



UNIVERSITÀ  
DEGLI STUDI DI MILANO-BICOCCA

## SYLLABUS DEL CORSO

### Machine Learning and Statistical Physics of Macromolecular Systems

2627-1-F1703Q025

---

#### Obiettivi

Le macromolecole biologiche sono esempi paradigmatici di sistemi complessi: Da un lato, la loro modellizzazione fisica è complicata dalla presenza di interazione su diverse scale di lunghezza. Dall'altro, frustrazione e barriere sia energetiche che entropiche rendono computazionalmente estremamente costosa l'esplorazione dello spazio conformazionale e lo studio delle proprietà sia dinamiche che di equilibrio di questi sistemi.

Questo corso si propone di condurre lo studente dai principi di base fino alla frontiera dello studio teorico e la simulazione dei sistemi macromolecolari complessi. In particolare, si mostrerà come l'integrazione di tecniche di fisica statistica sviluppate negli ultimi 20 anni con algoritmi di machine learning renda oggi possibile affrontare i principali ostacoli computazionali, offrendo un nuovo strumento di indagine chimico-fisico per lo studio processi fondamentali in biofisica, biologia, scienze dei materiali e ricerca farmacologica.

Il corso è concepito per adattarsi a due tipologie diverse di studenti:

studenti provenienti da un percorso teorico in fisica, chimica o scienze dei materiali interessati ad approfondire maggiormente aspetti teorici e concettuali della fisica statistica dei sistemi complessi, della dinamica stocastica e i fondamenti teorici di algoritmi di machine learning che sono utilizzati in questo ambito.

studenti provenienti da un percorso sperimentale (in biofisica, biochimica, o biotecnologie) interessati ad approfondire gli aspetti pratici, ovvero imparare ad eseguire ed interpretare di tecniche di simulazione molecolare. Gli studenti potranno scegliere tra due diverse modalità di esame che riflettono questi due diversi obiettivi formativi.

#### Contenuti sintetici

Dalla dinamica stocastica alle proprietà di equilibrio di sistemi complessi. Tempi di rilassamento. Approcci multi-scala, campi statistici e stocastici, meccanica statistica degli eventi rari e problemi computazionali ad essa

collegati. Cenni di fenomenologia delle macromolecole e in particolare di biopolimeri. Algoritmi AI per l'esplorazione, il campionamento e la analisi di dinamica molecolare. Algoritmi per la predizione di strutture molecolari.

## **Programma esteso**

### **Capitolo 1: Biomolecole come sistemi aperti complessi**

I fondamenti teorici della simulazione delle biomolecole basata sulla fisica  
Il concetto di descrizione multiscala della materia soffice e biologica

### **Capitolo 2: Dinamica stocastica**

Dinamica nei sistemi macromolecolari aperti: scale caratteristiche e processi di fluttuazione-dissipazione  
Derivazione microscopica dell'equazione di Langevin generalizzata. Limite ohmico e limite sovrasmorzato  
Elementi di calcolo stocastico e misura stocastica  
Equazione di Fokker-Planck e integrali di cammino stocastici

### **Capitolo 3: Meccanica statistica dei processi termicamente attivati**

Eventi reattivi e insieme dei percorsi di transizione (transition path ensemble)  
Descrittori stocastici della dinamica reattiva: funzione di committor, densità dei percorsi di transizione e corrente dei percorsi di transizione  
Cinetica dei tempi medi di primo passaggio (Mean First Passage Times) e del tempo di percorrenza della transizione (Transition Path Time)  
Modelli a stati di Markov (Markov State Models)

### **Capitolo 4: Dinamica Molecolare (MD)**

Integratori ergodici  
Dinamica molecolare negli ensemble NPT e NVT  
Simulazioni pratiche con GROMACS: preparazione, esecuzione e analisi di simulazioni di dinamica molecolare (6 ore)

### **Capitolo 5: Calcolo statistico per le simulazioni molecolari**

Variabili collettive e potenziale della forza media (Potential of Mean Force)  
Distanza diffusiva e mappe di diffusione (Diffusion Maps)  
Clustering delle configurazioni e costruzione di Modelli a Stati di Markov  
Varietà intrinseca e dimensionalità intrinseca

Capitolo 6: Metodi di campionamento avanzato e apprendimento automatico per le simulazioni molecolari  
Sfide aperte nella simulazione della materia biologica complessa  
Algoritmi di campionamento avanzato  
Apprendimento automatico (Machine Learning) per le simulazioni molecolari

## **Prerequisiti**

Meccanica analitica Lagrangiana e Hamiltoniana, termodinamica rudimenti di fisica statistica (teoria degli ensemble), meccanica quantistica, rudimenti di metodi matematici per la fisica (distribuzioni, spazi di funzioni).

## **Modalità didattica**

Gli argomenti del corso saranno trattati in lezioni frontali .

Gli studenti interessati potranno svolgere un breve progetto applicativo che costituirà parte dell'esame.

## **Materiale didattico**

Dispense fornite dal docente.

M. Tuckermann: "Statistical mechanics: Theory and Molecular Simulations", Oxford Graduate Texts

T. Schlick: "Molecular Modeling and Simulation: An Interdisciplinary Guide (Interdisciplinary Applied Mathematics)" Springer.

## **Periodo di erogazione dell'insegnamento**

Secondo semestre

## **Modalità di verifica del profitto e valutazione**

Il voto verrà assegnato sulla base di una prova orale finale, tenuto anche conto del contributo dello studente alle sessioni tematiche e relative discussioni.

## **Orario di ricevimento**

In qualsiasi momento, previo accordo con il docente via email

## **Sustainable Development Goals**

IMPRESE, INNOVAZIONE E INFRASTRUTTURE

---