines plus circles) using the hot replicas statistics projected at the target replica (c_1=1) with MBAR. The distributions obtained with state sampled form the target replica alone is shown in the inset as thick solid lines.*

Conclusion

We have described a computational strategy that can be effectively used for determining free energy difference between two isomeric compounds *A* and *B* in condensed phase, a problem that is of paramount importance in drug design. Such a technology combines on parallel platforms bidirectional free energy perturbation techniques based on the Crooks theorem with the Hamiltonian replica exchange method. By selectively applying the Hamiltonian scaling to the interactions between the atoms involved in the isomerization only and the remainder of the system, one can keep the amount of processes (replicas) at minimum. With such scaling protocol, two HREM simulations are performed with the solute in the two isomeric states described by the Hamiltonians H_A and H_B. From these simulations, one can straightforwardly builds the two time reversal related histograms for the energy corresponding to the isomeric exchange starting either from state A or B. The tails of the distributions can be accurately reconstructed using all the configurational statistics of the generalized ensemble by reweighting the histograms computed for hot replicas for the proper sampling at normal conditions using the multiple Bennett acceptance ratio recently proposed.[17] The methodology has been practically demonstrated for the determination of the solvation free energy difference between the tautomer δ and ε of the histidine side chain.

Acknowledgements

This work was supported by Regione Toscana. We wish to thank Samuele Pierattini of ENEA for his constant and competent technical support on the CRESCO facility.

REFERENCES

[1] R. W. Zwanzig. High-temperature equation of state by a perturbation method. i. nonpolar gases. *J. Chem. Phys.*, 22:1420–1426, 1954.

[2] T. Simonson, G. Archontis, and M. Karplus. Free energy simulations come of age: protein-ligand recognition. *Acc. Chem. Res.*, 35:430–437, 2002.

[3] C. Jarzynski. Nonequilibrium equality for free energy differences. *Phys. Rev. Lett.*, 78:2690, 1997.

[4] H. Oberhofer and C. Dellago abd P. L. Geissler. Biased sampling of non-equilibrium trajectories: Can fast switching simulations outperform conventional free energy calculation methods? *J. Phys. Chem. B*, 109:6902, 2005.

[5] S. Park and K. Schulten. Calculating potentials of mean force from steered molecular dynamics simulations. *J. Chem. Phys.*, 120:5946–5961, 2004.

[6] G. Hummer. Fast-growth thermodynamic integration: error and efficiency analysis. *J. Chem. Phys.*, 114:7330–7337, 2001.

[7] M. R. Shirts and V. S. Pande. Comparison of efficiency and bias of free energies computed by exponential averaging, the Bennett acceptance ratio, and thermodynamic integration. *J. Chem. Phys.*, 122:144107, 2005.

[8] G. E. Crooks. Path-ensemble averages in systems driven far from equilibrium. *Phys. Rev. E*, 61:2361–2366, 2000.

[9] P. Procacci, S. Marsili, A. Barducci, G. F. Signorini, and R. Chelli. Crooks equation for steered molecular dynamics using a Nose'-Hoover thermostat. *J. Chem. Phys.*, 125:164101, 2006.

[10] Collin D., Ritort F., Jarzynski C., Smith S. B., Tinoco I. Jr, and Bustamante C. Verification of the Crooks fluctuation theorem and recovery of rna folding free energies. *Nature*, 437:231–234, 2005.

[11] S. Marsili and P. Procacci. Free energy reconstruction in bidirectional force spectroscopy experiments: the effect of the device stiffness. *J. Phys. Chem. B*, 114:2509, 2010.

[12] P. Procacci and S. Marsili. Energy dissipation asymmetry in the non equilibrium folding/unfolding of the single molecule alanine dacapeptide. *Chem. Phys*, 375:8–15, 2010.

[13] M. R. Shirts, E. Bair, G. Hooker, and V. S. Pande. Equilibrium free energies from nonequilibrium measurements using maximum likelihood methods. *Phys. Rev. Lett.*, 91:140601, 2003.

[14] R. H. Swendsen and J. S. Wang. Replica monte carlo simulation of spin-glasses. *Phys. Rev. Lett.*, 57:2607–2609, 1986.

[15] E. Marinari and G. Parisi. Simulated tempering: a new monte carlo scheme. *Europhys. Lett.*, 19:451, 1992.

[16] K. Hukushima and K. Nemoto. Exchange monte carlo method and application to spin glass simulations. *J. Phys. Soc. Jpn.*, 65:1604–1608, 1996.

[17] M. R. Shirts and J. D. Chodera. Statistically optimal analysis of multiple equilibrium simulations. *J. Chem. Phys.*, 129:124105, 2008.

[18] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.*, 79:926, 1983.

[19] Y. Duan, C. Wu, S. Chowdhury, M. C. Lee, G. Xiong, W. Zhang, R. Yang, P. Cieplak, R. Luo, T. Lee, J. Caldwell, J. Wang, and P. Kollman. A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *J. Comput. Chem.*, 24:1999–2012, 2003.

[20] S. Marsili, G. F. Signorini, R. Chelli, M. Marchi, and P. Procacci. Orac: A molecular dynamics simulation program to explore free energy surfaces in biomolecular systems at the atomistic level. *J. Comp. Chem.*, 31:1106–11161, 2010.

[21] P. Procacci, T. Darden, E. Paci, and M. Marchi. Orac: a molecular dynamics program to simulate complex molecular systems with realistic electrostatic interactions. *J. Comp. Chemistry*, 18:1848, 1997.

[22] A detailed derivation of the weight $w(x)$ for the target replica can be found in the orac manual at http://www.chim.unifi.it/orac/orac-manual.pdf, pages 45-46.

AB-INITIO CALCULATIONS OF EXCITED STATE PROPERTIES OF COMPLEX SYSTEMS

Margherita Marsili[1,2], Paola Gori[1,3], and Olivia Pulci[1,2,3]

[1] European Theoretical Spectroscopy Facility

[2]Dipartimento di Fisica Università di Roma 'Tor Vergata'

Via della Ricerca Scientifica 1 - 00133 Roma

[3]CNR-ISM, Via del Fosso del Cavaliere Roma

Introduction

In this report we review briefly the results that have been obtained in the last year also thanks to the use of the ENEA CRESCO facility. All the works concern the calculation of the electronic and optical properties of materials employing ab-initio theoretical methods based on density-functional theory (DFT) and/or many-body perturbation theory (MBPT).

The report is organized as follows: an overview of the theoretical framework is given, then we present the results of the different studies concerning (i) the optical and electronic properties hydrogenated diamond surfaces and graphane; (ii) the influence of steps on the optical properties of C(100):H; (iii) the influence of different time-dependent DFT (TDDFT) long-range kernel in the calculation of the optical properties of Si(111)2x1; (iv) the optical spectra of bulk ZnO.

While the calculations running on the ENEA CRESCO facility have been mainly carried out by the authors of this report, the complete studies to which we will refer are due to the work of many collaborators, which appear as coauthors in the publications.

Codes used:

ESPRESSO (www.quantum-espresso.org), EXC, YAMBO, GW, DP (www.etsf.eu)

Methods (DFT and Many Body techniques in a nutshell)

In this paragraph the theoretical approaches used are reviewed: DFT for ground state properties, GW for band structure calculations (charged excitations), and the Bethe–Salpeter approach and/or the Time Dependent DFT for optical spectra (neutral excitations).

Density Functional Theory (DFT) treats the case of an external, time independent potential, and is based on the seminal paper of Hohenberg and Kohn of 1964 [1]. For a review on DFT see for example [2]. In their paper, Hohenberg and Kohn proved that all the ground state properties of an interacting electronic system, including, in principle, the many-body wave function, could be expressed as unique functionals of the electronic density alone. In particular this assertion is valid also for the total energy E of the system. The energy functional finds its minimum at the ground state density n_{GS} , for which $E_{GS} = E[n_{GS}]$. This theorem leads, in principle, to a straightforward method for computing ground state properties, but the explicit form of the functional in terms of the density is not known and such an explicit functional may not exist.

Ground state properties are instead computed through the application of the so-called Kohn-Sham scheme [3], mapping the study of the (complicated) interacting system into the study of a (simpler) non interacting fictitious system that has, by construction, the same ground state density of the interacting system. The single-particle eigenvalues obtained within the solution of the Kohn-Sham scheme are often interpreted as one electron excitation energies corresponding to the excitation spectra of the system upon removal or addition of an electron, and DFT is in this way used to calculate band structures. The qualitative agreement with experiments is often remarkable, but not quantitative: the electronic gaps of semiconductors are, as a matter of fact, always systematically underestimated within DFT.

Underestimation of the electronic gaps in DFT calculations finds its origin not in a deficiency of the theory, but in the incorrect use we make of the Kohn–Sham equations. In fact, even if the Kohn–Sham equations have the form of single-particle Schrödinger-like equations, their eigenvalues are not the excited state energies. DFT is an exact theory for ground state properties, but there is no strict theoretical justification to use it to obtain excitation energies.
In the determination of the electronic band structure, we are dealing with energy differences between the system with N-electrons and the system with N – 1 (direct photoemission spectroscopy) or N + 1 (inverse photoemission spectroscopy) electrons. It is hence natural to change over to the Green's function formalism [4], where the electron addition and removal energies are the poles of the single-particle Green's function G. The single-particle Green's function obeys the Dyson equation $G=G_0+G_0\Sigma G$, which links it to the non-interacting Green's function G_0 and the electronic self-energy Σ. A closed set of equations, the so called Hedin's equations [5], is available which together with the Dyson equation for G defines implicitly the self-energy Σ. Approximation for Σ, may be found within this framework: all the work presented in this report makes use of the GW approximation [6], in which Σ is expressed as the product of the single-particle Green's function and the screened Coulomb interaction W. Band structures computed within the single-particle approximation are typically in excellent agreement with experiments (for a review on this topic see for example [7]).

In order to describe optical spectra, we have to take into account excited states of a completely different nature with respect to the ones involved in photoemission spectroscopy.
In the latter the final state of the system is charged since one electron has been removed or added to it; in the first, instead, the system exchanging energy with the probe, is left in what is called a neutral excited state. The physical quantity which is directly connected to these spectra is the macroscopic dielectric function ε_M, linked to the *inverse* microscopic dielectric function through the so called macroscopic average, as shown in [8]. The different levels of sophistication on the determination of ε_M correspond to the different physical effects that described. For instance avoiding the macroscopic average step and computing directly ε_M from the microscopic dielectric function, is computationally favorable, but does not include the effects of the system inhomogeneity (i.e. the so called local-field effects are neglected [9]). Going beyond a single-(quasi)particle description, the electron-hole interaction may be included in the determination of ε_M mainly through two approaches: the solution of the Bethe-Salpeter equation (BSE) in the framework of many-body perturbation theory, and time-dependent DFT. The two methods are carefully compared and reviewed in [10].

Results:

Physics of Carbon surfaces

With the aid of *ab initio*, parameter free calculations based on density-functional and many-body perturbation theory, the electronic band structure and electron affinity of diamond surfaces have

been investigated. We have focused on clean, ideal (001) and (111) surfaces and on the effect of hydrogen adsorption. Also single sheets of graphane, that is graphene functionalized upon hydrogen, were investigated. At full H-coverage nearly free electron states (NFESs) appear near the conduction band minimum in all the systems under study. An example of NFES is shown in *Figure 1* for the case of the (001) hydrogenated diamond surface. The electron affinity is strongly reduced becoming negative for the hydrogenated diamond surfaces, and almost zero in graphane [11].

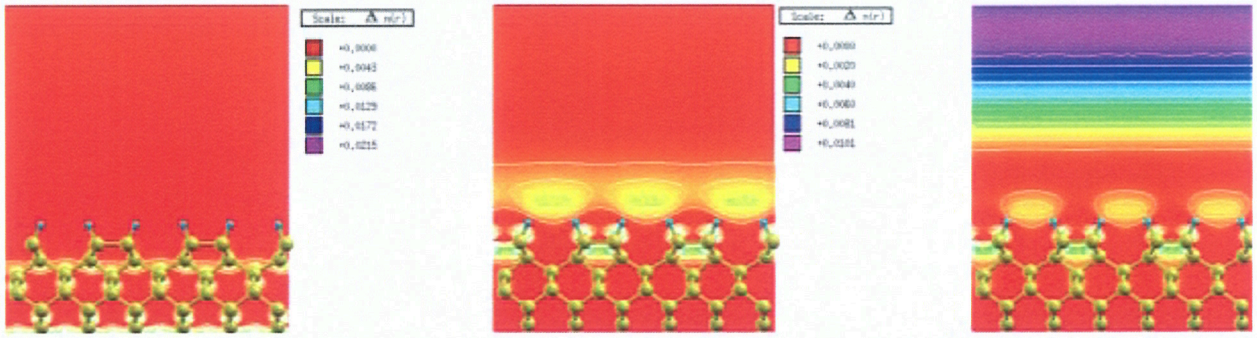


***Figure 1**: From left to right: square modulus of the electronic KS wave function $|\psi(r)|$ of the HOMO, LUMO and LUMO+1 states at theC(0 0 1)2 × 1 :H surface. The LUMO+1 state has a NFE character.*

Contribution of steps to optical properties of vicinal diamond (100):H surfaces

We have carried out one of the very few studies of optical properties of surfaces with the inclusion of steps. Due to the large supercells that one has to use in order to include terraces, and the large number of k-points that have to be employed for the convergence of the optical properties, the use of high performance parallel platforms, as CRESCO is, was mandatory.

This work has been done in strong collaboration with experimentalists. Reflectance anisotropy spectroscopy (RAS) measurements of hydrogen-terminated vicinal C(100) surfaces were performed in Liverpool (UK) and our aim was to connect the empirical RAS signatures with the presence of single (S) and double (D) height steps on these surfaces. We compared the empirical step spectral signatures with the results of theoretical calculations, within DFT and the GW method, for the flat and stepped surfaces using the models suggested by Chadi [12] and Tsai and Yeh [13]. We have simulated the S_A, S_B, S'_B, D_A, and D'_A steps of both models and computed the corresponding RAS at a single-particle level . The theoretical RAS spectra are presented in *Figure 2* [14]. The comparison between experimental RAS and theoretical simulations for the flat and vicinal surfaces suggests that, in the energy range studied here (1–5 eV), the optical signal arises from single steps S_B and double height steps D_A steps of the Chadi type.

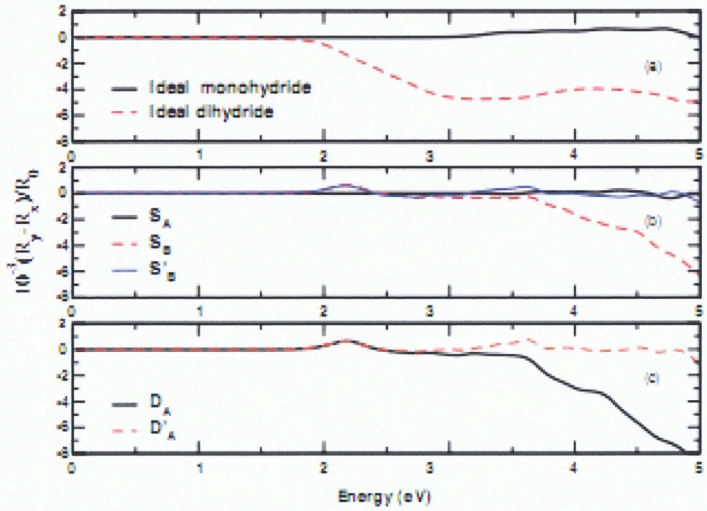


***Figure 2**: DFT-simulated RAS for (a) ideal flat 2 × 1 C(100):H and 1 × 1 C(100):2H, (b) single steps, and (c) double steps.The y direction is along the dimer rows in (a) and along the steps edge in (b) and (c).*

Test of long-range exchange-correlation kernels of time-dependent density functional theory at surfaces: Application to Si(111)2x1

We have carried out the first Time-Dependent Functional Theory (TDDFT) study of a surface. The aim was to explore the quality of well known TDDFT kernels in the description of bound and not bound excitons in 2D.

The Si(111)2x1 surface, being a well characterized system, has been selected as test case. Experimental optical spectra in fact do exist, for this surface, from the infrared to the ultraviolet energy range. We have carried out ab-initio calculations using different methods, from DFT to BSE and, within TDDFT, we have tested the ability of different kernels to describe the optical features in a wide range of energies. We found good agreement between TDDFT and BSE results, by using in TDDFT a long-range frequency-dependent exchange-correlation kernel derived from the many-body formalism, as shown in *Figure 3*.

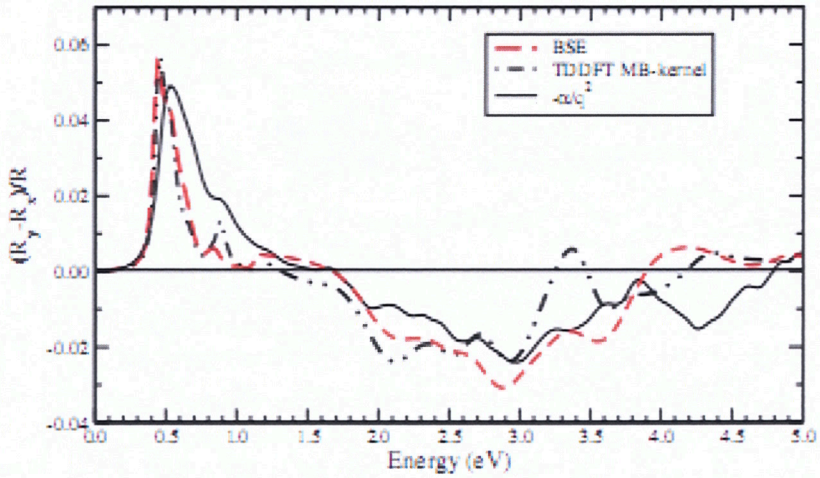


***Figure 3**: Calculated RAS for Si(111)2x1 within TDDFT- many-body kernel, TDDFT-RORO kernel with α=+0.2, and BS equation.*

The agreement between theory and experiment is very good in the whole frequency range studied. Excitonic effects, important in the infrared part of the spectrum, are less pronounced in the visible and UV ranges. The static long range kernel (RORO kernel [15]) also gives reasonably good results, in the full energy range. This fact, together with the enormous computational advantages that this simplified kernel carries, opens a way to cheap TDDFT calculations on surfaces [16].

Optical spectra of ZnO in the far ultraviolet: First-principles calculations and ellipsometric measurements

We have performed a joint experimental-theoretical study based on ellipsometry data of the dielectric function of wurtzite ZnO in a wide energy range 2.5-32 eV. The ordinary and extraordinary components show a strong anisotropy above 10 eV, a feature for which ZnO deviates from the other II-VI wurtzite compounds. With the aid of ab initio calculations, performed within many-body perturbation theory MBPT and within time-dependent density-functional theory TDDFT, we have analyzed the origin of the measured optical structures (see *Figure 4*) [17]. TDDFT, with the use of a static long-range exchange-correlation kernel [15], proves to be a cheaper computational tool than MBPT to yield a good description of the whole spectrum.

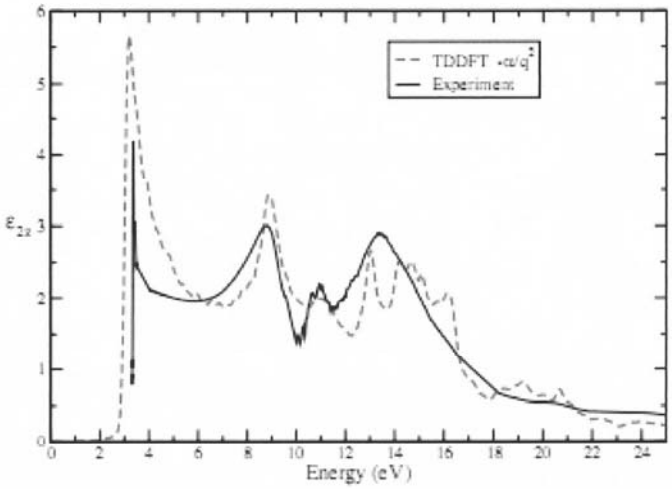


***Figure 4**: ZnO dielectric function. Experimental results for the extraordinary component, compared with the TDDFT calculations.*

BIBLIOGRAPHY

[1] P. Hohenberg and W. Kohn, Phys. Rev. B 136, 864 (1964).

[2] R. M. Dreizler and E. K. U. Gross, Density Functional Theory (Springer-Verlag, Heidelberg, 1990). R. O. Jones and O. Gunnarsson, Rev. Mod. Phys. 61, 689 (1989).

[3] W. Kohn and L. J. Sham, Phys. Rev. A 140, 1113 (1965).

[4] L. Fetter and J. D. Walecka, Quantum Theory of Many Body Systems (McGraw-Hill, New York, 1981).

[5] L. Hedin, Phys. Rev. A 139, 796 (1965). R. D. Mattuck, A Guide to Feynman Diagrams in the Many Body Problem (McGraw-Hill, New York, 1976).

[6] L. Hedin and B. J. Lundquist, in: Solid State Phys., Vol. 23, edited by H. Ehrenreich, F. Seitz, and D. Turnbull (Academic Press, New York, 1969), p. 1.

[7] F. Aryasetiawan and O. Gunnarsson, Rep. Prog. Phys. 61, 237 (1998). Aulbur W. G. Johsson and J. W. Wilkins, Solid State Phys. 54, 1–218 (2000).

[8] S. L. Adler, Phys. Rev. 126, 413 (1962). N. Wiser, Phys. Rev. 129, 62 (1963).

[9] O. Pulci et al. Phys. Stat. Sol. (b) 13, 2737 (2005).

[10] G. Onida, L. Reining, and A. Rubio, Rev. Mod. Phys. 74, 601 (2002).

[11] Margherita Marsili and Olivia Pulci, J. Phys. D: Appl. Phys. 43, 374016 (2010).

[12] D. J. Chadi, Phys.Rev.Lett. 59, 1691 (1987).

[13] M. H. Tsai and Y. Y. Yeh, Phys. Rev. B 58, 2157 (1998).

[14] M.Schwitters, D. S. Martin, P. Unsworth, T. Farrell, J. E. Butler, M. Marsili, O. Pulci, and P. Weightman, Phys. Rev. B 83, 085402 (2011).

[15] L. Reining, V. Olevano, A. Rubio, and G. Onida, Phys. Rev. Lett. 88, 066404 (2002).

[16] O. Pulci, A. Marini, M. Palummo, and R. Del Sole, Phys. Rev. B 82, 205319 (2010).

[17] Paola Gori, Munise Rakel, Christoph Cobet, Wolfgang Richter, Norbert Esser, Axel Hoffmann, Rodolfo Del Sole, Antonio Cricenti, and Olivia Pulci, Phys. Rev. B 81, 125207 (2010).

A MULTISCALE STOCHASTIC VOLATILITY MODEL IN MATHEMATICAL FINANCE

Lorella Fatone

Dipartimento di Matematica e Informatica
Università di Camerino, e-mail: lorella.fatone@unicam.it

Francesca Mariani

CERI- Centro di Ricerca Prevenzione e Controllo dei Rischi Geologici
Università di Roma "La Sapienza", e-mail: fra_mariani@libero.it

Maria Cristina Recchioni

Dipartimento di Scienze Sociali
Università Politecnica delle Marche, e-mail: m.c.recchioni@univpm.it

Francesco Zirilli

Dipartimento di Matematica "G. Castelnuovo"
Università di Roma "La Sapienza", e-mail: f.zirilli@caspur.it

An explicitly solvable multiscale stochastic volatility model that generalizes the Heston model has been introduced and studied. The model describes the dynamics of the price of an asset and of its two stochastic variances using a system of three Ito stochastic differential equations. The two stochastic variances vary on two different time scales and can be regarded as auxiliary variables introduced to model the dynamics of the asset price. Let R and R^+ be the sets of real and of positive real numbers respectively, t be a real variable that denotes time, S_t be a stochastic process describing the asset (stock, commodity, index) price at time t, $t \geq 0$, $x_t = \log(S_t/S_0)$, $t \geq 0$, be the corresponding log-return and $v_{1,t}$, $v_{2,t}$, $t \geq 0$, be the associated stochastic variances. The following system of Ito stochastic differential equations is assumed to describe the dynamics of the stochastic process $(x_t, v_{1,t}, v_{2,t})$, $t > 0$:

$$dx_t = (\hat{\mu} + a_1 v_{1,t} + a_2 v_{2,t})dt + b_1\sqrt{v_{1,t}}\,dW_t^{0,1} + b_2\sqrt{v_{2,t}}\,dW_t^{0,2}, \quad t > 0, \quad (1)$$

$$dv_{1,t} = \chi_1(\theta_1 - v_{1,t})dt + \varepsilon_1\sqrt{v_{1,t}}\,dW_t^{1}, \quad t > 0, \quad (2)$$

$$dv_{2,t} = \chi_2(\theta_2 - v_{2,t})dt + \varepsilon_2\sqrt{v_{2,t}}\,dW_t^{2}, \quad t > 0, \quad (3)$$

with initial conditions:

$$x_0 = \tilde{x}_0, \quad (4)$$

$$v_{1,0} = \tilde{v}_{1,0}, \quad (5)$$

$$v_{2,0} = \tilde{v}_{2,0}. \quad (6)$$

where the quantities $\hat{\mu}$, $\chi_i, \theta_i, \varepsilon_i$, a_i, b_i, i = 1,2, are real constants satisfying the following conditions: : $\chi_i \geq 0$, $\varepsilon_i \geq 0$, $\theta_i \geq 0$, $b_i \geq 0$, i = 1,2. These conditions are due to the financial meaning of the constants. The stochastic processes $W_t^{0,1}$, $W_t^{0,2}$,W_t^1, W_t^2, t>0, are standard Wiener processes such that $W_0^{0,1}= W_0^{0,2}=W_0^1=W_0^2=0$, $dW_t^{0,1}$, $dW_t^{0,2}$,dW_t^1, dW_t^2 are their stochastic differentials and we assume that $<dW_t^{0,1}dW_t^{0,2}>= 0$, $<dW_t^{0,1}dW_t^1>= \rho_{0,1}dt$, $<dW_t^{0,2}dW_t^2>= \rho_{0,2}dt$, $<dW_t^1dW_t^2>=0$, where $<\cdot>$ denotes the mean of $\cdot$ and $\rho_{0,1}$, $\rho_{0,2} \in [-1,1]$ are constants known as correlation coefficients. The fact that the model is a two scale stochastic volatility model is translated in the assumption that $\chi_1 << \chi_2$. The random variables $\tilde{x}_0$, $\tilde{v}_{1,0}$, $\tilde{v}_{2,0}$ are assumed to be concentrated in a point with probability one and $\tilde{x}_0$ is chosen equal to zero. Moreover $\tilde{v}_{1,0}$, $\tilde{v}_{2,0}$ are assumed to be positive. The model (1), (2), (3) introduced in [1] is the announced multiscale stochastic volatility model and generalizes the Heston stochastic volatility model [2]. In fact choosing a_1=-1/2, a_2 =0, b_1 =1, b_2 = 0 the equations (1), (2) reduce to the Heston model and they become decoupled from equation (3).

Under the assumptions made on the correlation coefficients between the Wiener processes, the transition probability density function of the stochastic process solution of (1), (2), (3), (4), (5), (6) can be represented as a one dimensional integral of an explicitly known integrand. In this sense the model is explicitly solvable [1]. In mathematical finance multiscale stochastic volatility models are used for several purposes such as, for example: i) the study of commodity prices, in fact usually commodity prices are characterized by spikes that can be modelled using a fast time scale volatility together with an intermediate time scale volatility (see [3]), ii) the study of financial products that live for long time periods (such as life insurance contracts) that can be modelled using a long time scale volatility together with an intermediate time scale volatility (see [5]). In alternative to multiscale stochastic volatility models stochastic jump models, such as, for example, Lévy's processes with stochastic volatility, can be used. However this last type of models is explicitly solvable only in very special circumstances and in general the transition probability density functions associated to them are defined as solution of integro-differential equations that must be solved numerically. That is the use of jump models in practical circumstances is, in general, cumbersome when compared to the use of the model proposed here.

In [1] a risk neutral measure associated to the multiscale stochastic volatility model proposed and the associated risk premium parameters are defined. Formulae to price European vanilla options (call and put) in the multiscale stochastic volatility model are derived [1]. The formulae obtained are one dimensional integrals of explicitly known integrands and in a special case reduce to the corresponding formulae of the Heston model [2]. The option price formulae obtained are used to study the values of the model parameters, of the correlation coefficients of the Wiener processes defining model (1), (2), (3), (4), (5), (6) and of the initial stochastic variances (see (5), (6)) implied by the "observed" option prices using both synthetic and real data. That is we generalize to the model proposed here the so called term structure analysis of the implied volatility used in the Black Scholes model. This analysis is translated in the solution of constrained nonlinear optimization problems with objective functions containing integrals that must be evaluated numerically. The real data studied are the S&P 500 index and the associated European option prices in the year 2005 [1] and two series of electric power prices taken from the US electricity markets [3]. Using a slightly different model the HFRI (Hedge Fund Research) Equity index and its relation with the S&P 500 index in the years going from January 1990 to June 2007 are studied [4]. The real data analysis presented shows that the multiscale stochastic volatility model is a useful model and that, in particular, gives a satisfactory explanation of the observed option prices including those of the out of the money options and that the values of the model parameters, of the correlation coefficients and of the initial stochastic variances implied by the option prices can be used to obtain high quality forecasts of the underlying asset (S&P 500,

electric market prices,..) log-return and of the option prices. A general reference to the work of the authors and of their co-authors in mathematical finance is the website: http://www.econ.univpm.it/recchioni/finance, [6].

Let us consider for the model (1)-(6) the computational effort required to derive the values of the model parameters from price data.

Let R^n be the n dimensional real Euclidean space. Let us introduce the vector $\underline{\Theta}=(\varepsilon_1,\theta_1,\rho_{0,1},\chi_1,\tilde{v}_{0,1},\hat{\mu},\lambda_1,\varepsilon_2,\theta_2,\rho_{0,2},\chi_2,\tilde{v}_{0,2},\lambda_2) \in R^{13}$ where the parameters λ_i, i=1,2, are the risk premium parameters that must be used when we consider asset and option prices and the remaining parameters have already being introduced. The set $M \subset R^{13}$ of the feasible vectors is defined as follows:

$$M = \{\underline{\Theta}=(\varepsilon_1,\theta_1,\rho_{0,1},\chi_1,\tilde{v}_{0,1},\hat{\mu},\lambda_1,\varepsilon_2,\theta_2,\rho_{0,2},\chi_2,\tilde{v}_{0,2},\lambda_2)\in R^{13}|$$
$$\varepsilon_i,\chi_i,\theta_i \geq 0,\ \frac{2\chi_i\theta_i}{\varepsilon_i^2}\geq 1,\ -1\leq\rho_{0,i}\leq -1,\ \tilde{v}_{0,i}\geq 0,\chi_i+\lambda_i\geq 0,i=1,2\}. \qquad (7)$$

The constraints that define M translate some elementary properties of model (1), (2), (3). Moreover, let t_i, i = 0,1,…,n, be time values such that $t_i < t_{i+1}$, i = 0,1,…,n-1, without loss of generality we choose $t_0 = 0$, and for later convenience we define $t_{n+1} = +\infty$. Let $\left(\tilde{x}_i,\tilde{C}_i,\tilde{P}_i\right)$ be respectively the log-return of the asset price S_t and the prices of European vanilla call and put options on the asset having maturity time T_i ($T_i > t_i$) and strike price K_i, observed at time $t = t_i$, i=0,1,…,n. We suppose that the option prices observations $\tilde{C}_i$, $\tilde{P}_i$ are affected by a Gaussian error with mean zero and known variance ϕ_i, i=0,1,…,n, and that the asset log-returns $\tilde{x}_i$, i = 0,1,…,n, are observed without error. The calibration problems considered in [1], [3] are the following:

1) Given the observations $F_t = \{\left(\tilde{x}_i,\tilde{C}_i,\tilde{P}_i\right)$: i such that $t_i \leq t\}$, $t > 0$, find the vector $\underline{\Theta} \in$ M that "makes most" likely the occurrence of the observations $F_t = \{\left(\tilde{x}_i,\tilde{C}_i,\tilde{P}_i\right)$: i such that $t_i \leq t\}$, $t > 0$;
2) Given the observations $F_t = \{\left(\tilde{C}_i,\tilde{P}_i\right)$: i such that $t_i \leq t\}$, $t > 0$; of the option prices associated to the random variables x_t, $v_{1,t}$ and $v_{2,t}$, $t > 0$, solution of (1), (2), (3), (4), (5), (6) find the vector $\underline{\Theta} \in$ M that minimizes the distance between the observed option prices and the theoretical option prices [1]. Note that in this case since only option prices are used as data the risk premium parameters λ_1 and λ_2 are not necessary.

Problem 1) is a maximum likelihood problem and Problem 2) is a least square problem. Note that to solve Problem 1 we must first solve the following filtering problem:

3) Filtering problem: Given the vector $\underline{\Theta} \in$ M find the probability density function $p(x,v_1,v_2,t|F_t, \underline{\Theta})$, $(x,v_1,v_2) \in R\times R^+\times R^+$, $t > 0$, of the random variables x_t, $v_{1,t}$ and $v_{2,t}$, $t > 0$, solution of (1), (2), (3), (4), (5), (6), conditioned to the observations F_t =

$\{\left(\tilde{x}_i, \tilde{C}_i, \tilde{P}_i\right)$: i such that $t_i \le t\}$, $t > 0$, and forecast the values of the state variables x_t, $v_{1,t}$, $v_{2,t}$ for $t > 0$ and in particular for $t \neq t_i$, $i = 0,1,\ldots,n$ and for $t > t_n$.

Note that Problem 1 can be easily reformulated and solved when only log-returns are available as data [3], [4].

In our numerical experiments several thousands problems like Problem 1 and Problem 2 have been solved in the process of calibrating model (1)-(6) on real data. A steepest ascent algorithm is used to solve Problem 1 and a steepest descent algorithm is used to solve Problem 2. Note that Problem 1 and Problem 2 have been solved for many different sets of data and that these problems are independent of each other, this means that they can be solved in parallel.

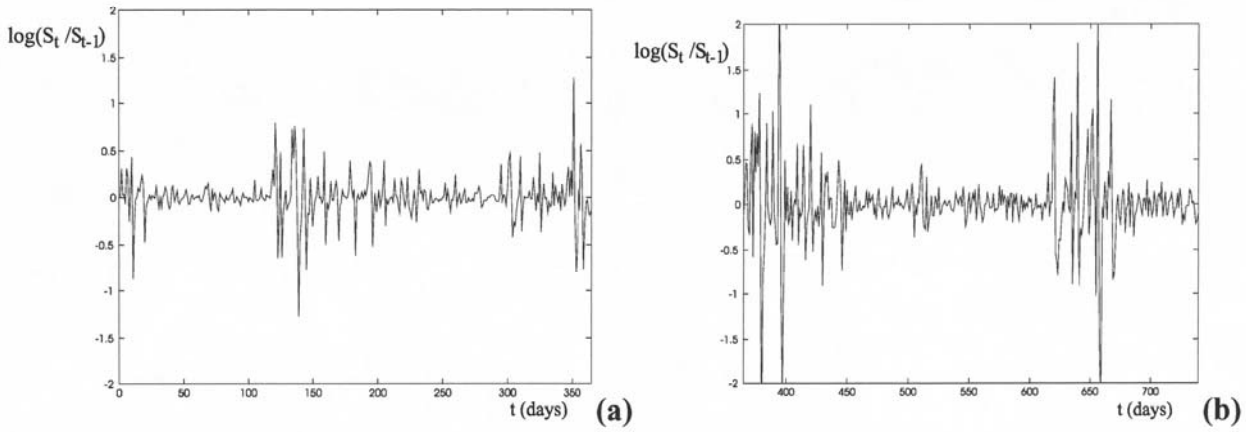


Figure 1: *(a) Log-return increment $log(S_t/S_{t-1})$ of the electric power prices versus time t (days),*

(a) $\hat{S}_i$, *i=0,1,2,...,364 first year of observation, daily data,*

(b) $\hat{S}_i$, *i=365,366,...,764, second year of observation, daily data.*

Let us give an example of the computational work needed to analyze a time series of real data. Let us consider a time series of electric power price data taken from the U.S. electricity market. The time series consists of 765 daily price observations $\hat{S}_i$, i = 0,1,...764. In this experiment no option price data are used. This time series is used to calibrate the multiscale model and a data window made of 26 consecutive daily observations is chosen. This window is moved along the time series substituting the first observation of the window with the next observation after the window. Using this procedure in the case of the time series shown in *Figure 1(a), (b)* we solve 740(740=765-26+1) calibration problems using the maximum likelihood approach (Problem 1). These problems are solved in parallel thanks to the fact that they are independent of each other.

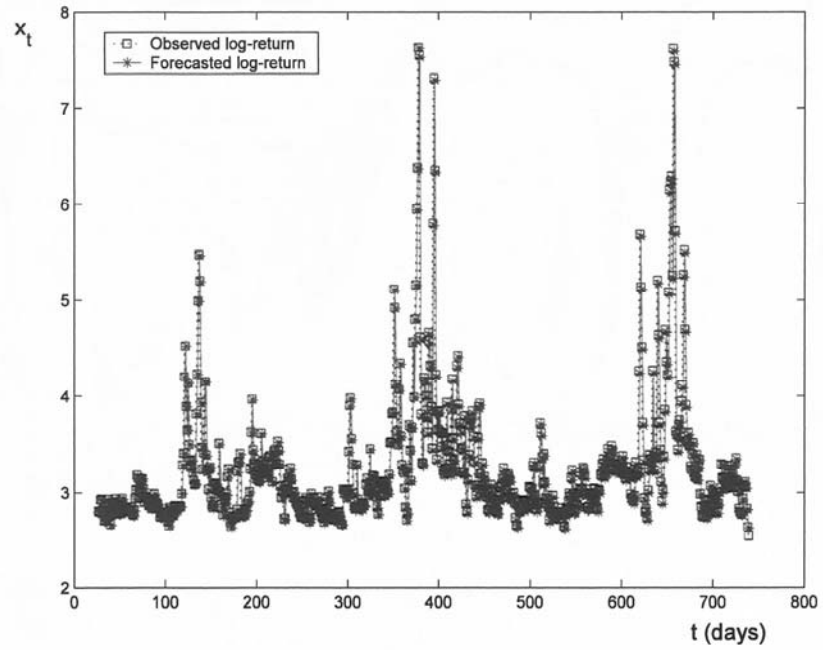


Figure 2: *Forecast values (one day in the future) (stars) and observed values (squares) of the log-returns of electric power prices (time series with spikes, Figure 1) versus time t (days).*

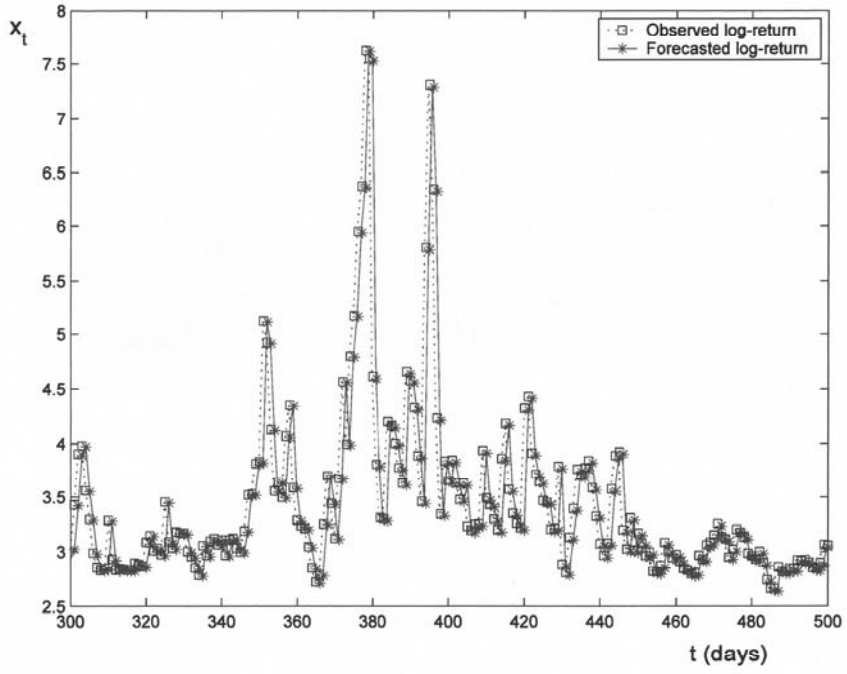


Figure 3: *Forecast values (one day in the future) (stars) and observed values (squares) of the log-returns of electric power prices (time series with spikes, Figures 1(b)) versus time t (days).*

To establish the quality of the solution of the calibration problem we do a forecasting exercise. The estimated values of the parameters and of the initial stochastic variances obtained solving the calibration problems (Problem 1) are used to produce through a tracking procedure forecasts of the electric power prices [3] (or of S&P 500 index [3], or of HFRI Equity index [4]), of the associated stochastic variances and of the option prices. The forecasts of the asset prices and of the option prices are compared with the prices actually observed (see [1], [3], [4]). This comparison shows that the forecasts are of very high quality even when we consider "spiky" electric power price data.

Figures 2 and *3* shows (on two different time scales) the forecasted values of the log-return (one day in the future) and the time series of the log-returns corresponding to the time series shown in Figure 1 (a), (b) as a function of time t (days). *Figures 2* and *3* show that the (one day in the future) forecast log-returns are of very high quality. In [1] the formulation given in Problem 2 to the calibration problem has been used to study time series of real data. In this study different sets of data produce problems independent of each other and these last problems are solved in parallel.

REFERENCES

[1] FATONE L., MARIANI F., RECCHIONI M.C., ZIRILLI F. *"An explicitly solvable multi-scale stochastic volatility model: option pricing and calibration problems"*, Journal of Futures Markets **29** (2009) pp. 862-893. (http://www.econ.univpm.it/recchioni/finance/w7).

[2] HESTON S.L. *"A closed-form solution for options with stochastic volatility with applications to bond and currency options"*, Review of Financial Studies **6** (1993) pp. 327-343.

[3] FATONE L., MARIANI F., RECCHIONI M.C., ZIRILLI F. *"The analysis of real data using a multiscale stochastic volatility model"*, to appear in European Financial Management. (http://www.econ.univpm.it/recchioni/finance/w9).

[4] CAPELLI P., MARIANI F., RECCHIONI M.C., SPINELLI F., ZIRILLI F. *"Determining a stable relationship between hedge fund index HFRI-Equity and S&P 500 behaviour, using filtering and maximum likelihood"*, Inverse Problems in Science and Engineering **18** (2010) pp. 83-109. (http://www.econ.univpm.it/recchioni/finance/w8).

[5] BERTOCCHI M.I., GIACOMETTI R., RECCHIONI M.C., ZIRILLI F. *"Pricing life insurance contracts as financial options: the endowment policy case"*, submitted to Insurance: Mathematics and Economics (2010). (http://www.econ.univpm.it/recchioni/finance/w12).

[6] http://www.econ.univpm.it/recchioni/finance/.

EXCHANGE FLOW THROUGH THE STRAIT OF GIBRALTAR AS SIMULATED BY A NON-HYDROSTATIC MODEL

Gianmaria Sannino[1], Luca Liberti[2], J. C. Sánchez-Garrido[3]

[1]ENEA – UTMEA-CLIM. Casaccia Research Centre.

[2]ISPRA Via Curtatone, 3 00185 Roma.

[3]Grupo de Oceanografia Fısica. Dpto. Fisica Aplicada II, Campus de Teatinos s/n, 29071, University of Malaga, Malaga, Spain.

The Strait of Gibraltar is a remarkable place in the World Ocean connecting the Mediterranean Sea and the Atlantic Ocean through a narrow channel with a minimum width of 14 km and a complex bottom topography (see *Figure 1b and c*). The oceanography of the entire Mediterranean basin is deeply influenced by the water flows at the Strait of Gibraltar where a two-way baroclinic water exchange takes pace: salty and dense Mediterranean Water (MW) flows towards the Atlantic in the bottom layer at an average rate of 0.8 Sv (Sverdrup, 1 Sv = 10^6 m^3 s^{-1}) and a slightly higher volume rate (0.85 Sv on average) of Atlantic Water (AW) enters the Mediterranean in the surface layer to compensate for the net evaporative losses over the basin. Several works have pointed out that correct prediction of the overall mediterranean circulation, and sea-atmosphere interaction, require accurate modelization of these exchanges [1,2,3] which is particularly challenging because of the complex topography of the area, the density related (baroclinic) effects in the flow field and the variability in the flow pattern.

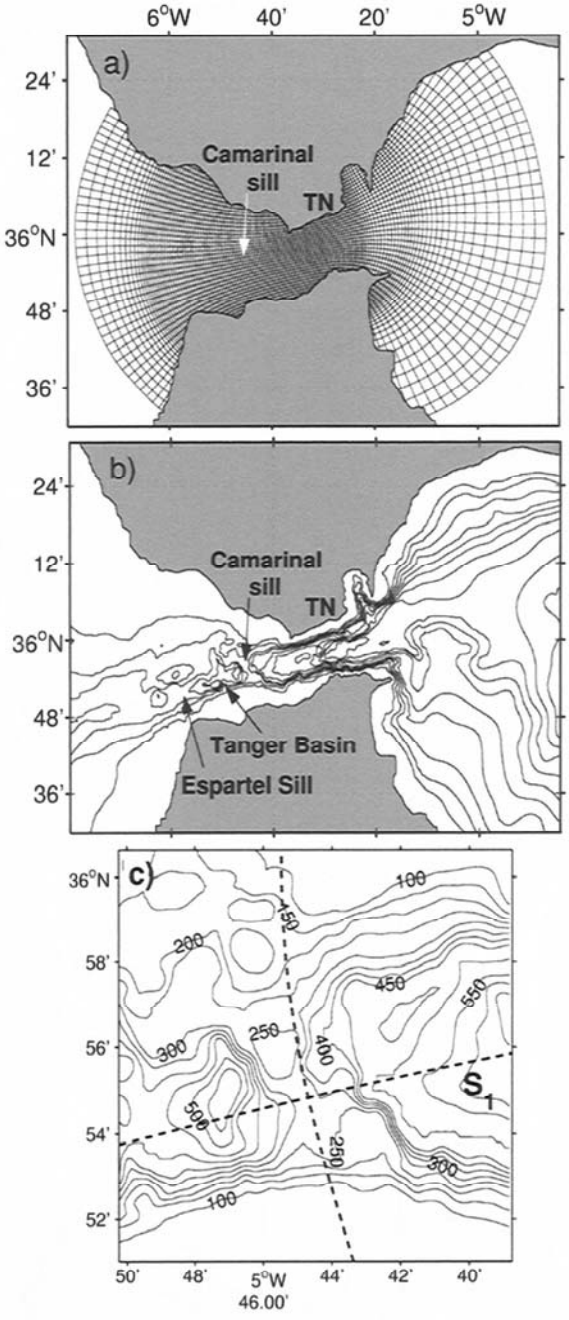


Figure 1:

a) Computational grid used in current model. Only 2% of actual grid is shown in the Figure.

b) Bottom topography of the Strait of Gibraltar (isobath shown every 100 m).

c) Detailed bottom topography of the Strait of Gibraltar in the area of Camarinal Sill (CS), Tangier Basin (TB) and Espartel Sill (ES).

In fact, the two-way exchange is far from being steady and fluctuates on a very broad spectrumof time scales including tidal, meteorological, seasonal and inter-annual variations [4]. The exchange intemittently undergoes hydraulic control over the main sills in the Strait and in the narrowest section of the Strait when flow condition becomes super-critical. Hydraulic control has significant influence on water exchanges in the strait and in their sensitivity to external perturbations.

The interaction between the stratified flow, tidal forcing and the main sill of the Strait, Camarinal Sill (CS) leads to periodic generation of internal waves with amplitudes exceeding 100 m (Large Amplitude Internal Wave or LAIW). After their generation LAIWs propagate eastward in the Mediterranean (*Figure 2*) carrying a large amount of energy which is eventually dissipated in turbulence and mixing in shallow areas where the internal waves break.

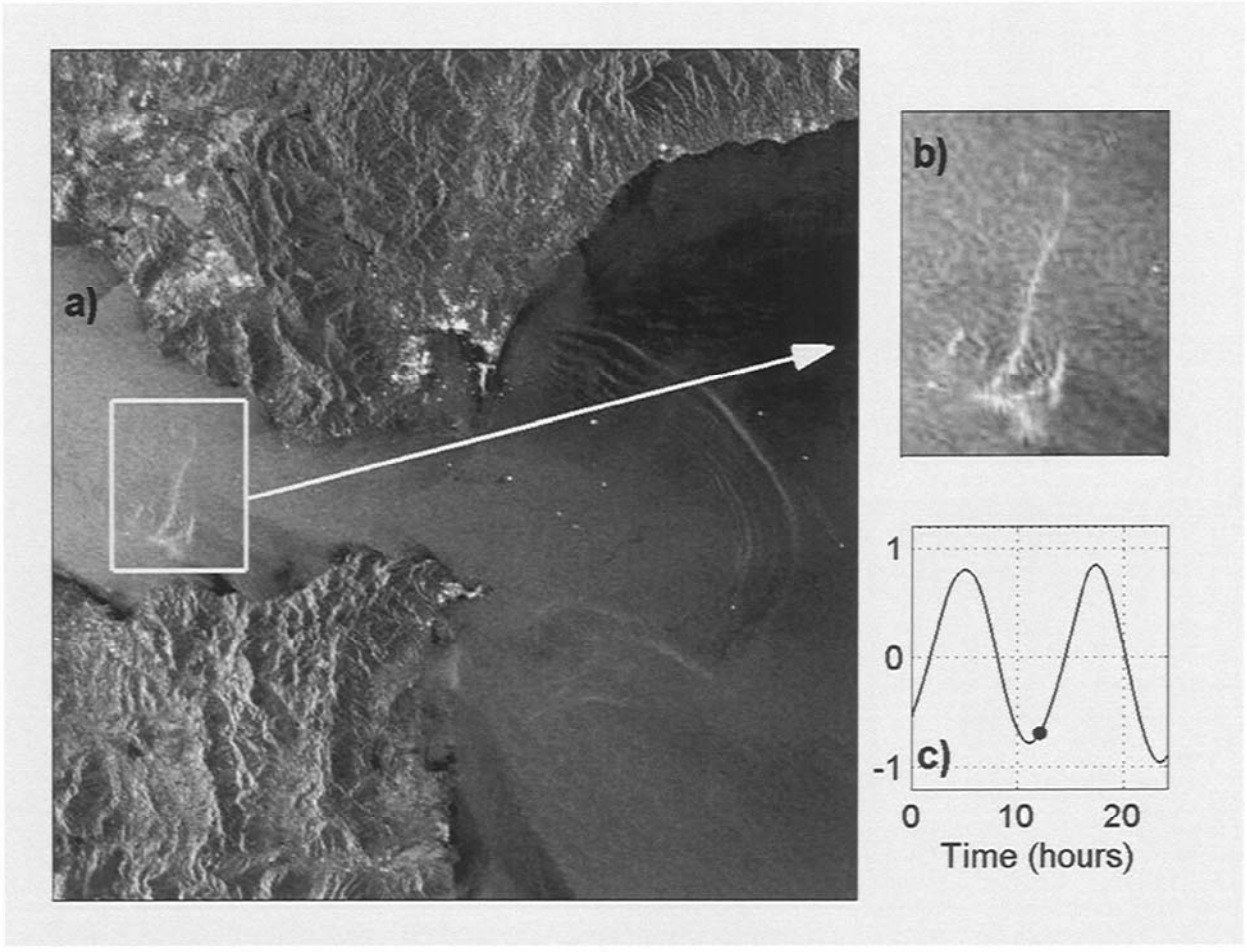


Figure 2: *a) Synthetic aperture radar (SAR) image of the Strait of Gibraltar showing the development of internal waves in CS and the propagation of previously generated packet of LAIWs in the Alboran Sea (date of reception: 2005-11-28 22:09. Envisat ASAR Data provided by ESA Project AO.SMOS4657) b) Detail of CS Area c) Barotropic current velocity (depth-averaged velocity) over CS around the time of image acquisition estimated by harmonic analysis of ADCP measurements. The dot indicates the time of acquisition.*

Although the LAIW generation mechanism at CS is well understood some of the processes concerning their evolution and propagation are not yet addressed by numerical studies available in literature. Previous modelling studies were based on simplified (laterally averaged) formulations that neglect important effects due to the complex topography and earth rotation. Furthermore, these studies included only the M_2 constituent of the tidal forcing and could not reproduce the tidal cycle variability due to the modulation exerted by other tidal constituents (S_2, O_1, K_1).

On the other hand, some authors found that the strong entrainment and mixing in the Strait leads to the formation of a thick interfacial layer where density and velocity gradually change along the water column and argued that hydraulic control theory based on a two-layer system is inadequate to correctly account for the flow regime in the Strait [5,6]. Therefore, hydraulic control theory based on a two-layer system might be insufficient to correctly describe this process in the Gribraltar Strait and a more advanced formulation based on three layers is required.

In order to address all the above-mentioned issues a novel numerical study taking advantage of the computational power made available by the CRESCO supercomputer was carried over by ENEA UTMEA-CLIM in collaboration with researches from University of Malaga and ISPRA.

In the present study an extremely high resolution, three dimensional and fully non-linear model based on the MITgcm general circulation code was developed. The MITgcm solves the fully nonlinear, non-hydrostatic Navier-Stokes equation under the Boussinesq approximation for finite-volume discretization on a curvilinear computational grid. The model formulation includes free surface and partial step topography, and the code is higly parallelized following a domain decomposition approach and suitable to be used on large computer clusters. The model domain extends from 6.3° W to 4.78° W and was discretized by a non-uniform curvilinear orthogonal grid of 1440 x 210 points (*Figure 1a*). Horizonal spatial resolution along the axis of the Strait, Δx (across the Strait,Δy) ranges between 46 -63 m (175-220 m) in the Camarinal Sill Area and mesh sizes is always less than 70 m (340 m) in the middle of the Strait between Espartel Sill and Camarinal Sill and less than 70 m (200) between Camarinal Sill and Tarifa Narrow. In order to adequately resolve the pycnocline the model has 53 vertical z-levels with a width of 7.5 m in the first 300 m gradually increasing to a maximum of 105 in the bottom 13 levels. Bottom topography was obtained merging the ETOPO2 bathymetry [U.S. Department of Commerce and NOAA/NGDG, 2001] with a very high resolution bathymetry chart of Sanz et Al. Inital conditions for temperature and salinity were derived from the climatologic MedarMedAtlas Database [MEDAR Group 2002] for the month of April. The two way exchange in the Strait is achieved laterally forcing the model through the imposition of the mean baroclininc velocities and tracers obtained from the intermediate resolution model by [6].

The new model represents the state of the art of the numerical models of the Gibraltar Strait thanks to the combination of extremely detailed spatial resolution, non-hydrostatic formulation and external forcing accuracy. The simulation takes place during a full tropical month (approx. 30 days) in order to catch the entire neap-spring cycle tidal variability. The model ran with a 4 s time step on the CRESCO computing facilities using an average of 800 CPUs. In this configuration the computational time was approximately the same as the physical time and a complete simulation was carried out in about 30 days. The MITgcm model features a pure MPI parallelization strategy that scales extremely well even on such a large number of computational nodes. Performace analysis reveal almost linear speedup up to the limit of the computational units available on CRESCO.

Model results are in good agreement with both observational data and results obtained by previous models. The generation and propagation of LAIWs in the Strait of Gibraltar (*Figures 3,*

4) were simulated with great detail and accuracy. The influence barotropic tidal forcing strength on the generation and propagation mechanism was analyzed in depth and the study results were included in a publication [7]. A new study of hydraulic controls in the Gibraltar Strait based on a three layer theory was also conducted and the results will be included in a separate paper. [8]. Further analysis of the model output will likely lead to additional scientific production.

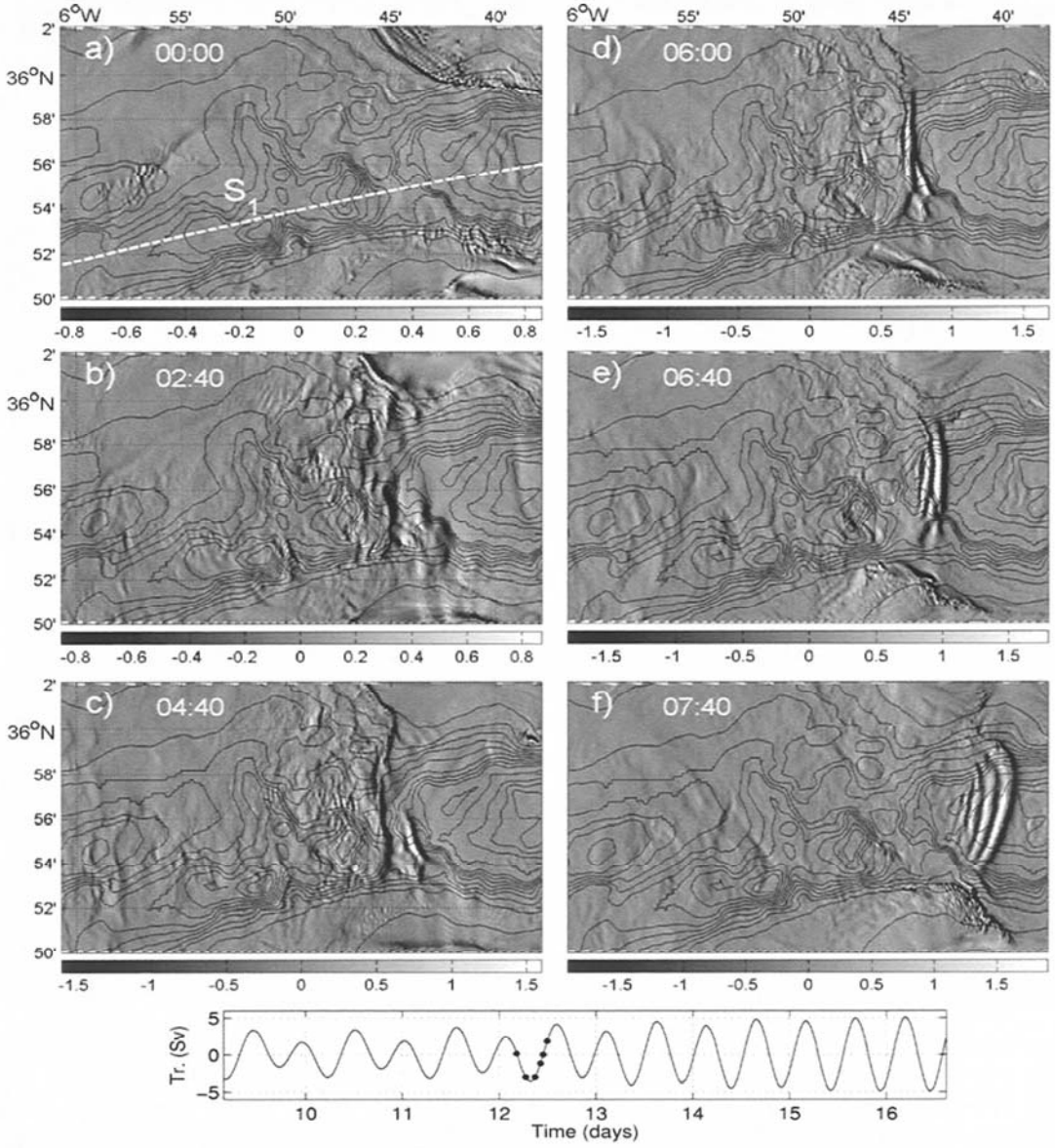


Figure 3: *From a) to f) Time evolution of surface zonal velocity gradient (in 10^3 s^{-1}) during the flood tide. S_1 indicates longitudinal cross-section mentioned in Figure 4. Bottom: barotropic tidal transport during the time-series.*

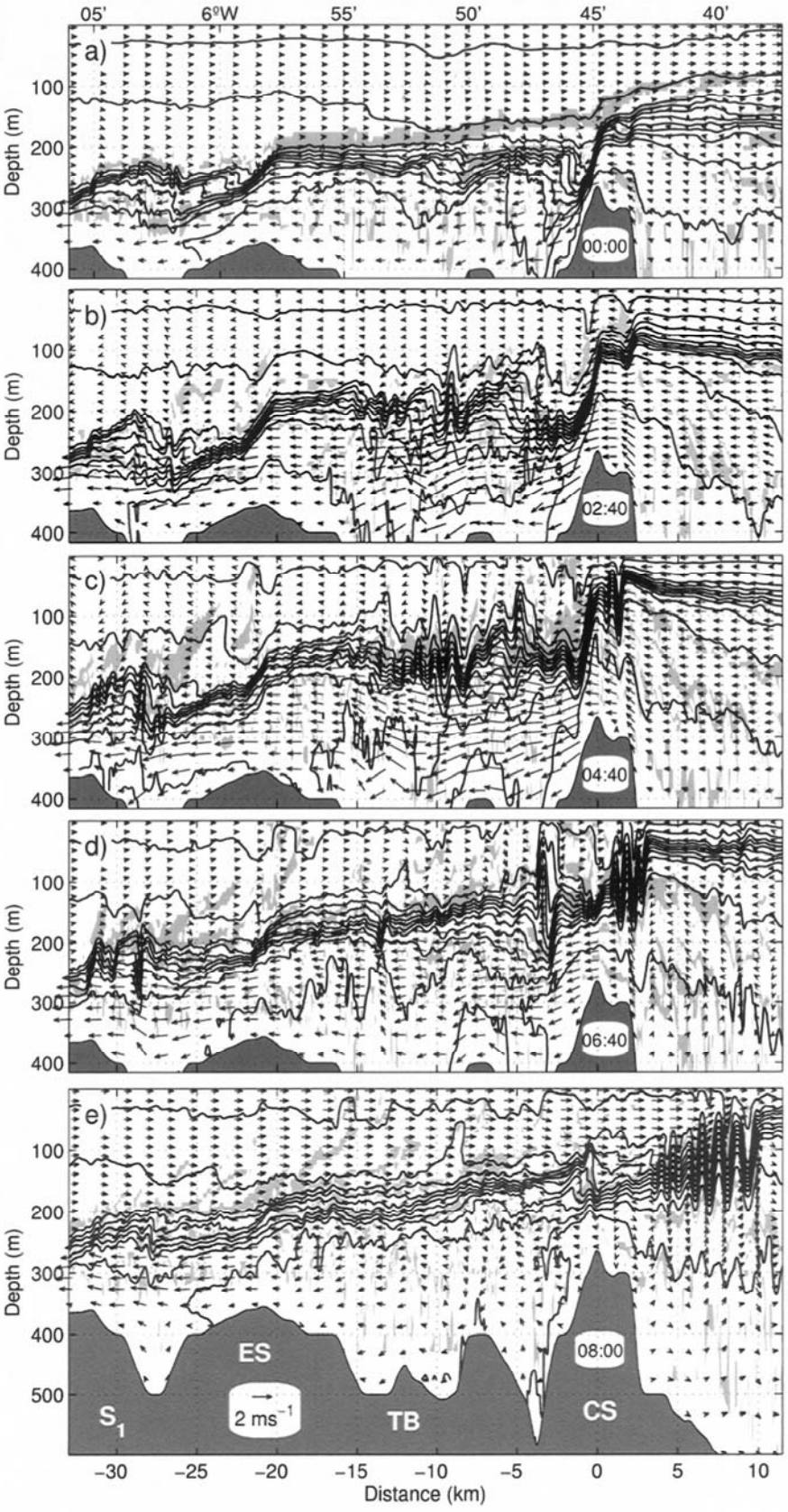


Figure 4: *Time evolution of potential density along the axis of the strait (cross-section S_1 in Figure 3). Isopycnal contours 26.80,27.05,...,28.80,29.02 are shown in the figure. Arrows indicate local current velocity. Gray contours indicate areas where Ri<1/4 and intense water mixing and energy disspation take place.*

REFERENCES

[1] SANNINO G., BARGAGLI, A., ARTALE, V. "*Numerical modeling of the mean exchange through the Strait of Gibraltar*" J. Geophys. Res. 107 (8) (2002) pp. 9 1–24.

[2] SANNINO G., BARGAGLI, A., ARTALE, V. "*Numerical modeling of the semidiurnal tidal exchange through the strait of Gibraltar.*" J. Geophys. Res. 109, (2004) C05011, doi:10.1029/2003JC002057.

[3] SANNINO G., HERRMANN M., CARILLO A., RUPOLO V., RUGGIERO V., ARTALE V. HEIMBACH P. "*An eddy-permitting model of the Mediterranean Sea with a two-way grid refinement at the Strait of Gibraltar*". Ocean Modelling, Vol 30, (2009) 1, 56-72. DOI: 10.1016/j.ocemod.2009.06.002

[4] SÁNCHEZ ROMÁN A., G. SANNINO, GARCIA-LAFUENTE, A. CARILLO, F. CRIADO-ALDEANUEVA. "*Transport estimates at the western section of the Strait of Gibraltar: a combined experimental and numerical modelling study*", Journal of Geophysical Research, Vol 114, (2009) C06002, DOI: 10.1029/2008JC005023.

[5] SANNINO G., CARILLO, A., ARTALE, V. "*Three-layer view of transports and hydraulics in the strait of Gibraltar: A three-dimensional model study*". J. Geophys. Res. 112, (2007) C03010, doi:10.1029/2006JC003717.

[6] SANNINO G., PRATT, L., CARILLO, A. "*Hydraulic Criticality of the Exchange Flow through the Strait of Gibraltar*". J. Phys. Oceanogr. 39 (11) (2009), 2779–2799.

[7] SANNINO G., SÁNCHEZ GARRIDO, JC., LIBERTI, L. "*Exchange flow through the Strait of Gibraltar as simulated by a sigma-coordinate hydrostatic model and a z-coordinate non-hydrostatic model*" in preparation.

[8] SÁNCHEZ GARRIDO, JC., SANNINO, G., LIBERTI, L. "*Generation of Large-Amplitude Internal Waves in 1the Strait of Gibraltar: a Three-Dimensional 2Numerical Study*" submitted JGR.

A FAST METHOD TO DETERMINE THE BEST HIGH PERFORMANCE LINPACK (HPL) PARAMETERS FOR A GIVEN SUPERCOMPUTER

Giovanni Bracco[1], Silvio Migliori[2], Salvatore Podda[1], Salvatore Raia[3]

[1]ENEA UTICT-HPC C. R. Frascati
[2]ENEA UTICT Roma Sede
[3]ENEA UTICT-HPC C. R. Portici

Astract

The performances of parallel applications are evaluated to test both parallel algorithms and advanced computer systems. One of the software widely used to evaluate the floating point operations (flops) rate is HPL (High Performance Linpack) [1] [2]. The run of the HPL application requires a set of input parameters and the performances on a given supercomputer system depend strongly on the values of such parameters. Finding the optimal configuration is a time consuming task when thousands of processors are used. The more it is known that small performances improvements in peak performances allow to climb the rank in top500 list [3] which is based on HPL results. We present here a novel approach to obtain the best set of HPL input parameters among a group of tests by stopping the execution flow of the recursive algorithm. Our main assumption is that information about cache, memory and network performances are already available at the beginning of iterations. We check our hypothesis numerically and report the application of our method for some typical HPL benchmarks. Our examples show that we can save about 80 – 90% of the time respect to the time required by the whole computation. In addition, the approach gives a better understanding of the evolution of the flops rate and it can be used to implement a framework capable to auto detect the optimal set of input parameters.

Introduction

The relative performances of parallel computing systems are evaluated by software benchmarks [4]. They allow to measure, for example, floating point operations (flops) rate, memory bandwidth, latency and bandwidth of the interconnection network. Among various benchmarks ([2], [5], [6]) one of the most used by the HPC (High Performances Computing) community for evaluating flops rate on supercomputers is HPL [2]. It solves in parallel a dense linear system using LU factorization [7]. Moreover HPL is the reference software to draw up the top500 list [3]. In this paper we address the optimization and the tuning issues of the input parameters of the HPL software. Finding the optimal set of parameters is a time consuming task. Each test may require several hours of computing time when thousands of processors are used. A typical parameter search requires to run groups of tests changing one parameter and keeping the other ones fixed. We show how to analyze and evaluate the performances of HPL during the calculation and particularly at the beginning of the iterations of the execution flow. Because of the recursive formulation of LU factorization [8] it is possible to detect and determine an optimal set of parameters by stopping the program after a few iterations and decide which would be the best configuration among a given ensemble of tests. As it will be shown, this allows us to reduce the total execution time by 80 – 90% with respect to a full run of all the tests so that only the optimal cases are run to their completion. A similar approach concerning HPL has been shortly

mentioned in a recent conference [9] but this paper provides the first complete illustration of the method together with the analysis of the basis of its operation. The approach can be also be applied in a more general context not only to save time in choosing the optimal input configuration but also it can be used to implement auto tuning software. An efficient utilization of a given HPC system depends critically on the matching of the algorithm parameters with the characteristics of the system, as processor cache, memory size and bandwidth, interconnection network capabilities. In the case of HPL, the package provides a timing and testing program that allows to achieve the best performances of a system by tuning a set of input parameters. Our approach allows to set up a programmed sequence of tests capable to reveal within a few iterations the optimal set of parameters to be used. The method could be applied to optimize other time consuming scientific applications or to stop inefficient running applications. Moreover this work resulted also in a deeper and fundamental understanding of the dynamic evolution of the program and of its performances.We have choosen HPL software as case study because it represents a clear example where many parameters can be tuned to achieve the best performances. Moreover the Top500 list shows that the peak performances as a function of the rank drop off polynomially with rank [3] [10]. It follows then that small improvements due to case optimization can allow to gain many positions in the ranking of Top500 list.

HPL Algorithm

In this section we briefly revisit the HPL algorithm pointing out its main features and in particular its recursive formulation.

Main algorithm. HPL solves a dense linear system of order N by first computing the LU factorization:

$$A \cdot X = B \qquad (1)$$

where A = LU is the coefficients matrix, L and U are the lower and upper triangular factors respectively and b is the known term. Because L is applied to b (Ly = b) as the factorization go ahead, the solution x is obtained by solving the upper triangular system Ux = y. The data is partitionated into bs×bs blocks and distributed onto P × Q processors grid using 2 – D block cyclic scheme [11]. This assures load balancing as well as scalability. The total number of processors is indicated by p = P × Q and bs is the block size of the blocked algorithm.

Recursive formulation. In the recursive formulation [12] there are kmax = N/bs iterations. The recursion works as follow: at the first iteration (k = 1) a panel of N ×bs elements is factorized by P processes. If required row swap occurs to have the pivot in the main diagonal. The matrix A is decomposed as follows:

$$\begin{aligned} L_{00}U_{00} &= A_{00} \\ L_{10}U_{00} &= A_{10} \\ L_{00}U_{01} &= A_{01} \\ L_{10}U_{01} + L_{11}U_{11} &= A_{11} \end{aligned} \qquad (2)$$

where the panels Li0 have bs columns and U0j have bs rows. *Figure 1* depicts the decomposition given by the equations (2). The panel L10 and L00 are broadcast to the column processes. The upper triangular system is solved to obtain U01 and then it is broadcast to row processes. After the broadcast phase local rank-bs update takes place [13]:

$$A'_{11} = A_{11} - L_{10}U_{01} = L_{11}U_{11} \qquad (3)$$

At the second iteration (k = 2) the decomposition of the equations (2) is applied to the submatrix A_{11} of the equation (3) continuing until k = kmax. In other words each iteration consists of three main steps: the panel factorization, the broadcast of factorized panel and the local rank-bs update. The flops execution rate is calculated with the following equation:

$$R_{\max} = \frac{2/3N^3 + 3/2N^2}{T_p} \qquad (4)$$

where $(2/3N^3 + 3/2N^2)$ is the total number of flops to solve the linear system (1) and Tp is the time (parallel) to complete the task.

Relevant optimization parameters. The performances of parallel applications depend on many factors: cache and memory sizes, latency and bandwidth of interconnection network, number of cores on shared memory, etc. HPL allows to choose various parameters and variants of the blocked algorithm to match the system characteristics. The most relevant parameters are:

N: problem size. Memory occupation depends on this parameter and the execution time is roughly of order $o(N^3)$;

bs: it determines the efficiency of matrix-matrix multiplication in the update phase and the size of the communication volume;

P × Q: the processors grid. The blocked matrix is mapped to that grid using 2 – D cyclic scheme;

bcast: the communication topology used during the panel broadcast.

There are many other parameters but we used the previous ones to show how to use our methodology and check our hypothesis.

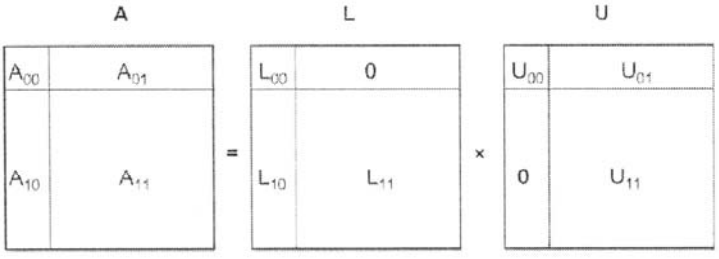


Figure 1: *Decomposition of the matrix* ***A*** *=* ***LU****.*

The Used HPC System

The HPL experiments have been performed on the HPC CRESCO platform [14]. CRESCO (Computational Research Center for Complex Systems) is an ENEA (the Italian National Agency for New Technologies, Energy and Sustainable Economic Development) Project, cofunded by the Italian Ministry of University and Research. The performance for the CRESCO HPC system has ranked 125 in the June 2008 top500 list with Rmax = 17.14 Tflops in the HPL benchmark. The system consists of two main sections respectively oriented for high memory request and moderate parallel scalability and for limited memory and high scalability. Both sections are interconnected by a common InfiniBand 4X DDR network and can operate as a single large integrated system. CRESCO Section 1 (we refer to it as Sect.1 cluster in the rest of the paper), for large memory application is composed by 42 fat nodes IBM x3850 – M2 with 4 Xeon Quad-Core Tigerton E7330 processors (the total number of cores in the first section is equal to 672). CRESCO Section 2 (Sect.2 cluster), for high scalable applications is composed by 256 blades IBM HS21 each supporting dual Xeon Quad- Core Clovertown E5345 processors, 16 GB RAM for total of 2048 cores. The larger system created by joining the two main sections has 2720 cores. The third system where we performed numerical experiments is composed by 6 nodes with 4 sockets AMD Dualcore 8222 (AMD cluster). Further details are reported in Ref. [15].

Methodology, Assumption and Observations

Our methodology is based on a few simple assumptions, on experimental observations and on a basic complexity analysis of the recursive algorithm. The applications of the method and timing results will be reported in sect. (5).

Main assumptions. In the HPL program network features influence the performances of the broadcast steps while the cache and memory performances are critical for the local rank-bs update phase. For each iteration the same type of operations are applied to a matrix with an order scaling as $(N - k \cdot bs)$, where $k = 1, 2, ...,N/bs$. The aim is to deduce the best configuration among a group of tests by analyzing and stop the HPL program after a few iterations. Our basic assumption is that, for a given group of tests, the ranking of flops rate emerges at the beginning of the iterations because the three main steps of the HPL recursive algorithm contain the information about cache, memory and network performances. From the point of view of a single process the update phase is a sequential task carried out with the same efficiency of the matrix multiplication. The communication load depends by the block size, and the single message length does not change as the algorithm proceeds. Each process broadcasts its message to the other row or column processes. What changes during the iterations is number of submatrix of order $bs \times bs$ that each process has to update and the number of messages to be sent. In other words we assume that if two or more tests are ranked in some way, then they will be ranked in the same way at the end of iterations. The latest hypothesis is true only if the communication overhead is roughly constant. Then the requirement is that the ratio between communication time Tbc (bc means broadcast, see next two paragraphs) and total time Ttot be constant as function of k. We check our assumptions numerically analyzing some HPL tests. The results are presented in the next two paragraphs where we first deduce the number of floating point operations as a function of the iterations and then comment some experimental observations.

Complexity analysis as function of the iterations. To calculate the flops rate as a function of k we need to know how the number of flops varies with the iteration. The computational complexity of the various components (LU factorization, triangular solve, etc.) of the HPL algorithm can be

found in literature [16] [2]. Applying the complexity analysis to the recursive formulation, it is possible to treat the number of flops as function of k. By such derivation the number of flops for the panel factorization (F_{pf} (k)), the upper triangular solve (F_{ts}(k)) and the local rank–bs update (F_{up}(k)), are given by the following equations:

$$
\begin{aligned}
F_{pf}(k) &= (N - k \cdot bs)bs^2 - bs^3/3 - bs^2 \\
F_{ts}(k) &= (N - k \cdot bs)bs(bs - 1) \\
F_{up}(k) &= 2 \cdot bs(N - k \cdot bs)^2
\end{aligned}
\tag{5}
$$

In general, the total number of flops and the flops rate at a given k will be:

$$
\begin{aligned}
F_{tot}(k) &= \sum_{k'=1}^{k} [F_{pf}(k') + F_{ts}(k') + F_{up}(k')] \\
R(k) &= F_{tot}(k)/T_{tot}^{acc}
\end{aligned}
\tag{6}
$$

where T(k) is the total time accumulated for all k□ <= k. If in equation (6) the sum is for k = k_{max} we obtain again R_{max} of equation (4).

Experimental observations. We have choosen to monitor and analyze the evolution of the flops rate as function of the iterations by timing the main steps. That is for a single HPL tests we measure: T_{pf} (k), the time to complete the panel factorization at k iteration, T_{bc}(k), the time for the panel broadcast to the row processes and T_{up}(k) the time for the update step. The time T^{acc}_{pf} (k) = $\sum_{k'=1}^{k}$ T_{pf} (k') indicates the total time of panel factorization accumulated for all k□ <= k and similarly for T^{acc}_{bc} (k) and T^{acc}_{up} (k). Note that T^{acc}_{tot} (k_{max}) is equivalent to the total execution time T_p used in equation 4. The timing is made in parallel by all processes but it is possible to analyze a single output because of the parallel nature of the algorithm. This also allows to have a process independent treatment. There is just one limitation on the minimum valid iteration from which the execution time per iteration can be considered equal for all processes. The panel factorization and broadcast steps are made by a single processes column, so computing is not balanced in the first iteration. The process with rank = 0 (or any other rank one may want to monitor) will have a peak in the broadcast step at k = 2, 2 +Q, 2 +2Q, ... (e.g. see *Figure 3* that reports the case p = 64 and P × Q = 2 × 32), where Q is number of row processes in the grid. The recursive formulation entails a periodicity in the main computational steps. We have checked that from the second or the third iteration the processes tend to synchronize (see *Figure 2*) and that this property is independent by the number of processes. By such considerations k = 3 will be the minimum iteration that we consider when monitoring execution times of a single process. Most of our experiments were performed using a row-major map for the cyclic distribution data and the ring modified broadcast topology.

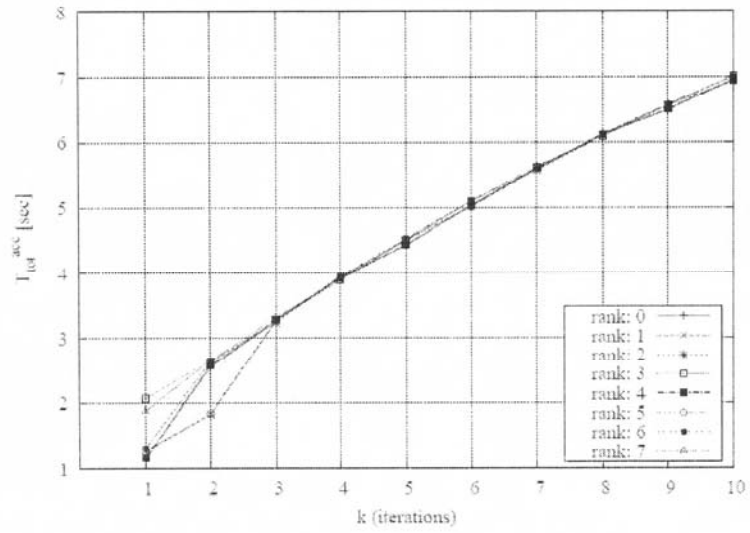


Figure 2: *It is reported the total accumulated (measured) time as function of iterations k ($T^{acc}_{tot}(k) = T^{acc}_{pf}(k)+T^{acc}_{bc}(k)+T^{acc}_{up}(k)$). After two or three iterations all processes (in the graph the number of the processes is 8) measure roughly the same accumulated time.*

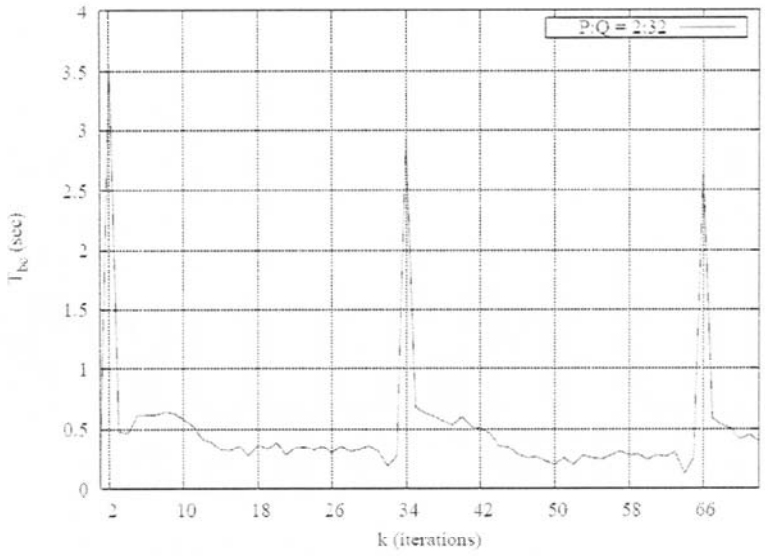


Figure 3: *The figure shows the peaks for the broadcast time (T_{bc}) as function of the iterations k for the process p0 with rank 0. In this example $P \times Q = 2 \times 32$ and a ring modified broadcast topology is used. For $k = 1$, p_0 will send a message and proceed with calculations, while at $k = 2, 2 + Q, 2 + 2Q, \ldots$ it has to receive and send messages. Because the HPL algorithm, the panel factorization is executed by P processes. In the broadcast phase the P processes send the owned portion of the panel to the Q row processes.*

For that reasons is better to stop execution for a $k = k_{min}$ such that:

$$2 < K_{\min} < 2+Q \qquad (7)$$

where the factor 2 is due to the broadcast topology. The flops rate can be seen as a mean over the iterations weighted with respect to the number of the flops. Another important observation concerns the evolution of the ratio T^{acc}_{bc} (k)/T^{acc}_{tot} (k) and T^{acc}_{up} (k)/T^{acc}_{tot} (k). *Figure 4* shows the evolution of the times measured as function of k. It shows that the above ratio is roughly constant and that supports our main assumption.

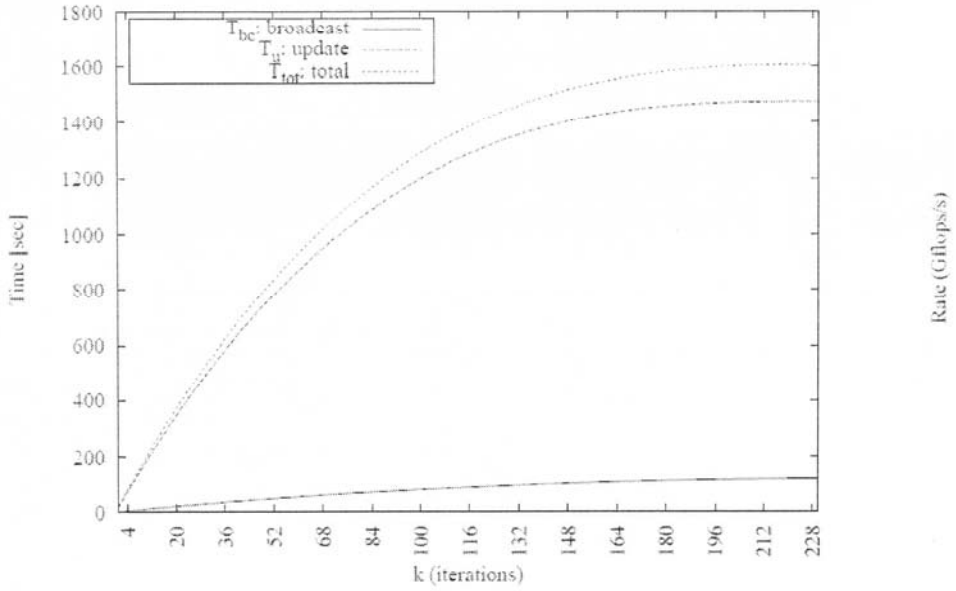


***Figure 4**: At k-th iteration the ratios T^{acc}_{bc} (k)/T^{acc}_{tot} (k) and T^{acc}_{up} (k)/T^{acc}_{tot} (k) are roughly constant. The figure reports the total accumulated time in all of the steps as function of iterations k. We have checked that a such pattern is independent by problem size and number of processes. In the figure is not reported the fraction of time for the panel factorization, that is T^{acc}_{pf} (k)/T^{acc}_{tot} (k).*

Results and Methodology Applications

In this section we show some numerical results for a typical optimal parameters set up. Our method allows to save time in the search of optimal configurations but also gives a better understanding of performances evolution. We use the equations (5) and (6) to build the curves of flops rate versus the iteration index and detect the best rate by stopping the execution flow at the beginning of calculation (following the condition of equation (7)). We also run the tests for all iterations to check the results. In sub sect. (5.1) we analyze some simple HPL tests, while sect. (5.2) reports a practical application of our method.

Analysis of the flops rate at the beginning of the Iterations. In this subsection we analyze some simple tests to show the behavior of flops rate at the beginning of the iterations. The analysis is useful for a better understanding of the relations between memory occupation, block size bs and communication cost. We treat three cases for p = 32 and P × Q = 4 × 8 on Sect.2 cluster :

- N/bs fixed: changing N to allocate 10, 20, 40, 80% of memory per node and varying bs.
- bs fixed: N is changed as previous case but bs is kept at fixed value.
- N fixed: only bs is changed.

The typical slope of flops execution rate curve versus the iteration index is like one of those represented in *Figure 5*: for smalls k it increases rapidly, reaches a maximum and then

decreases slowly. This behavior has to be related with the ratio T_{bc}/T_{tot} and the block size. *Figure 5* and *Figure 6* show the flops rate as function of iterations for $k = k_{max}$ and $k_{min} = 9$ respectively. By looking at Fig. 6 we note that when bs increase with the problem size the slope of the curve is smoother. The explanation is that a small bs increments the intrinsic cache misses in the update step and, as we observed numerically, the cost communication is bigger than in the rest of iterations. We observe also that T_{bc}/T_{tot} depends mainly (not only) by the memory occupation. This is confirmed by another experiment where we keep bs fixed and solves the same linear system (with the same N) as the previous case. *Figure 7* represents the rate for the bs fixed case. Tests stopped at $_{min} = 9$ give the same set of curves as Fig. 6. There are no qualitative differences and, as before we observe that T_{bc}/T_{tot} depends by the memory allocation. The maximum is reached when the ratio T_{bc}/T_{tot} is roughly constant. The slowly decrease in flops rate after the maximum is because T_{bc}/T_{tot} roughly constant and the load imbalance is proportional to bs. This is clear when it is fixed the problem size and let vary bs.

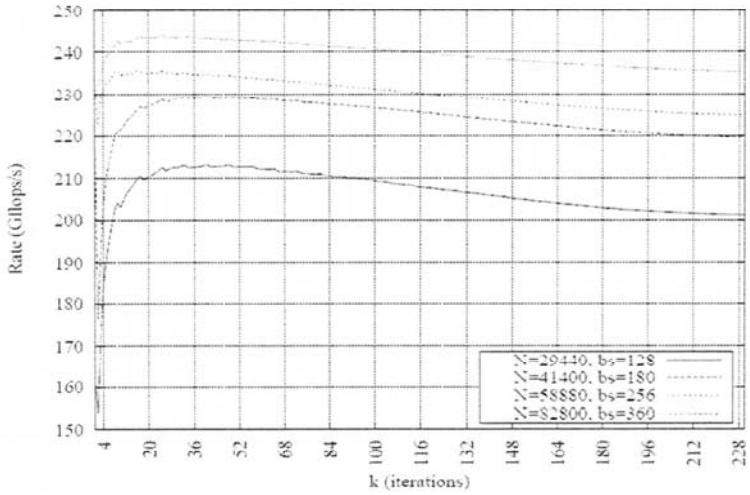


Figure 5: *The figure shows the trend of flops execution rate (Rate is expressed in Gflops/s) for all iterations k. Are reported various problem size and the block size is chosen to have N/bs constant. The results were obtained using $p = 32$ cores and $P \times Q = 4 \times 8$ on Sect.2 cluster, although analogous behaviour of flops rate is found in Sect.1 cluster and AMD cluster.*

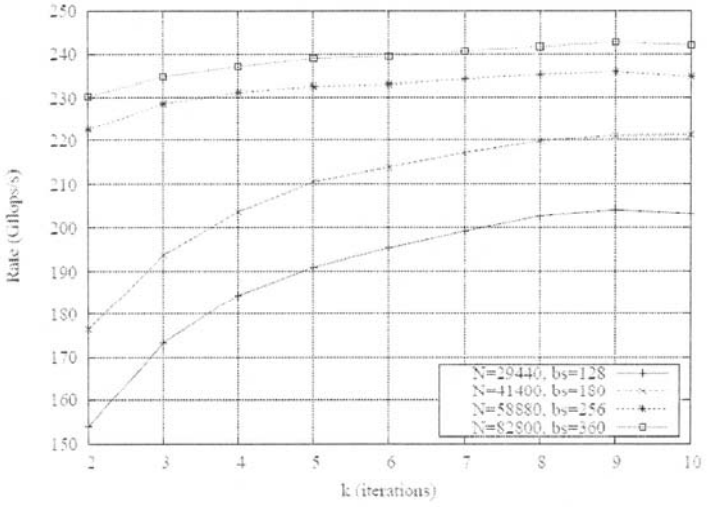


Figure 6: *The same HPL tests as Fig. 5 but runs are cut-off at $k_{min} = 9$ (see condition of eq. (7)). The broadcast topology used is the ring modified.*

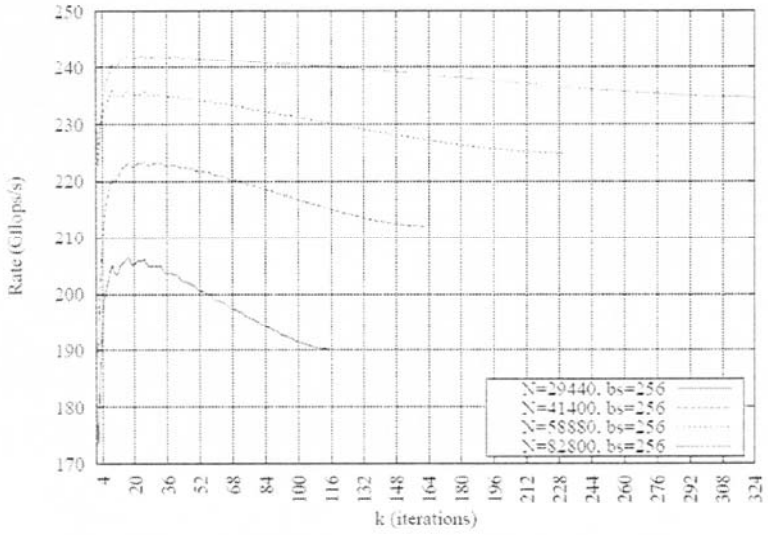


Figure 7: *The same tests as Fig. 5 respect N but keeping bs = 256 constant. For k_{min} = 9 we obtained the same qualitative trend as reported in Fig. 6. The differences after the maximun value are because load imbalance.*

As shown in *Figure 8* the smaller bs will give the smoother slope. However in the worst case between the maximum rate and the final rate there is a difference of about 3 – 5%. We have checked our results for the cases p = 128 and p = 1024 in the Sect.2 cluster and the cases p = 16 and p = 32 for the Sect.2 and AMD clusters: the qualitative behaviour of the flops rate is kept. Finally, we note that the results and the flops rate curves are obtained by monitoring the first process. An analogous treatment can be made for a process with generic rank, provided that the stopping criterium T_{bc}/T_{tot} of equation (7) is modified according with the process rank and the broadcast topology.

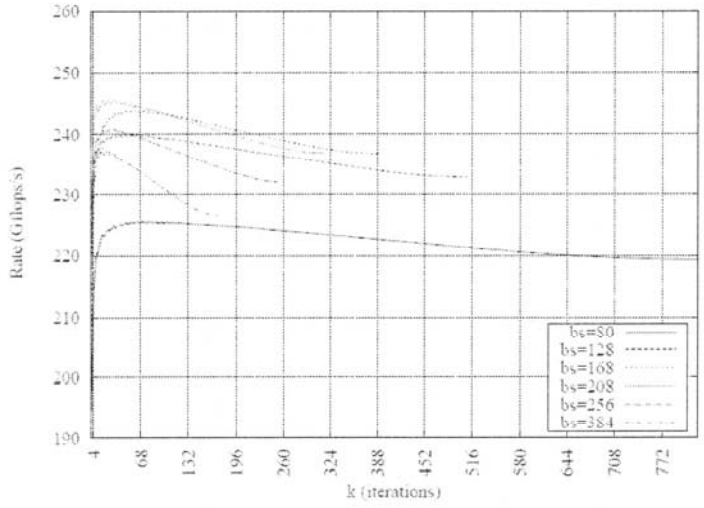


Figure 8: *The rank of flops rate at beginning of iterations are the same as the completed tests reported in figure.*

Example of the Aapplication of the Method

As an example of the application we have determined, for a given problem size N, the optimal bs, processors grid P × Q and the broadcast topology (bcast) for a system of p = 64 processors. We consider the previous parameters to show the time we saved in the HPL tests by stopping the calculations at k given by the condition (7). Once we have the parameters from all of the three group of tests, we run the best configuration for all k. Our examples tests were using p = 64 cores of Sez.2 cluster and a fixed problem size with N = 82440 (the memory allocation is about 40% in this case). We first try bs = 80, 128, 208, 384 and P × Q = 8 × 8. Timings and flops rate are summarized in *Table 1*. It can be seen at $k_{min} = 9$ that the best bs is 208. The table reports also the results for the complete run of all the cases, so that an estimation of the saved time can be performed.

In the second group of tests we try all combinations of processors grid: P × Q = 2 × 32, 4 × 16, 8 × 8, 16 × 4, 32 × 2. The *Table 2* reports times and flops rates.

bs	$T_{tot}^{acc}(k_{min})$	$R(k_{min})$	$T_{tot}^{acc}(k_{max})$	$R(k_{max})$
80	25.4	381.9	883.3	422.8
128	35.0	441.7	827.9	451.2
208	53.2	466.4	808.2	462.2
384	96.4	463.1	842.8	443.2

Table 1: *In this group of tests we let vary just the block size bs. The execution flow was stopped k_{min} at = 9 that is a consequence of condition (7). Times T^{acc}_{tot} and flops rate R are expressed in seconds and Gflops/s respectively.*

$P \times Q$	$T_{tot}^{acc}(k_{min})$	$R(k_{min})$	$T_{tot}^{acc}(k_{max})$	$R(k_{max})$
2 × 32	154.2	456.7	871.9	428.4
4 × 16	79.2	473.5	810.5	461.0
8 × 8	44.4	454.2	812.8	459.6
16 × 4	27.8	406.8	902.1	414.1
32 × 2	19.7	346.3	1098.9	339.9

Table 2: *With p = 64 we try various processors grids. The execution flow was stopped at kmin following the condition (7). For example for the test with P × Q = 4 × 16 we have $k_{min} = 17$.*

In the third group of tests we try all of five broadcast topologies that HPL allow to use: bcast = 0, 1, 2, 3, 4, 5. See *Table 3* for the results. The total number of run tests is 15. The three groups above give as the following best (likely) parameters: bs = 208, P × Q = 4 × 16 and bcast = 1. Running entirely that test requires: Tp = 808.0 sec and returns Rmax = 462.5 Gflops/s. The total time to run 15 tests for k = k_{min} and one test for k = k_{max} is about 1610 sec. If we have to run all of the 15 tests for k = k_{max} total time would be 12832 sec. So we saved the 87.4% of the time. It is not possible to give a general rule for the fraction of time saved because it depends on the number of group of tests. However, when considering many tests on many thousands of cores, our method allows to find the best HPL configuration with a speed-up of 8–10 with respect to the standard methods.

$bcast$	$T_{tot}^{acc}(k_{min})$	$R(k_{min})$	$T_{tot}^{acc}(k_{max})$	$R(k_{max})$
0	43.6	463.1	819.4	456.0
1	43.4	464.6	812.6	459.8
2	44.6	452.3	823.9	453.5
3	43.9	459.2	818.7	456.4
4	46.1	437.6	863.3	432.8
5	46.1	438.1	836.1	446.9

Table 3: *HPL allows to select the broadcast topology. In this group of tests we try all of them. Because here we set* $P \times Q = 8 \times 8$, $k_{min} = 9$.

Summary and Discussion

The problem of performances evaluation is becoming increasingly important for programmers, computer scientists and computers vendors. Running and tuning the HPL program requires a lot of time when the number of processors and the problem size increases. In this paper we have presented a new method to detect the best HPL parameters for a given computer system. Our approach allows to stop the program after a few iterations and decide which is the best configuration among a group of tests. In this way it is possible to save about 80–90% of the wall clock time. The main assumptions are supported by the reported numerical experiments showing the application of the method. Although our target is not a precise prediction of the final flops rate, the examples show that the flops rate at $k = k_{min}$ differs from the R_{max} by only 0.5–3%. This is a further confirmation of the validity of our method and opens the prospective for a more accurate prediction of the HPL performances at the beginning of calculations. To achieve that goal it is required to analyse and understand the slope of flops rate curves as a function of the iteration index. In future works we also want to extend and check the method for almost all HPL input parameters. The presented results show also that the peak performances can be be quite accurately predicted opening the possibility to implement a complete and accurate framework that rapidly allows to obtain the best performances of a given parallel system. Finally, another important consequence would be the possibility to include in the future Linear Algebra libraries an automated detection of the performances during the computation. This can be used to stop inefficient scientific applications that makes use Linear Algebra software and detect, for example, software or system faults.

REFERENCES

[1] Petitet, A., Whaley, R.C., Dongarra, J.J., Cleary, A. "*Hpl: A portable implementation of the high-performance linpack benchmark for distributed-memory computers*". Innovative Computing Laboratory, Available online (2000). URL http://icl.cs.utk.edu/hpl/ and http://www.netlib.org/benchmark/hpl/.

[2] Dongarra, J., Luszczek, P., Petitet, A. "*The linpack benchmark: past, present and future. Concurrency and Computation*". Practice and Experience 15(9), 803–820 (2003).

[3] Feitelson, G. "*On the interpretation of top500 data*". Intl. J. High Performance Comput. Applications 13(2), 146–153 (1999).

[4] Gustafson, J.L., Todi, R. "*Conventional benchmarks as a sample of the performance spectrum*". The Journal of Supercomputing 13(3), 321–342 (1999).

[5] Luszczek, P., Dongarra, J., Kepner, J. "*Design and implementation of the hpc challenge benchmark suite*". CTWatch Quarterly 2(4A) (2006).

[6] Bailey, D., Barscz, E., Barton, J. et al. "*Nas parallel benchmarks*". NAS Technical Report RNR-94-007 (1994).

[7] Dongarra, J., van de Geijn, R.A., Walker, D.W. "*Scalability issues affecting the design of a dense linear algebra library*". J. Parallel Distrib. Comput. 22(3), 523–537 (1994).

[8] Gustavson, F.G. "*Recursion leads to automatic variable blocking for dense linear-algebra algorithms*". IBM Journal of Research and Development 41(6), 737–756 (1997).

[9] Peterson, K. "*Cray scaling efforts to reach a petaflop*". In: CUG. Atlanta (2009).

[10] Feitelson, D.G. "*The supercomputer industry in light of the top500 data*". Comput. in Science Engineering 7(1), 42–47 (2005).

[11] Petitet, A.P., J.Dongarra, J. "*Algorithmic redistribution methods for block-cyclic decompositions*". IEEE Transactions on Parallel and Distributed Systems 10(12), 201–220 (1999).

[12] Dongarra, J., Eijkhout, V., Luszczek, P. "*Recursive approach in sparse matrix LU factorization*". Scientific Programming 9(1), 51–60 (2001).

[13] Agarwal, R.C., Gustavson, F.G., Zubair, M. "*A high performance matrix-multiplication algorithm on a distributed-memory parallel computer, using overlapped communication*". IBM Journal of Research and Development 38(6) (1994).

[14] Migliori, S., et al. "*CRESCO project: Computational RESearch center for COmplex systems*". ENEA: Italian National Agency for New Technologies, Energy and Sustainable Economic Development (2007). URL http://www.cresco.enea.it/.

[15] Migliori, S., Bracco, G., D'Angelo et al. "*Architecture and performances of the CRESCO HPC system*". International Supercomputing Conference. Dresden (2008).

[16] Toledo, S. "*Locality of reference in LU decomposition with partial pivoting*". SIAM Journal on Matrix Analysis and Applications 18(4), 1065–1081 (1997).

Edito dall'ENEA
Unità Comunicazione

Stampato presso il Laboratorio Tecnografico ENEA - C.R. Frascati
Finito di stampare nel mese di giugno 2011