

UNIVERSITA' degli STUDI di NAPOLI FEDERICO II  
DIPARTIMENTO DI MATEMATICA E APPLICAZIONI "RENATO CACCIOPPOLI"

RICHIESTA DI FINANZIAMENTO DI UN PROGETTO DI RICERCA DIPARTIMENTALE A VALERE SUI FONDI  
"RICERCA DIPARTIMENTALE 2015" – DD 25 del 19/3/2015

**Titolo del progetto:** Approcci Innovativi per la risoluzione di **Modelli di Interesse** nelle simulazioni **Computazionali (AMICo)**

**Responsabile scientifico:** Luisa D'Amore.

**Partecipanti.** P. Festa, G. Laccetti, M. Lapegna, M. Longobardi, E. Pirozzi,

**Assegnisti:** Valeria Mele.

**Studenti di dottorato:** Simone Celestino, Giuseppe D'Onofrio, Daniele Ferone, Antonio Napoletano, Giuseppe Vettigli.

Settore ERC: PE1, PE6

### Descrizione proposta progettuale

#### a. stato dell'arte

La Computational Science and Engineering (CSE) è un campo multidisciplinare di ricerca, che si avvale di esperienze e competenze di matematica applicata, computer science e delle discipline delle scienze e dell'ingegneria. Suo principale scopo è quello di sviluppare e utilizzare metodi computazionali per l'avanzamento dell'innovazione scientifica e tecnologica. La CSE è attualmente uno dei maggiori campi di intervento per la ricerca e l'alta formazione universitaria ed è basata, come pilastro fondamentale, sull'High Performance Computing (HPC) rivestendo un ruolo centrale nella "data revolution". In questo ambito si colloca il Progetto **AMICo**.

Più precisamente, i partecipanti al progetto **AMICo** svolgono la loro attività di ricerca nel contesto scientifico della modellistica e simulazione computazionale di sistemi complessi (ovvero multi-scala e/o multi- physics e/o di grandi dimensioni) e della risoluzione di tali problemi su sistemi di calcolo ad alte prestazioni, anche in collaborazione con Università e Centri di Ricerca di rilevanza internazionale; in particolare in questa proposta progettuale l'attenzione sarà rivolta verso **modelli climatici, biologici, neuronali**.

#### b. obiettivi del progetto

L'attività della presente proposta sarà articolata attraverso l'individuazione di opportuni **casi di studio (testbed)** sui quali sarà studiata la sperimentazione e validazione di **approcci innovativi** per la loro risoluzione computazionale. In particolare, i casi di studio su cui si prevede di rivolgere l'attenzione, in riferimento rispettivamente a ciascuno dei sistemi complessi precedentemente citati sono:

1. Simulazione della circolazione oceanica su scala locale in mari chiusi e semi-chiusi (Mar Caspio e Mar Mediterraneo)[Shiller2011]. Attività in collaborazione con il Centro Euro Mediterraneo sui Cambiamenti Climatici, (IT), Imperial College London (UK), University of California (USA), Science & Technology Facilities Council (Daresbury), (UK).
2. Problemi di ottimizzazione combinatoria fra cui i cosiddetti problemi di consenso ed in particolare il Far From Most String Problem, problemi di clusterizzazione e di bi-clusterizzazione di grandi moli di dati e i problemi di Feature Selection, ai fini della creazione di sonde diagnostiche per le infezioni batteriche o individuazione di potenziali bersagli farmacologici [Blum2014]. Attività in collaborazione con il Department of Computer Science and Artificial Intelligence, University of the Basque Country (Spagna), l'AT&T Labs Research, NJ (USA), l'Istituto di Analisi dei Sistemi ed Informatica -- Consiglio Nazionale delle Ricerche (Roma) e l'Istituto di Scienze dell'Alimentazione -- Consiglio Nazionale delle Ricerche (Avellino).
3. Processi stocastici ristretti da barriere assorbenti e/o riflettenti, temporalmente non omogenei [Higham2001,Taillefumier2013]. Attività in collaborazione con il Laboratorio di Fisica Matematica della Rockefeller University di New York (USA),
4. Misura della mutua informazione tra le componenti estreme di campioni casuali descrittivi tempi di vita, condizionate da osservazioni compiute ad istanti differenti [Longobardi2015a]. Attività in collaborazione con Ferdowsi University of Mashhad (Iran), Rajiv Gandhi Institute of Petroleum Technology, Rae Bareilly, (India) e University of Murcia (Spagna).

L'innovazione risiede nella metodologia introdotta e/o nel progetto dell'algoritmo e/o nelle scelte implementative adottate oramai indispensabili ai fini della realizzazione di strumenti computazionali affidabili e in grado di sfruttare appieno l'enorme potenza di calcolo che sarà disponibile nei prossimi anni (Exascale Computing, linee guida della Ricerca e Innovazione, Horizon 2020) [D'Amore2014a, Dongarra2014, Laccetti2014a].

#### c. metodologia

Si prevede la sperimentazione e validazione delle metodologie numeriche di Data Assimilation (DA), nella sua formulazione Variazionale tri-dimensionale e quattro dimensionale nei modelli di simulazione della circolazione oceanica. Se da una lato la DA consente di migliorare l'affidabilità dei modelli stessi dall'altro aumenta in maniera significativa la complessità dell'intera simulazione computazionale, pertanto nasce l'esigenza di introdurre paradigmi di parallelismo multilivello nei relativi algoritmi risolutivi [D'Amore2014b]. Si prevede inoltre la sperimentazione di algoritmi di simulazione stocastica di modelli neuronali, al fine di investigare e validare proprietà della densità del

tempo di primo passaggio per processi di diffusione temporalmente non omogenei e per i quali risultati teorici e simulativi sia in regime transiente che asintotico e in presenza di barriere assorbenti e/o riflettenti sono carenti e frammentari [Pirozzi2010, Pirozzi2014]. Per risolvere istanze reali dei problemi oggetto di studio nei modelli biologici, si vuole perfezionare la dimostrazione (formale e sperimentale) della appartenenza dei problemi di ottimizzazione combinatoria di interesse nei modelli biologici, alla classe dei problemi computazionalmente intrattabili al fine di progettare nuovi e più efficienti metodi di soluzione che ne sfruttino le proprietà e caratteristiche peculiari in termini di complessità computazionale [Festa2012, Festa2013, Ferone2013]. Infine, per la determinazione di caratteristiche della mutua informazione indipendenti dalle distribuzioni sottostanti si intende sviluppare procedure numeriche basate sulla rappresentazione mediante copule [Longobardi2015b]. Ai fini della loro implementazione efficiente su architetture emergenti, gli algoritmi sviluppati dovranno tenere conto dei differenti "livelli architetturali" dei sistemi di calcolo su cui verranno implementati e sperimentati. Si pensi ad esempio ad architetture multiprocessore-multicore, più nodi connessi tra loro, ciascun nodo composto da più cores (computing cores, acceleratori grafici, ecc ..). Ciò porta alla progettazione di algoritmi (e software) di tipo "ibrido", che sappiano sfruttare le diverse potenzialità e caratteristiche dei "diversi" livelli. [Laccetti2014b, Laccetti2015]. Grande importanza inoltre riveste il problema della gestione di grandi quantità di dati (Big Data) in ambito High Performance Computing [Laccetti2014a]. Un tale approccio alla implementazione e sperimentazione si avvarrà della collaborazione con l'Argonne National Laboratory del Department of Energy degli USA, e con la University of Chicago.

#### d. risultati e prodotti attesi

Nel campo delle scienze computazionali il principale risultato atteso - a lungo termine - è la realizzazione di prodotti software effettivamente utilizzabile nelle simulazioni scientifiche. Tuttavia, poiché nessun singolo ricercatore ha le competenze necessarie per far fronte a tutto il dominio di esperienza (dall'analisi dei modelli, alla progettazione di nuovi algoritmi, fino alla loro implementazione in modo efficiente sulle architetture di calcolo, e all'analisi dei risultati, ...), questa proposta progettuale si pone come risultato lo sviluppo di attività inter - disciplinari essenziali per l'avanzamento della conoscenza nelle simulazione di sistemi complessi. Ad esempio, i risultati ottenuti nei vari casi di studio, anche in itinere, potranno essere immediatamente sfruttati nell'ambito del progetto stesso, migliorando o integrando il progetto e lo sviluppo degli algoritmi. Ciò sarà possibile anche con la partecipazione a convegni, l'organizzazione di eventi e seminari, e tenendo anche conto della collaborazione scientifica già in atto con Università ed Enti di Ricerca nazionali e internazionali, attraverso la mobilità dei (alcuni) partecipanti al progetto ed in particolare si prevede che l'assegnista di ricerca si rechi per un mese presso l'STFC ( Science & Technology Facilities Council), Daresbury , (UK), (N.B. per almeno 2-3 dottorandi, nell'ambito del progetto, si prevede un periodo di mobilità; tuttavia tale spesa non graverà sul finanziamento del progetto). Si prevede altresì la pubblicazione di articoli scientifici su riviste internazionali e la presentazione di risultati in conferenze internazionali.

#### Bibliografia

- [Blum2014] Blum C., **Festa P.**, A Hybrid Ant Colony Optimization Algorithm for the Far From Most String Problem, Lecture Notes in Computer Science, Vol. 8600, Springer-Verlag, Berlin Heidelberg, 2014, pag. 1-12.
- [D'Amore2013] **D'Amore L.**, Arcucci R., Marcellino L., Murli A., HPC computation issues of the incremental 3D variational data assimilation scheme in OceanVar software, Journal of Numerical Analysis, Industrial and Applied Mathematics (JNAIAM), vol. 7, n. 3-4, 2013, pag. 91-105.
- [D'Amore2014a] **D'Amore L.**, Murli A., Boccia V., Carracciolo L., Insertion of PETSc in the nEMO stack software. Driving NEMO towards Exascale Computing, IEEE Proceedings of the International Conference on High Performance Computing & Simulation 2014, pa. 724-731, 2014.
- [D'Amore2014b] **D'Amore L.**, Arcucci R., Carracciolo L., Murli A., A Scalable Approach for Variational Data Assimilation, Journal of Scientific Computing, vol. 61, n. 2, 2014, pag. 239-257.
- [Dongarra2014] Dongarra J., Reed D. A., Exascale Computing and Big Data: The Next Frontier, Comm. of the ACM, 2014 (submitted).
- [Ferone2013] Ferone D., **Festa P.**, Resende M.G.C., Hybrid Metaheuristics for the Far From Most String Problem, Lecture Notes in Computer Science, vol. 7919, Springer-Verlag, Berlin Heidelberg, 2013, pag. 174-188.
- [Festa2012] **Festa P.**, Pardalos P.M., Efficient solutions for the Far From Most String Problem, Annals of Operations Research, vol. 196, n. 1, 2012, pag. 663-682.
- [Festa2013] **Festa P.**, A biased random-key genetic algorithm for data clustering, Mathematical Biosciences, vol. 245, n. 1, 2013, pag. 76-85.
- [Higham2001] Higham D.J., An algorithmic introduction to numerical simulation of stochastic differential equations. SIAM Review vol. 43, n. 3, 2001, pag. 525-546.
- [Laccetti2014a] **Laccetti G.** et al., Virtualizing high-end GPGPUs on ARM clusters for the next generation of high performance cloud computing, Cluster Computing, vol. 17, n. 1, 2014, pag. 139-152.
- [Laccetti2014b] **Laccetti G., Lapegna M.** et al., A study on Adaptive Algorithms for Numerical Quadrature on Hybrid GPU and Multicore Based System, Lecture Notes in Computer Science, vol. 8384, Springer, 2014, pag. 704-713.

- [Laccetti2015] **Laccetti G., Lapegna M.** et al., A loosely coordinated model for heap-based priority queues in multicore environments, *International J. of Parallel Programming*, 2015 (in stampa).
- [Longobardi2014] **Longobardi, M.** Cumulative measures of information and stochastic orders. *Ricerche Mat.* 63 (2014), pag. 209-223.
- [Longobardi2015a] Di Crescenzo, A. and **Longobardi, M.**, Some properties and applications of cumulative Kullback-Leibler information *Applied Stochastic Models in Business and Industry*, to appear.
- [Longobardi2015b] Ahmadi, J., Di Crescenzo, A. and **Longobardi, M.**(2015) On dynamic mutual information for bivariate lifetimes, preprint arXiv:1411.6257
- [Pirozzi2010] Buonocore A., Caputo L., **Pirozzi E.** , Ricciardi L.M., On a stochastic leaky integrate-and-fire neuronal model, *Neural Computation*, vol. 22, n. 10, 2010, pag. 2558–2585.
- [Pirozzi2014] Buonocore A., Caputo L., Nobile A.G., **Pirozzi E.**, Gauss-Markov processes in the presence of a reflecting boundary and some applications in neuronal models, *Appl. Math. Comput.*, vol. 232, 2014, pag. 799–809.
- [Shiller2011] Schiller A., Brassington G.B. (eds.), *Operational Oceanography in the 21st Century*, Springer, 2011.
- [Taillefumier2013] Taillefumier T., Magnasco M. O., A phase transition in the first passage of a Brownian process through a fluctuating boundary with implications for neural coding, *Proceedings of the National Academy of Sciences of the United States of the America (PNAS)*, vol. 110, n. 16, 2013, pag. E1438–E1443.