Part II Courses

Module 2.1:	**Data Structures and Sorting & Searching**
Lecturer	Axel Kohlmeyer (Temple U. & ICTP)
Class Duration	10 hours
Laboratory Duration	24 hours
Module Description	Introduction to fundamental data structures and their impact on performance and memory consumption. Study parallelization issues.
Main Topics	● Implementation of data structures like arrays, linked lists, hash tables, trees, maps in different programming languages ● Comparison of performance for data access, data insert, data removal, data append, scaling with problem size, and need for auxiliary storage ● Accessing and updating data structures in parallel (using locks and lock-free) ● Data structures for numerical problems (example: sparse matrix storage) ● Some Considerations on Data Structure Design ● Implementation of selected popular sort algorithms ● Comparison of performance for unsorted and presorted data ● Searching in unsorted or presorted data
Objectives	On successful completion of this module students should be able to determine which data structure is best suited for a given task.

Module 2.2:	CUDA Fortran for material scientists
Lecturer	ICTP school (http://indico.ictp.it/event/8150/)
Class Duration	12 hours
Laboratory Duration	12 hours
Module Description	Introduce the audience to GPU and CUDA FORTRAN WHILE the main focus is hands-on and lab demonstrations about how to use CUDA FORTRAN with a large FORTRAN codes
Main Topics	● CUDA FORTRAN
Objectives	On successful completion of this module students should be able to write code in CUDA FORTRAN.

Module 2.3:	**Cluster Analysis**
Lecturer	Alex Rodriguez (SISSA)
Class Duration	12 hours
Laboratory Duration	12 hours
Module Description	Theory and applications of Clustering algorithms.
Main Topics	● Motivation for Clustering ● Similarities and Distances ● Flat, fuzzy and Hierarchical clustering methods ● Clustering methods examples ● External and Internal Validation ● Clustering applications
Objectives	The students should be able to implement a clustering algorithm method and to choose the one that fits better the problem that they want to solve.

Module 2.4:	**Spatial locality algorithms**
Lecturer	Riccardo Valdarnini (SISSA)
Class Duration	8 hours
Laboratory Duration	16 hours
Module Description	Theory and applications of algorithms for spatial locality
Main Topics	● Space filling curves-theory ● Morton and Peano-Hilbert orders ● Tree codes, general framework ● Quad tree, R- tree, Kd-tree ● Nearest neighbor search using space filling curves.
Objectives	Upon completion students will be able to deal with spatial locality algorithms.

Module 2.5:	**Reduced Basis Method**
Lecturer	Gianluigi Rozza (SISSA)
Class Duration	12 hours
Laboratory Duration	12 hours
Module Description	In this course we present reduced basis (RB) approximation and associated a posteriori error estimation for rapid and reliable solution of parametrized partial differential equations (PDEs).
Main Topics	● Introduction to RB methods, offline-online computing, elliptic coercive affine problems ● Sampling, greedy algorithm, POD ● A posteriori error bounds ● Primal-Dual Approximation ● Time dependent problems: POD-greedy sampling ● Non-coercive problems ● Approximation of coercivity and inf-sup parametrized constants ● Geometrical parametrization ● Reference worked problems ● Examples of Applications in CFD
Objectives	On successful completion of this module students should know the basic aspects of numerical approximation and efficient solution of parametrized PDEs for computational mechanics problems (heat and mass transfer, linear elasticity, viscous and potential flows).

Module 2.6:	**Parallel Linear Algebra**
Lecturer	Francesco Sanfilippo (La Sapienza - Rome)
Class Duration	16 hours
Laboratory Duration	16 hours
Module Description	Introduction to the solution of large linear systems in parallel
Main Topics	• Overview of real-world occurrences of the problem • Comparison of direct and iterative solvers • Preconditioning of a problem • Accuracy of the solution, efficiency of different solvers • Extensive discussion on implementation of a solver in massive parallel architectures
Objectives	Upon completion students will be able to implement a parallel conjugate gradient with different level of parallelism

Module 2.7:	**Monte Carlo method**
Lecturer	Stefano Baroni (SISSA) Federico Becca (SISSA) Pietro Delugas (SISSA) Roberto Innocente (SISSA)
Class Duration	20 hours
Laboratory Duration	20 hours
Module Description	Theory and applications of the Monte Carlo methods. Hands-on with examples, analysis of simulations and parallelization
Main Topics	• Review of Probability theory • Pseudo-random numbers generators, Parallel RNG (SPRNG) • Parallel Libraries for random generators • Sampling multi-variate Gaussian deviates • Importance sampling • Stochastic processes • Metropolis Monte Carlo • Langevin dynamics • Introduction to Quantum Monte Carlo
Objectives	Upon completion students will be able to optimize software for the general application of the Monte Carlo Method on parallel systems having a good knowledge of the most advanced library for RNG.

Module 2.8:	**Fast Fourier Transforms in Parallel and Multiple Dimensions**
Lecturer	Ralph Gebauer (ICTP) Ivan Girotto (ICTP)
Class Duration	8 hours
Laboratory Duration	12 hours
Module Description	Introduction to the Discrete Fourier Transform (DFT) and its application to real problems. From the Discrete to the "Fast" version (FFT). Analysis of a most common algorithm for the solution of a multi-dimensional FFT on parallel systems
Main Topics	● Discrete and Continuous Fourier Transform ● The Cooley-Tukey FFT algorithm ● Implementation of an MPI parallel multi-dimensional FFT based on a 1d FFT ● Example use in a diffusion problem
Objectives	On successful completion of this module students should be able to integrate parallel FFTs into applications.

Module 2.9:	**Approximation and interpolation of simple and complex functions**
Lecturer	Nicola Seriani (ICTP)
Class Duration	12 hours
Laboratory Duration	12 hours
Module Description	Introduction to several techniques for efficient approximation of numerical functions to varying degrees of accuracy
Main Topics	• Interpolation (linear, spline) and errors (number and spacing of interpolation points) • Cost of interpolation vs. explicit function evaluation of complex functions • Approximation (taylor/maclaurin, pade). • Lookup plus newton-raphson • Range reduction as an efficient way to approximate exp(), log(),sin(), cos() and others in combination with a spline table or pade approximation for a small interval. • Floating point math tricks
Objectives	Upon completion of this module students will be able to implement efficient approximations and tabulations of numerical functions and determine the accuracy of the approximations.

Module 2.10:	**Molecular Dynamics**
Lecturer	Giovanni Bussi (SISSA) Sabine Reisser (SISSA)
Class Duration	6 hours
Laboratory Duration	12 hours
Module Description	Theory and applications of molecular dynamics simulations."Hands-on exercises on neighbor-list, linked cells, and parallelization strategies
Main Topics	● Introduction to molecular dynamics ● Neighbor lists and linked cells ● Parallelization using MPI ● Implementation of multi-replica algorithms (parallel tempering)
Objectives	Upon completion students will be able to run molecular dynamics simulations of a Lennard-Jones system and to optimize and parallelize a molecular dynamics code. Additionally, they will be able to implement parallel tempering algorithms.

Module 2.11:	**Supervised Machine learning**
Lecturer	Valerio Consorti (Generali)
Class Duration	12 hours
Laboratory Duration	12 hours
Module Description	Supervised Machine Learning

Main Topics	● The data analytics process
	● Regression problems
	  ○ Linear regression
	  ○ Overfitting
	● Classification problems
	  ○ Logistic regression
	  ○ Decision trees
	  ○ Quality measures
	  ○ Class unbalancing
	● Basic methodological approach to the Natural Language Processing
	● Ensemble models:
	  ○ bagging
	  ○ random forest
	  ○ boosting

Objectives	● To be able to design and implement a simple end-to-end data analytics process ● To understand the various logical steps that constitute a typical data analytics process ● To be able to build the proper set of variable to describe the data ● To choose the proper model to solve the specific problem ● To be able to properly evaluate the performance of the analysis ● To become familiar with the python package scikit-learn

	● To be able to implement custom transformers and estimators ● To be able to implement complex multi-step analysis

Module 2.12:	**Electronic structure: from blackboard to source code**
Lecturer	Stefano de Gironcoli (SISSA)
Class Duration	16 hours
Laboratory Duration	16 hours
Module Description	
Main Topics	
Objectives	

Module 2.13:	**The Finite Element Method Using deal.II**
Lecturer	Luca Heltai (SISSA) Martin Kronbichler (TUM)
Class Duration	12 hours
Laboratory Duration	16 hours
Module Description	Hands-on module that guides the students to solve a simple poisson problem
Main Topics	● Generation simple meshes ● Degrees of Freedom - Matrix Sparsities ● A Laplace Solver in 2D ● Dimension independent Laplace Solver ● Adaptively refined meshes ● Hanging nodes and other constraints ● A Parallel Laplace Solver in 2D
Objectives	On successful completion of this module students should be able to understand existing codes for the solution of PDEs, and to develop efficient HPC enabled scientific codes dedicated to the solution of PDEs using existing parallel libraries and tools, dealII in particular.

Module 2.14:	**CMAKE**
Lecturer	Matthias Maier (Minnesota University, USA)
Class Duration	16 hours
Laboratory Duration	16 hours
Module Description	Advanced usage of cmake
Main Topics	- Linking - Test suite
Objectives	Learn how to link different libraries exploiting functionalities of cmake. Learn how to setup an advanced testsuite

Useful links and resources

- **MPI**
  - http://www.mpi-forum.org/
    (MPI documents, including standard specifications with examples, forum for discussing standard specifications)

- **OPENMP**
- http://openmp.org/wp/
    (OpenMP specifications, examples, news on conferences and schools on parallel programming, forum for discussing standard specifications)

- **OPENACC**
    (OpenACC standards for accelerator directives, forums and tutorials)

- **GPGPU**
  - http://hgpu.org/
  - http://gpgpu.org/

  (These are two very frequently updated blogs with software release news, specifications, papers, thesis and codes on GPGPU programming as well as conferences, workshops and schools on HPC world)

  - http://developer.nvidia.com/cuda-zone

(Commercial NVIDIA CUDA website with news and tutorials)

  - http://www.khronos.org/opencl/

(OpenCL standards, reference cards and information on commercial and open source implementations)

- **OTHER MATERIALS**

  - **Intel Parallel Universe Magazine**

  (This is a quarterly publication devoted to exploring innovations in the world of software development with stories and tutorials, especially from Intel software enginners and updates on new Intel products and features)

  - **HPCWIRE**

  http://www.hpcwire.com/
  (News from HPC world and vendors)

➢ **SOME INTERESTING MATERIAL ON EXASCALE**

- **http://www.cse.nd.edu/Reports/2008/TR-2008-13.pdf**
- **http://science.energy.gov/~/media/ascr/pdf/research/am/docs/EMWGreport.pdf**