Micro-Specialization: High-performance computing and applications

Department/School/Center: Centre for Computational and Data Sciences (CD)

Brief Description: This Micro-specialization course will primarily focus on introducing the high-performance computing (HPC) environment to students which helps them to speed up their computations. The course has been designed to have a foundation subject, where the students will learn the basic and essential tools needed to work efficiently on the HPC platforms, followed by interdisciplinary subjects that focus on the applications of HPC systems. It is suitable for students with or without a computer science background.

Number of Subjects needed to earn the Micro-Specialization: 3 Subjects

Credits needed to earn the Micro-Specialization: 12 credits

Structure:

Component I: One Subject (3-1-0)

Component II: Two subjects, with a requirement that at least one subject must be selected from the subjects offered by CCDS (3-1-0/2-0-3)

A. Component- I: Mandatory Requirement (4 Credit Foundation Course)

Table-I

Sub No.	Sub Name	LTP	Credits	Offering Semester	Pre-Requisite (if any)
CD61002	High performance scientific computing	3-1-0	4	Autumn	Basic linear algebra and PDS (first year) or equivalent courses in computer programming.

B. Component- II: Two Subjects (8 Credits) From Table-II

Table-II

Sub No.	Sub Name	LTP	Credits	Offering Semester	Pre-Requisite if any
CD61001	Quantum mechanics and quantum computing	3-1-0	4	Spring	-
CD61006	Quantum methods in molecular simulations	2-0-3	4	Spring	-

CD61004	High performance computing and its applications in complex physical systems	3-1-0	4	Autumn and Spring	-
CS61060	Computational biophysics: algorithms to applications	3-1-0	4	Autumn	-
CY62204	Classical methods of molecular simulation	2-0-3	4	Spring	-
CY62206	Modeling of quantum systems and processes	2-0-3	4	Spring	-
CS61064	High Performance Parallel Programming	3-1-0	4	Spring	-

CD61002: HIGH PERFORMANCE SCIENTIFIC COMPUTING (3-1-0)

SYLLABUS:

SPARSE MATRICES: discretization of differential equations, storage schemes for sparse matrices, permutations and reorderings, direct solution methods

ITERATIVE METHODS AND CONVERGENCE: SOR, gradient search methods: steepest descent, conjugate gradient algorithm, Krylov subspaces methods: Arnoldi's method, GMRES, symmetric Lanczos algorithm, convergence analysis, block Krylov methods, preconditioning techniques, ILU factorization preconditioners, multigrid methods.

DOMAIN DECOMPOSITION: Schwarz algorithms and the Schur complement, graph partitioning: geometric approach, spectral techniques

PARALLEL COMPUTING: architectures for parallel computing, shared and distributed memory performance metrics, parallelization of simple algorithms

MPI and OpenMP: basic MPI and OpenMP calls, parallelizing matrix solvers using domain decomposition

CUDA: GPGPU architecture, thread algebra for matrix operations, accelerating matrix solvers using CUDA

CD61001: QUANTUM MECHANICS AND QUANTUM COMPUTING (3-1-0)

SYLLABUS:
FUNDAMENTAL CONCEPTS: Stern-Gerlach experiment, kets, bras and operators, base kets and matrix representations measurements, observables and the uncertainty relations, change of basis, position, momentum and translation, wave functions in position and momentum space, time evolution and the Schrödinger equation.

THEORY OF ANGULAR MOMENTUM: rotations and angular momentum commutation relations Spin 1/2 systems and Finite rotations, SO(3), SU(2) and Euler rotations, density operators and pure versus mixed ensembles, eigenvalues and eigenstates of angular momentum addition of angular momenta, Schwinger oscillator model of angular momentum spin correlation measurement and Bell's inequality, tensor operators.

QUANTUM COMPUTING: qubits, quantum entanglement; reversible computation, quantum gates; quantum parallelism and simple quantum algorithms; quantum Fourier transforms and its applications, quantum search algorithms; introduction to quantum error correcting codes; entanglement assisted communication; elements of quantum information theory and quantum cryptography.

CD61006: QUANTUM METHODS IN MOLECULAR SIMULATIONS (2-0-3)

SYLLABUS:
THEORY COMPONENT: analytical solution of hydrogen and hydrogenic systems; Born-Oppenheimer approximation; formulation of Hartree-Fock method for molecular systems; post HF methods overview; exchange and correlation concepts; configuration interaction techniques; coupled cluster technique for many-body systems; density functional theory; exchange-correlation functionals, dispersion corrections; time-dependent DFT; geometry optimization techniques; basis sets in quantum chemistry; general framework of molecular simulations; quantum chemical simulations of solids and crystalline materials; simulation of molecular catalysis; vibrational analysis and frequency calculations; simulation of molecular properties.

LAB COMPONENT: building molecules; Z-matrices; internal coordinates; SCF calculations of atoms and small-to-medium sized molecules, restricted and unrestricted Hartree Fock calculations; SCF convergence criteria; geometry optimization techniques; vibrational frequency calculations and analysis of vibrational modes; calculation of Gibbs free energy of formation and Gibbs reaction energy; CI, MP2, coupled-cluster calculations of small molecules; multi-configuration SCF calculations; concept of active space; density functional calculations of large molecules; cost-benefit analysis; effects of basis sets and electron correlation - a comparison of methods; framework for crystalline solids; building crystal planes and supercells; DFT calculations of crystalline solids; simulation of molecular catalysis; biomolecular simulation using hybrid QM/MM approach.

CD61004: HIGH PERFORMANCE COMPUTING AND ITS APPLICATIONS IN COMPLEX PHYSICAL SYSTEMS (3-1-0)

SYLLABUS:

INTRODUCTION TO HPC ARCHITECTURE AND PARALLEL PROGRAMMING: basic architecture and organization: memory hierarchy, shared and distributed memory architectures, multiprocessor architecture, parallel programming, introduction to MPI, optimizing cluster operations: running jobs in HPC environment, job scheduler, cluster level load balancing

SPECIAL METHODS FOR STUDYING COMPLEX SYSTEMS: basics of statistical mechanics, potential energy surface, an introduction to molecular mechanics, Simulation methods: molecular dynamics and Monte Carlo simulations, enhanced sampling methods, coarse-grain modelling

APPLICATIONS TO COMPLEX SYSTEMS: Open-source software: MD and MC simulation packages, parallelization in software: domain/spatial decomposition, distribution of nonbonded interactions, dynamic load balancing, multiprocessor communication, modeling of soft matter systems such as biomolecules, polymers, carbon nanostructures, etc., computation of thermodynamic, kinetic and mechanical properties of different complex systems

CS61060: COMPUTATIONAL BIOPHYSICS: ALGORITHMS TO APPLICATIONS (3-1-0)

SYLLABUS:

INTRODUCTION: central dogma of molecular biology, Relevant databases in computational biophysics, molecular visualization software.

ALGORITHMIC TECHNIQUES FOR MODELLING: Monte Carlo simulation, replica-exchange Monte Carlo simulation, simulated annealing, neural network method.

PROTEIN/NUCLEIC ACID STRUCTURE MODELLING: methods for protein secondary structure prediction, comparative modelling, threading and fold recognition, ab initio modelling, combined modelling approaches, CASP: a blind protein structure prediction competition.

PROTEIN-PROTEIN/NUCLEIC ACID INTERACTION PREDICTION: Fast Fourier Technique, geometric hashing, designing scoring function, protein-protein docking algorithms, protein-nucleic acid docking methods, CAPRI: a blind protein interaction prediction competition.

Selected topic (any one of the following will be covered): (i) protein function annotation: gene ontology, enzyme classification, sequence and structure-based function annotation, meta servers. (ii) protein design: force field design, simulation techniques, ab initio design, interaction design

CY62206: MODELLING OF QUANTUM SYSTEMS AND PROCESSES (2-0-3)

SYLLABUS:

MODULE 1 THEORY: 1. Overview of variational principle and perturbation theory 2. Multi-electron wave functions, Self-consistent field method for atoms and molecules, Fock operator, canonical HF equations, Hartree-Fock Roothan Equation 3. Basis Sets: Hydrogenic functions, Slater type, contracted Gaussian functions, splitvalence, polarization, correlation-consistent basis functions, Effective core potentials 4. Post Hartree-Fock methods (Brief introduction to CI/MP2/CC/DFT) 5. Wave function analysis, Population analysis and atomic charges: Mulliken, Lowdin, natural population 6. Molecular Electrostatic Potential, ESP charge, restrained charge fitting 7. Theory of atoms-in-molecules 8. Frontier Molecular Orbital Theory 9. Inter-molecular interactions 10. Hybrid QM/MM methods: additive/subtractive schemes 11.Molecular response to electric and magnetic fields

MODULE 2: LABORATORY: 1. Z-matrix, Models for small molecules and large bio/nano-complexes 2. Hartree-Fock calculation for small molecules: Analysis of density matrix, Fock matrix, HF Convergence analysis 3. Geometry optimization with density functional methods: Comparison of functionals, energy and other molecular properties 4. Hessian Calculation, Frequency analysis, IR/Raman/VCD spectra simulation 5. Thermochemistry: Evaluation of zero-point energy, enthalpy, entropy, and Gibbs free energy of simple chemical reactions 6. Excited state calculation with TDDFT, Simulation of UV-Vis and emission spectra 7. Simulation of Vibronic spectra with Franck-Condon factors, Duschinsky matrix 8. Techniques of potential energy surface scanning, linear interpolation of geometry, reaction coordinates 9. Transition state search with different techniques, IRC, transition state characterization 10. Wave function analysis: Molecular orbital analysis, partial charge, bond order, density of states 11.

Visual study of weak interaction: atoms-in-molecules, critical points, non-covalent interaction plots 12. Calculation of reactivity descriptors: electronegativity, hardness, polarizability, electrophilicity index; local reactivity descriptors: local softness, Fukui functions 13. Comparison of solvation models: PCM, SMD models, solvation parameters, solvation free energy 14. QM/MM modelling: Layer definition, link atom, embedding schemes

CY62204: CLASSICAL METHODS OF MOLECULAR SIMULATION (2-0-3)

SYLLABUS:

MODULE 1: THEORY: 1. Introduction: Definition of simulation and its history of development 2. Concept of Force-Fields for atomic and molecular systems 3. Handling bulk systems: periodic boundary convention and minimum image convention 4. Molecular dynamics method: Generalized coordinates and equations of motions (EOM): Newton's, Lagrange's and Hamilton's EOM 5. Finite difference methods to integrate EOM: Predictor-corrector, Verlet and VelocityVerlet methods 6. Shifted potential and shifted-force potential; Verlet neighbour list method 7. A few important concepts of Statistical Mechanics 8. Generating initial configuration of a system in a simulation 9. Measurement of temperature and its control in a simulation; Equilibration process 10. Algorithm to develop a molecular dynamics code for a simple Lennard-Jones system 11. Monte Carlo Method: Metropolis scheme and derivation of acceptance criteria 12. Algorithm to develop Metropolis Monte Carlo code for a simple Lennard-Jones system 13. Biased Monte Carlo methods to study flexible molecules 14. Analysis of simulated trajectory: Concepts of distribution functions and time correlation functions 15. Introduction to Extended phase-space approach: NVT and NPT ensemble simulations 16. Introduction to rare event simulation methods 17. Concept of collective variables 18. Simulation of free energy along a chosen collective variable using umbrella sampling and metadynamics 19. Estimation of the rate of activated barrier crossing 20. Application of machine learning and deep learning methods in molecular simulation

MODULE 2: LABORATORY: 1. Generation of Random numbers 2. Calculations of simple pair-wise interaction potentials and their plots 3. Application of periodic boundary conditions and minimum image conventions 4. Calculation of forces acting between interacting particles 5. Development of a routine to integrate EOM using Verlet method 6. Generating initial configurations by random insertion method and from regular lattice

Module 3 Short project:
PROJECT-1: Development of a molecular dynamics code for a Lennard-Jones system.
PROJECT-2: Development of a Metropolis Monte Carlo code for a Lennard-Jones system.
PROJECT-3: Molecular dynamics simulation of liquid water.
PROJECT-4: Potential-of-mean-force for a pair of simple solutes in water.

PROJECT-5: Free energy surface of alanine dipeptide in water.
PROJECT-6: Monte Carlo/Molecular dynamics simulation of a small protein in water.
PROJECT-7: Machine learning-based analysis of MD trajectories

CS61064: HIGH PERFORMANCE PARALLEL PROGRAMMING (3-1-0)

SYLLABUS:

Introduction to HPC Systems - Introduction to basic architecture and OS concepts - Multi-core CPUs - High-speed interconnects - High performance file systems - GPU systems - High performance clusters
Parallel Programming Models, Runtime Systems - OpenMP and MPI - Thread Management, Workload manager and Job Schedulers - CUDA / OpenCL - HW schedulers, Software runtime systems - MapReduce System Architecture and runtime management
Benchmarks and Case Studies - Developer Libraries- Linear Algebra : MKL, BLAS, Lapack - Distributed Machine Learning using Spark Framework - Benchmarks: A. GPU Benchmarks : Parboil, Rhodinia, B. Computational Science : BLAST and SOM (Bioinformatics), Fluid Dynamics (CFD packages)- Hybrid Parallel Programming: Putting it together (Python, MPI, OpenMP, CUDA)
Tutorials - Familiarization with HPC softwares : OpenMP and MPI, Spark Framework for Map-Reduce - Benchmark based performance evaluation experiments on HPC systems - HPC Application development: Drug design, Fault Simulation, Machine Learning Application development