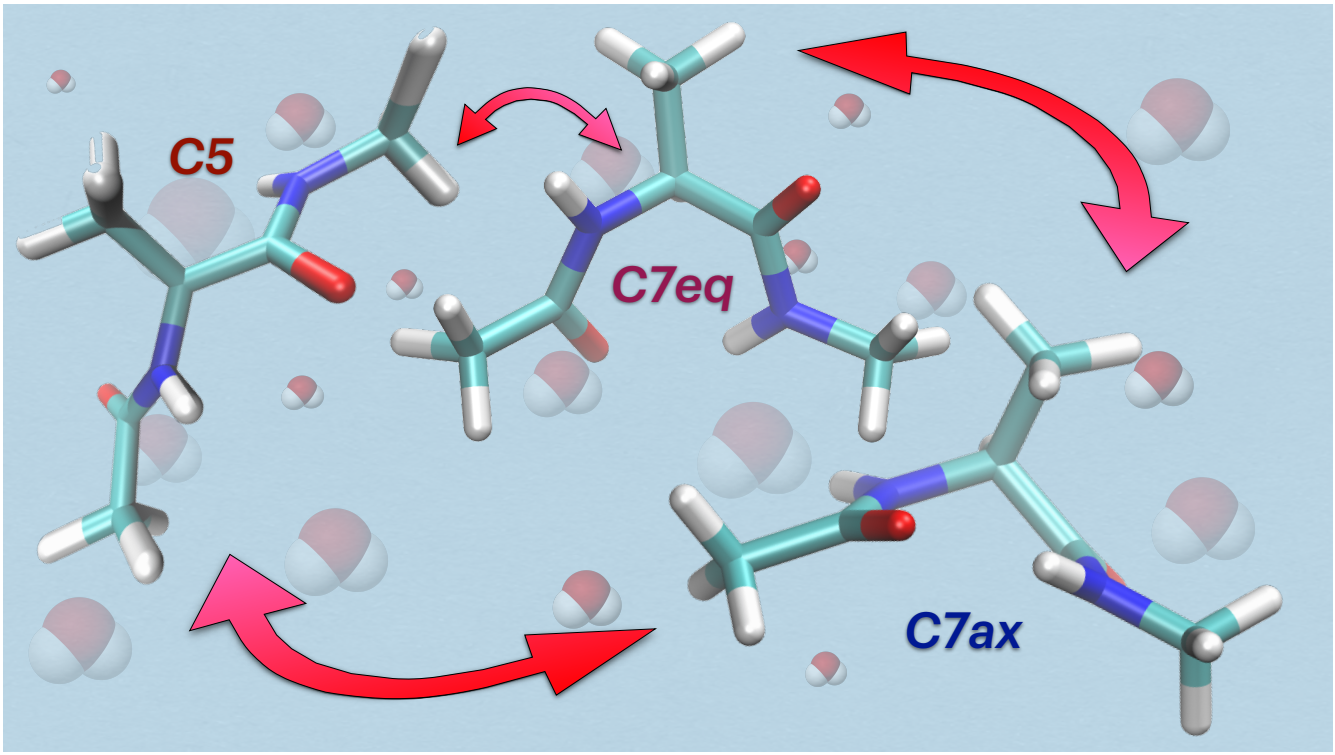

Workshop:

Rare Events: Analysis, Numerics, and Applications

February 27 – March 3, 2023

Brin Mathematics Research Center
University of Maryland
College Park, MD 20742, USA

Program & abstracts

Contents

Schedule at a Glance

	Monday	Tuesday	Wednesday	Thursday	Friday
8:00					
9:00	Breakfast and Opening	Breakfast	Breakfast	Breakfast	Breakfast
10:00	David Wales	Eric Vanden-Eijnden	Yannis Kevrekidis	Benedict Leimkuhler	Ron Elber
	Benoit Roux	K. Spiliopoulos	Xuhui Huang	Carsten Hartmann	Andrew Ballard
11:00					
	Andrew Ferguson	Tony Lelievre	Weiqing Ren	Todd Gingrich	Aaron Dinner
12:00	Lunch (served)	Lunch (on your own)	Lunch (served)	Lunch (on your own)	Lunch (boxes)
13:00					
	Pilar Cossio	Katie Newhall	Hugo Touchette	Grant Rotskoff	
14:00	Greg Bowman	Peter Koltai	Gabriel Stoltz	Sapna Sarupria	
15:00			Lightning talks		
	Bettina Keller	Z. Bezemek, L. Evans		Omar Valsson	
16:00	D. Wang, E. Beyerle	Brainstorming Session	Poster Session	Brainstorming session	
17:00					

Workshop Overview

Events occurring rarely on the timescale of the system affect the system the most dramatically. Often they are associated with transitions from one metastable state to another. The study of rare events is of great importance and of great challenge as their direct simulation is exceedingly difficult due to long waiting times. In recent years, significant progress in studying rare events has been made due to the development of novel numerical tools based on adapting ideas from the realm of data science and machine learning as well as harnessing classical mathematical theories and ideas from theoretical chemistry. This in-person workshop will bring together scientists from various disciplines working on various aspects of the study of rare events and facilitate the exchange of ideas regarding theory, numerics, and diverse applications to problems in biology, chemistry, and physics.

Organizing committee

Maria Cameron, University of Maryland

Pratyush Tiwary University of Maryland

Workshop Schedule

Monday, February 27, 2023

8:45 - 9:15	Breakfast
9:15 - 9:20	Doron Levy (University of Maryland/Director, Brin MRC) *Opening*
9:20 - 9:25	Pratyush Tiwary (University of Maryland) *Welcome*
9:30 - 10:10	David Wales (Cambridge University, UK) *Energy landscapes and rare events*
10:15 - 10:55	Benoit Roux (University of Chicago, USA) *Rare conformational transitions in biomolecular systems*
10:55 - 11:15	Coffee break
11:15 - 11:55	Andrew Ferguson (University of Chicago, USA) *Data-driven collective variable discovery and enhanced sampling*
11:55 - 1:25	Lunch
1:25 - 2:05	Pilar Cossio (Flatiron Institute, NY, USA) *Ensemble reweighting using Cryo-EM particles*
2:10 - 2:50	Greg Bowman (University of Pennsylvania, USA) *Accelerating cryptic pocket discovery with deep learning*
2:50 - 3:10	Coffee break
3:10 - 3:50	Bettina Keller (Free University Berlin, Germany) *Rare events explored by molecular simulations and dynamical reweighting*

3:55 - 4:15 Dedi Wang (University of Maryland, USA)
Introducing physics into representation learning

4:20 - 4:40 Eric Beyerle (University of Maryland, USA)
Machine-learned reaction coordinates for energy-entropy disentanglement

Tuesday, February 28, 2023

9:00 - 9:30 Breakfast

9:30 - 10:10 Eric Vanden-Eijnden (New York University, USA)
Learning to Sample Better

10:15 - 10:55 Konstantinos Spiliopoulos (Boston University, USA)
Novel perturbations for accelerating Langevin samplers

10:55 - 11:15 Coffee break

11:15 - 11:55 Tony Lelievre (Ecole des Ponts Paris, France)
How to compute transition times?

11:55 - 1:25 Lunch (on your own)

1:25 - 2:05 Katie Newhall (University of North Carolina, USA)
Predicting transition times in systems with both stochastically-switching forces and thermal noise

2:10 - 2:50 Peter Koltai (University of Bayreuth, Germany)
Collective variables in complex systems: from molecular dynamics to agent-based models and fluid dynamics

2:45 - 3:15 Coffee break

3:10 - 3:30 Zachary Bezemek (Boston University, USA)
Large Deviations and Importance Sampling for Weakly Interacting Diffusions

3:35 - 3:55 Luke Evans (University of Maryland, USA)
Computing committors in collective variables with Mahalanobis Diffusion Maps

4:00 - 4:40 Brainstorming Session (Moderator: Andrew Ferguson)

Wednesday, March 1, 2023

9:00 - 9:30 Breakfast

9:30 - 10:10 Xuhui Huang (University of Wisconsin-Madison, USA)
Non-Markovian Dynamic Models of Protein Conformational Changes

10:15 - 10:55 Yannis Kevrekidis (Johns Hopkins University, USA)
Staying the course: a one-sided, one-dimensional search with adaptively revealed collective variables

10:55 - 11:15 Coffee break

11:15 - 11:55 Weiqing Ren (National University of Singapore, Singapore)
Computing committor functions for the study of rare events using deep learning

11:55 - 1:25 Lunch

1:25 - 2:05 Hugo Touchette (Stellenbosch University, South Africa)
Machine learning of large deviations

2:10 - 2:50 Gabriel Stoltz (Ecole des Ponts Paris, France)
Coarse-graining molecular systems with autoencoders

3:00 - 3:50 Lightning talks (Students and postdocs, University of Maryland, USA)

3:50 - 5:00 Poster Session (Rotunda of W. E. Kirwan Hall.)
Refreshments will be served

Thursday, March 2, 2023

9:00 - 9:30 Breakfast

9:30 - 10:10 Benedict Leimkuhler (University of Edinburgh, UK)
Parameterisation and convergence of Langevin sampling algorithms

10:15 - 10:55 Carsten Hartmann (BTU Cottbus-Senftenberg, Germany)
Importance sampling for rare events and some pathologies of the exit problem

10:55 - 11:15 Coffee break

11:15 - 11:55 Todd Gingrich (Northwestern University, USA)
Tensor Networks for Chemical Reaction Network Rate Calculations

11:55 - 1:25 Lunch (on your own)

1:25 - 2:05 Grant Rotskoff (Stanford University, USA)
Enhanced sampling with auxiliary models: from coarse-graining to rare events

2:10 - 2:50 Sapna Sarupria (University of Minnesota, USA)
Path sampling of rare events: applications to nucleation

2:50 - 3:10 Coffee break

3:10 - 3:50 Omar Valsson (University of North Texas, USA)
Sampling Rare Event Energy Landscapes via a Birth-Death Process Augmented Langevin Dynamics

4:00 - 4:40 Brainstorming session (Moderator: Eric Vanden-Eijnden)

Friday, March 3, 2023

9:00 - 9:30 Breakfast

9:30 - 10:10 Ron Elber (University of Texas, Austin, USA)
Computing Observables with Exact Milestoning

10:15 - 10:55 Andrew Ballard (DeepMind, UK)
Learned free energy estimation

10:55 - 11:15 Coffee break

11:15 - 11:55 Aaron Dinner (University of Chicago, USA)
Computing transition-path statistics from short-trajectory dataComputing transition-path statistics from short-trajectory data

11:55 - 1:00 Lunch

Abstracts of talks

Energy landscapes and rare events

David Wales

Cambridge University, UK

Monday, February 27, 2023 @ 9:30 AM

The potential energy landscape provides a conceptual and computational framework for investigating structure, dynamics, and thermodynamics in atomic and molecular science. This talk will focus on new developments for efficient diagnosis and analysis of rare event dynamics. In particular, the discrete path sampling approach enables us to construct kinetic transition networks, and graph transformation provides a numerically stable method to extract observable properties for highly metastable systems. The computational energy landscapes approach is based on geometry optimisation, with post-processing using statistical mechanics and unimolecular rate theory, providing access to experimental time scales within a well-defined theoretical framework. We have recently advanced this methodology and applied it to rare events in atomic clusters, proteins, and nucleic acids. The first passage time distributions provide direct signatures of multi-funnel energy landscapes, which underpin metastable systems and the associated rare events, potentially encoding design of multi-functional materials. Selected Publications: J. Phys. Chem. Lett., 13, 6349-6358, 2022. Perspective: Dynamical Signatures of Multifunnel Energy Landscapes Phys Rev E (2021) 104, 015301. Numerical analysis of first-passage processes in finite Markov chains exhibiting metastability J. Chem. Phys. 155, 140901 (2021). Perspective: Nearly reducible finite Markov chains: Theory and algorithms Journal of Chemical Physics (2020) 153, 134115. Rare events and first passage time statistics from the energy landscape J Chem Theory Comput (2020) 16, 2661. Defining, Calculating, and Converging Observables of a Kinetic Transition Network J. Chem. Phys., 130, 204111 (2009). Calculating Rate Constants and Committor Probabilities for Transition Networks by Graph Tra nsformation Ann. Rev. Phys. Chem., 69, 401-425, (2017). Exploring Energy Landscapes Energy Landscapes, Cambridge University Press, Cambridge, 2003

Rare conformational transitions in biomolecular systems

Benoit Roux

University of Chicago, USA

Monday, February 27, 2023 @ 10:15 AM

Classical molecular dynamics (MD) simulations based on atomic models play an increasingly important role in a wide range of applications in physics, biology and chemistry. The approach consists of constructing detailed atomic models of the macromolecular system and having described the microscopic forces with a potential function, using Newton's classical equation, F=MA, to literally "simulate" the dynamical motions of all the atoms as a function of time. The calculated trajectory, though an approximation to the real world, provides detailed information about the time course of the atomic motions, which is impossible to access experimentally. While great progress has been made, producing genuine knowledge about biological systems using MD simulations remains enormously challenging. Among the most difficult problems is the characterization of large conformational transitions occurring over long-time scales. Issues of force field accuracy, the neglect of induced polarization, in particular are also a constant concern. Transition path theory (TPT) offers a powerful paradigm for mapping the conformational landscape of biomolecular systems is to combine free energy methods, string method, transition pathway techniques, and stochastic Markov State Model based massively distributed simulations.1-5 These concepts will be illustrated with a few recent computational studies of biomolecular systems.

References

1. Pan, A. C., Sezer, D., Roux, B. Finding transition pathways using the string method with swarms of trajectories. J. Phys. Chem. B 112, 3432-3440, (2008).

2. Pan, A. C., Roux, B. Building Markov state models along pathways to determine free energies and rates of transitions. J. Chem. Phys. 129, 064107, (2008).

3. Roux, B. String Method with Swarms-of-Trajectories, Mean Drifts, Lag Time, and Committor. J. Phys. Chem. A 125, 7558-7571, (2021).

4. Roux, B. Transition rate theory, spectral analysis, and reactive paths. J. Chem. Phys. 156, 134111, (2022).

5. He, Z., Chipot, C., Roux, B. Committor-Consistent Variational String Method. J. Phys. Chem. Lett. 13, 9263-9271, (2022).

Data-driven collective variable discovery and enhanced sampling

Andrew Ferguson

University of Chicago, USA

Monday, February 27, 2023 @ 11:15 AM

Enhanced sampling in collective variable (CV) space is a powerful technique to accelerate the exploration of the configurational phase space of molecular systems possessing rugged free energy landscapes and long-lived metastable states. The efficiency of these techniques is contingent on the availability of "good" CVs correlated with the important dynamical modes of the system. In this presentation, I will describe two approaches for simultaneous on-the-fly data-driven CV discovery and enhanced sampling. The first enables learning of large-scale collective variables incorporating both solute and solvent degrees of freedom by combining rigid graph theory, permutationally invariant system featurizations, and variational autoencoders. The second appeals to the variational approach to conformational dynamics, Girsanov path reweighting, and deep learning to discover slow CVs from biased simulation trajectories and exploit these slow modes to enhance barrier crossing between metastable states.

Ensemble reweighting using Cryo-EM particles

Pilar Cossio

Flatiron Institute, NY, USA

Monday, February 27, 2023 @ 1:25 PM

3D electron-density reconstruction by cryo-electron microscopy (cryo-EM) is limited to biomolecular samples with low conformational heterogeneity, where all the conformations can be well-sampled at many viewing angles. Despite that the conformational ensemble at room temperature can be conserved by sufficiently fast freezing, existing reconstruction tools cannot retrieve the ensemble density using the cryo-EM particles. These limitations hinder extracting the biomolecule's thermodynamic properties from cryo-EM. To overcome these limitations, we developed an ensemble refinement framework that estimates the ensemble density from individual cryo-EM particles by introducing a prior set of conformations, e.g., from molecular dynamics simulations or structure prediction tools. Building on a previous Bayesian approach, our method recovers the ensemble density in the conformational space, enabling the extraction of populations, and free energies for a simple toy model and from synthetic cryo-EM particles of a protein involving unfolded conformations.

Accelerating cryptic pocket discovery with deep learning

Greg Bowman

University of Pennsylvania, USA

Monday, February 27, 2023 @ 2:10 PM

Cryptic pockets that are absent in proteins' ground state structures but form during structural fluctuations could greatly expand the druggable proteome. However, identifying cryptic pockets is labor-intensive and slow. Here, I will describe how we are using deep learning in combination with physics-based simulations to accelerate the discovery and characterization of cryptic pockets.

Rare events explored by molecular simulations and dynamical reweighting

Bettina Keller

Free University Berlin, Germany

Monday, February 27, 2023 @ 3:10 PM

The relevant timescales of molecular processes range from femtoseconds to minutes or even days. While timescales up several tens or microseconds can now be modeled by direct molecular simulations (with atomistic force fields), slower processes need to be modeled by biased or enhanced sampling simulations. These slow timescales are particularly relevant for intermolecular interactions ranging from chemical reactions between small molecules to protein-protein binding. I will present two methods to recover the correct kinetics from enhanced sampling simulations. (1) In path reweighting one calculates kinetic properties as path ensemble averages. The weight of each path sampled at a biased potential is rescaled to correct weight at the unbiased molecular potential. I will discuss how to extract the path weights efficiently form molecular simulations. (2) By discretizing the Fokker-Planck equation of molecular dynamics, one can obtain kinetic models directly from multi-dimensional (free-)energy surfaces and diffusion profile. I will discuss how to meet the challenges that arise from steep and narrow barriers (such as barriers in chemical reactions) and the presence of nonlinear reaction coordinates.

Introducing physics into representation learning

Dedi Wang

University of Maryland, USA

Monday, February 27, 2023 @ 3:55 PM

Through building connections between artificial intelligence (AI) and physics, we propose methods to learn meaningful representations for molecular dynamics (MD) simulations and biophysical experiments in general. These methods will facilitate biophysics in two ways. First, they will guide the sampling process to accelerate MD simulations of biomolecular systems which would otherwise be prohibitively slow. Second, they will contribute to human understanding of the vast amount of data generated in MD simulations and biophysical experiments.

Machine-learned reaction coordinates for energy-entropy disentanglement

Eric Beyerle

University of Maryland, USA

Monday, February 27, 2023 @ 4:20 PM

The dynamics of physical systems at finite temperatures are governed by free-energy barriers, which are a sum of an energetic component and an entropic component. Here we propose an extension of the time-lagged variational autoencoder framework known as the state predictive information bottleneck (SPIB) that allows for an approximate disentanglement of the energy and entropy barriers in the learned latent space. We demonstrate results of the method on several analytical potentials and alanine dipeptide in vacuum and explain how the learned latent space improves the energy-entropy disentanglement over the original SPIB latent space.

Learning to Sample Better

Eric Vanden-Eijnden

New York University, USA

Tuesday, February 28, 2023 @ 9:30 AM

Sampling high-dimensional probability distributions is a common task in computational chemistry, Bayesian inference, etc. Markov Chain Monte Carlo (MCMC) is the method of choice to perform these calculations, but it is often plagued by slow convergence properties. I will discuss how methods from deep learning (DL) can help enhance the performance of MCMC via a feedback loop in which we simultaneously use DL to learn better samplers based e.g. on generative models, and MCMC to obtain the data for the training of these models. I will illustrates these techniques via several examples, including the sampling of reaction paths in metastable systems and the calculation of free energies and Bayes factors.

Novel perturbations for accelerating Langevin samplers

Konstantinos Spiliopoulos

Boston University, USA

Tuesday, February 28, 2023 @ 10:15 AM

We introduce a novel geometry-informed irreversible perturbation that accelerates convergence of the Langevin algorithm for Bayesian computation. It is well documented that there exist perturbations to the Langevin dynamics that preserve its invariant measure while accelerating its convergence. Irreversible perturbations and reversible perturbations (such as Riemannian manifold Langevin dynamics (RMLD)) have separately been shown to improve the performance of Langevin samplers. We consider these two perturbations simultaneously by presenting a novel form of irreversible perturbation for RMLD that is informed by the underlying geometry. Through numerical examples, we show that this new irreversible perturbation can improve estimation performance over irreversible perturbations that do not take the geometry into account. Moreover we demonstrate that irreversible perturbations generally can be implemented in conjunction with the stochastic gradient version of the Langevin algorithm. I will also discuss the connection of such novel perturbations to the transport map-accelerated Markov chain Monte Carlo algorithm and their similarities and differences in continuous and discrete time.

How to compute transition times?

Tony Lelievre

Ecole des Ponts Paris, France

Tuesday, February 28, 2023 @ 11:15 AM

We will present some recent works on numerical methods to compute transition times, using splitting techniques and the Hill relation. These are joint works in particular with Manon Baudel, Arnaud Guyader, Julien Reygner and Mouad Ramil.
References:
* M. Baudel, A. Guyader, and T. Lelievre. On the Hill relation and the mean reaction time for metastable processes, Stochastic Processes and their Applications, 155, 393-436, (2023).
* C.E. Bréhier, T. Lelièvre, and M. Rousset. Analysis of Adaptive Multilevel Splitting algorithms in an idealized case, ESAIM P&S, 19, 361 - 394, (2015).
* T. Lelievre, M. Ramil, and J. Reygner. Estimation of statistics of transitions and Hill relation for Langevin dynamics, https://arxiv.org/abs/2206.13264
* I. Teo, C. Mayne, K. Schulten, and T. Lelievre. Adaptive multilevel splitting method for molecular dynamics calculation of benzamidine-trypsin dissociation time, Journal of Chemical Theory and Computation, 12(6), 2983–2989, (2016).

Predicting transition times in systems with both stochastically-switching forces and thermal noise

Katie Newhall

University of North Carolina, USA

Tuesday, February 28, 2023 @ 1:25 PM

Biological systems under the influence of microscale active agents such as proteins are frequently modeled using stochastically-switching forces as the agents shift between different states. These rapidly switching forces are often on timescales faster than the time to reach thermal equilibrium, thus the system is in a constant state of disequilibrium. In one example system, a bead-spring polymer model of chromosomes with additional crosslinking stochastically-switching forces, long-lived stable condensed clusters of beads are observed consistent with experimental results. The lifetime of these clusters is linked to the stochastic switching rate which acts like an effective temperature: rapid switching produces low-temperature-like stable clusters, slow switching produces high-temperature-like amorphic arrangements, and intermediate switching times allow for dynamic clusters with beads exchanging between clusters. I will derive an effective energy or quasipotential for the system, showing that different limits concerning the two sources of noise (randomly switching protein binding forces and thermal noise) produce different predictions for the metastable transition times between clusters.

Collective variables in complex systems: from molecular dynamics to agent-based models and fluid dynamics

Peter Koltai

University of Bayreuth, Germany

Tuesday, February 28, 2023 @ 2:10 PM

The identification of persistent forecastable structures in complicated or high-dimensional dynamics is vital for a robust prediction (or manipulation) of such systems in a potentially sparse-data setting. Such structures can be intimately related to so-called collective variables known for instance from statistical physics. We have recently developed a first data-driven technique to find provably good collective variables in molecular systems. Here we will discuss how the concept generalizes to other applications as well, such as fluid dynamics and social or epidemic dynamics.

Large Deviations and Importance Sampling for Weakly Interacting Diffusions

Zachary Bezemek

Boston University, USA

Tuesday, February 28, 2023 @ 3:10 PM

In this talk, we design an importance sampling scheme for statistics related to rare events for the empirical measure on a system of N weakly interacting diffusions. We use the subsolution approach of Dupuis and Wang. It turns out that, due to the large deviation rate function's connection to mean-field control, the natural Hamilton-Jacobi-Bellman Equation to consider is posed on Wasserstein Space and involves derivatives in the sense of Lions . We identify conditions under which our scheme is provably asymptotically optimal in N in the sense of log-efficiency. We also provide evidence, both analytical and numerical, that our scheme can have vanishingly small relative error as N increases.

Computing committors in collective variables with Mahalanobis Diffusion Maps

Luke Evans

University of Maryland, USA

Tuesday, February 28, 2023 @ 3:35 PM

In this talk, I will describe a diffusion maps-based method for approximating a backward Kolmogorov operator in user-defined collective variables while alleviating typical dimensionality and timescale hurdles.The method incorporates position-dependent diffusion matrices through use of Mahalanobis kernels and incorporates MD importance sampling techniques such as metadynamics via reweighting to a user-defined "target measure". The resulting algorithm, "target measure Mahalanobis diffusion map" (tm-mmap), is suitable for a moderate number of collective variables in which one can approximate the position-dependent diffusion matrices and free energy. Particularly, I will illustrate that tm-mmap paves a tractable and provably accurate avenue for transition path theory in collective variables, and can calculate the committor as well as the reactive current delineating the transition channels and the transition rate.

Non-Markovian Dynamic Models of Protein Conformational Changes

Xuhui Huang

University of Wisconsin-Madison, USA

Wednesday, March 1, 2023 @ 9:30 AM

Protein's dynamic transitions between metastable conformational states play an important role in numerous biological processes. Markov State Model (MSM) built from molecular dynamics (MD) simulations provides a useful approach to study these complex dynamic transitions, but it is challenging to build truly Markovian models due to the limited length of lag time (bound by the length of relatively short MD simulations). In this talk, I will introduce our recent work on developing Generalized Master Equation (GME) methods based on the projection operator scheme that encodes the non-Markovian dynamics in a generally time-dependent memory kernel, whose characteristic decay time scale corresponds to the kernel lifetime. We show that GME methods can greatly improves upon Markovian models by accurately predicting long timescale dynamics using much shorter MD trajectories on complex conformational changes including clamp opening of RNA polymerase. Based on the projection operator scheme, I will introduce our recent development of the Encoder-neural-network based "RPnet" method for coarse-graining protein dynamics. The key insight of our RPnet method is that we designed a reverse projection scheme that allows us to quantify the difference (defined as the loss function) of transition modes between original dynamics and coarse-grained dynamics. RPnet provides a new way to define the loss function based on transition modes, which is different from other methods where the loss function is defined based on the variational principle of conformational dynamics. RPnet could yield comparable or better results than competing methods in terms of state partitioning and reproduction of slow dynamics in various systems. Finally, I will present our recent work on the development of Memory-Net (an encoder-neural-network) that can efficiently identify the slow collective variables (CVs) describing protein conformational changes by minimizing the integrals of memory kernels (defined as the loss function). We expect that these GME-based methods hold promise to be widely applied to study functional dynamics of proteins.

Staying the course: a one-sided, one-dimensional search with adaptively revealed collective variables

Yannis Kevrekidis

Johns Hopkins University, USA

Wednesday, March 1, 2023 @ 10:15 AM

We introduce a method to successively locate equilibria (steady states) of dynamical systems on Riemannian manifolds. The manifolds need not be characterized by an a priori known atlas or by the zeros of a smooth map. Instead, they can be defined by point-clouds and sampled as needed through an iterative process. If the manifold is an Euclidean space, our method follows isoclines, curves along which the direction of the vector field X is constant. For a generic vector field X, isoclines are smooth curves and every equilibrium lies on isoclines. We generalize the definition of isoclines to Riemannian manifolds through the use of parallel transport: generalized isoclines are curves along which the directions of X are parallel transports of each other. As in the Euclidean case, generalized isoclines of generic vector fields X are smooth curves that connect equilibria of X. Our algorithm can be regarded as an extension of the method of Newton trajectories to the manifold setting when the manifold is unknown.

This work is motivated by computational statistical mechanics, specifically high dimensional (stochastic) differential equations that model the dynamics of molecular systems. Often, these dynamics concentrate near low-dimensional manifolds and have transitions (saddle points) between metastable equilibria. We employ iteratively sampled data and isoclines to locate these saddle points. Coupling a black-box sampling scheme (e.g., Markov chain Monte Carlo) with manifold learning techniques (diffusion maps in the case presented here), we show that our method reliably locates equilibria of X. Additionally, we show that our methodology can be readily adapted to other methods, such as gentlest ascent dynamics.

We build on our previous iMapD method (intrinsic map dynamics) where the entire free energy surface was explored using locally identified and adaptively grown collective coordinates. Here, we use a scheme for finding saddle points without the knowledge of the final states (in the spirit of Newton trajectories, gradient extremals, and gentlest ascent dynamics) in a collective variable agnostic way, iteratively switching charts to discover one-dimensional curves joining critical points of the free energy surface.

Computing committor functions for the study of rare events using deep learning

Weiqing Ren

National University of Singapore, Singapore

Wednesday, March 1, 2023 @ 11:15 AM

The committor function is a central object in understanding transitions between metastable states in complex systems. It has a simple mathematical description – it satisfies the backward Kolmogorov equation. However, computing the committor function for realistic systems at low temperature is a challenging task, due to the curse of dimensionality and the scarcity of transition data. In this talk, I will present a computational approach that overcomes these issues and achieves good performance on complex benchmark problems with rough energy landscapes. The new approach combines deep learning, importance sampling and feature engineering techniques. This establishes an alternative practical method for studying rare transition events among metastable states of complex, high dimensional systems. If time allows, I will also discuss the computation of quasi-potentials using short trajectories.

Machine learning of large deviations

Hugo Touchette

Stellenbosch University, South Africa

Wednesday, March 1, 2023 @ 1:25 PM

The likelihood of nonequilibrium fluctuations is described within large deviation theory by two functions - the scaled cumulant generating function and the rate function - that are notoriously difficult to calculate analytically or numerically. In this talk, I will describe recent work with Grant Rotskoff and Jiawei Yan (PRE 2022) on learning and representing these functions using neural networks and stochastic optimisation. Using one long trajectory of a stochastic process, I will show how the scaled cumulant generating function can be estimated reliably by iteratively solving an optimal control problem in which control forces, representing physically the forces needed to generate rare fluctuations, are represented by a neural network. I will compare this approach with other works done recently, combining machine learning and large deviations, and will present applications for a simple diffusion and an interacting system of active particles.

Coarse-graining molecular systems with autoencoders

Gabriel Stoltz

Ecole des Ponts Paris, France

Wednesday, March 1, 2023 @ 2:10 PM

A coarse-grained description of atomistic systems in molecular dynamics is provided by reaction coordinates. These nonlinear functions of the atomic positions are a basic ingredient to compute more efficiently average properties of the system of interest, such as free energy profiles. However, reaction coordinates are often based on an intuitive understanding of the system, and one would like to complement this intuition or even replace it with automated tools. One appealing tool is autoencoders, for which the bottleneck layer provides a low dimensional representation of high dimensional atomistic systems. In order to have an efficient numerical method, autoencoders should be combined with importance sampling techniques based on adaptive biasing methods. The algorithm then iterates between an update of the reaction coordinate, and free energy biasing. I will discuss some mathematical foundations of this method, and present illustrative applications for biophysical systems, including alanine dipeptide and chignolin. Some on-going extensions to more demanding systems, namely HSP90, will also be hinted at. Depending on time, I will also mention current extensions aiming at better understanding the structure and behavior of autoencoders.

Lightning talks and Posters

Students and postdocs, University of Maryland, USA

Wednesday, March 1, 2023 @ 3:00 PM

- Christopher Moakler
  Using random graph theory to simplify hydrocarbon pyrolysis simulations
- Shashank Sule
  Error analysis of target measure diffusion maps and applications to transition path theory
- Margot Yuan
  Sampling rare events using model reduction and optimal control
- Bodhi Vani
  From sequence to Boltzmann ranking: AlphaFold2-RAVE
- Shams Mehdi
  Investigating ligand dissociation pathways in RNAs using Artificial Intelligence
- Zachary Smith
  Leveraging enhanced sampling for drug discovery
- Lukas Herron
  Predicting structural ensembles from RNA sequences

Poster Session

Refreshments will be served

Rotunda of W. E. Kirwan Hall.

Wednesday, March 1, 2023 @ 3:50 PM

Parameterisation and convergence of Langevin sampling algorithms

Benedict Leimkuhler

University of Edinburgh, UK

Thursday, March 2, 2023 @ 9:30 AM

I will discuss the design of numerical methods for enhanced sampling of invariant measures in diverse applications, including molecular dynamics. Our previous research has produced accurate numerical integration methods for Brownian (overdamped) dynamics, underdamped Langevin and "Adaptive" Langevin dynamics. I will discuss recent theoretical results on convergence of measures computed using these algorithms in different paramerization regimes. I will also show some intriguing results from numerical experiments in the low friction regime (relevant for rare-event studies) which is currently not well understood.

Importance sampling for rare events and some pathologies of the exit problem

Carsten Hartmann

BTU Cottbus-Senftenberg, Germany

Thursday, March 2, 2023 @ 10:15 AM

Importance sampling and multilevel splitting are key techniques for rare event simulation. Yet they may produce poor results or become inefficient for bad choices of a proposal distribution (in case of importance sampling) or an inappropriate design of the intermediate checkpoints or milestones (in case of multilevel splitting). The problem is known to be more severe for importance sampling, since the likelihood ratio that appears in the importance sampling estimator is not bounded a priori (in contrast to the particle weights in splitting), which may render importance sampling estimators useless in certain situations, especially for systems with high-dimensional state space or for processes involving long trajectories. We discuss importance sampling of exit problems that involve unbounded stopping times; examples are mean first passage times, transition rates or committor probabilities in molecular dynamics. The naive application of variance minimization techniques can lead to pathologies here, including proposals measures that are not absolutely continuous to the reference measure or importance sampling estimators that formally have zero variance, but that produce infinitely long trajectories. In this presentation, we will illustrate these issues with simple examples and discuss possible solutions that are based on a risk-sensitive optimal control framework of importance sampling.

Tensor Networks for Chemical Reaction Network Rate Calculations

Todd Gingrich

Northwestern University, USA

Thursday, March 2, 2023 @ 11:15 AM

Sampling-based approaches to rate calculations are typically necessitated by the impossibility to evolve complex dynamics in distribution. In my talk, I will explore situations in which that evolution can be executed via a controllable approximation using tensor network methodologies. In doing so, I will illustrate a way to numerically extract rates of metastable switching in stochastic chemical reaction networks without sampling. Challenges and outlook will be discussed.

Enhanced sampling with auxiliary models: from coarse-graining to rare events

Grant Rotskoff

Stanford University, USA

Thursday, March 2, 2023 @ 1:25 PM

In probability theory, the notion of "weak convergence" is often used to describe two equivalent probability distributions. This metric requires equivalence of the average value of well-behaved functions under the two probability distributions being compared. In coarse-grained modeling, Noid and Voth developed a thermodynamic equivalence principle that has a similar requirement. Nevertheless, there are many functions of the fine-grained system that we simply cannot evaluate on the coarse-grained degrees of freedom. In this talk, I will describe an approach that combines accelerated sampling of a coarse-grained model with invertible neural networks to invert a coarse-graining map in a statistically precise fashion. I will show that for non-trivial biomolecular systems, we can recover the fine-grained free energy surface from coarse-grained sampling.

Path sampling of rare events: applications to nucleation

Sapna Sarupria

University of Minnesota, USA

Thursday, March 2, 2023 @ 2:10 PM

The transition of metastable liquid (pure liquid and solutions) to solid crystalline states is of utmost importance in the field of materials science. From ice nucleation in clouds that govern the planet's climate and weather to stability of our bones – liquid-to-solid crystal plays a critical role. Most of these processes involve the nucleation step which relates to the formation of the solid embryo that catalyzes solid growth in the metastable liquid. The length and timescale associated with the nucleation step are beautifully suited for molecular simulation studies, however, the associated free energy barrier makes this a rare event in molecular simulations. Several advanced techniques have been used to address this challenge. Path sampling methods like forward flux sampling and transition interface sampling are powerful methods to study nucleation. However, there are a number of nuances that need to be considered in ensuring that these methods are performed correctly. Furthermore, given these calculations are extremely expensive it is important to be able to improve their efficiency as well as maximize the information that we can extract from these simulations. In my talk, I will describe our efforts on all these fronts. I will focus on quantifying the quality of sampling from forward flux sampling and transition interface sampling as well as discuss our efforts to develop novel methods that allow us to extract significant structural information to gain insights into the nucleation process.

Sampling Rare Event Energy Landscapes via a Birth-Death Process Augmented Langevin Dynamics

Omar Valsson

University of North Texas, USA

Thursday, March 2, 2023 @ 3:10 PM

A common problem that affects atomistic simulations within the computational physics and chemistry communities is the so-called sampling problem or rare event problem where proper sampling of energy landscapes is impeded by the presence of high kinetic barriers that hinder transitions between metastable states on typical simulation time scales. Many enhanced sampling methods have been developed to address this sampling problem and more efficiently sample rave event systems [1]. An interesting idea, coming from the field of statistics, was introduced in a recent work [2] in the form of a novel sampling algorithm that augments overdamped Langevin dynamics with a birth-death process. In this work [3], we show that this birth-death sampling scheme can efficiently sample prototypical rare event energy landscapes, and that the speed of equilibration is independent of the barrier height. We amend a crucial shortcoming of the original algorithm that leads to incorrect sampling of barrier regions by introducing a new approximation of the birth-death term. We establish important theoretical properties of the modified algorithm and prove mathematically that the relevant convergence results still hold. We investigate via numerical simulations the effect of various parameters, and we investigate ways to reduce the computational effort of the sampling scheme. We show that the birth-death mechanism can be used to accelerate sampling in the more general case of underdamped Langevin dynamics that is more commonly used in simulating physical systems. Our results show that this birth-death scheme is a promising method for sampling rare event energy landscapes.

[1] J. Henin, T. Lelievre, M. R. Shirts, O. Valsson, and L. Delemotte, arXiv:2202.04164 (2022)
[2] Y. Lu, J. Lu, and J. Nolen, arXiv:1905.09863 (2019)
[3] B. Pampel, S. Holbach, L. Hartung, and O. Valsson, arXiv:2209.00607 (2022)

Computing Observables with Exact Milestoning

Ron Elber

University of Texas, Austin, USA

Friday, March 3, 2023 @ 9:30 AM

I will describe how the theory and algorithm of exact Milestoning [1] are used to solve time-dependent and time-independent problems. The Free Energy profile, Mean First Passage Time, Committor function, Exit Time, and reaction coordinates will be discussed. [2] Emphasis will be made on small model systems, but a large example will be described briefly. Finally, software that automates the Milestoning calculations (ScMiles) will be presented. [3]

[1] Juan M Bello-Rivas and Ron Elber, "Exact Milestoning", Journal of Chemical Physics, 142,094102(2015)

[2] Piao Ma, Ron Elber and Dmitrii E Makarov, "Value of Temporal Information when Analyzing Reaction Coordinates", J. Chem. Theory and Comp. 2020, 16,6077-6090

[3] Alfredo E. Cardenas, Allison Hunter, Hao Wang, and Ron Elber, "ScMiles2: A script to conduct and analyze Milestoning trajectories for long time dynamics", J. Chem. Theory Comp., 18,6952-6965(2022)

Learned free energy estimation

Andrew Ballard

DeepMind, UK

Friday, March 3, 2023 @ 10:15 AM

Free energy computation is of key importance in the natural sciences yet remains a real challenge to estimate for complex systems exhibiting rare events. While traditional estimators such as Free Energy Perturbation and Bennett's Acceptance Ratio continue to be the workhorses of free energy estimation, recent ML-based extensions have made it possible to achieve more accurate estimates while using less data. In this talk I will describe how normalizing flows can be incorporated into traditional free energy methods (work going back to Jarzynski in 2002), and present a machine learning procedure for training normalizing flows that can result in gains in efficiency and accuracy. I will present two recent case studies that employ this technique, a prototype model of liquid solvation [Wirnsberger et al. 2020 J. Chem. Phys.] and atomic solids [Wirnsberger et al 2022 Mach. Learn.: Sci. Technol.], which use a novel normalizing flow architecture tailored to the underlying symmetries of the system of study.

Computing transition-path statistics from short-trajectory dataComputing transition-path statistics from short-trajectory data

Aaron Dinner

University of Chicago, USA

Friday, March 3, 2023 @ 11:15 AM

Understanding molecular mechanisms requires estimating statistics such as expected hitting times, reaction rates, and committors. In systems with well-defined metastable states and free energy barriers, these quantities can be estimated using enhanced sampling methods combined with classical rate theories. However, calculating such statistics for more complex processes with rugged landscapes and/or multiple pathways requires more general numerical methods. In this lecture, I will describe my group's recent efforts to develop methods for estimating transition-path statistics by combining information from many short molecular dynamics trajectories.

The Brin Mathematics Research Center

The Brin Mathematics Research Center is a research center that sponsors activity in all areas of pure and applied mathematics and statistics. The Brin MRC was funded in 2022 through a generous gift from the Brin Family. The Brin MRC is part of the Department of Mathematics at the University of Maryland, College Park.

Activities sponsored by the Brin MRC include long programs, conferences and workshops, special lecture series, and summer schools. The Brin MRC provides ample opportunities for short-term and long-term visitors that are interested in interacting with the faculty at the University of Maryland and in experiencing the metropolitan Washington DC area.

The mission of the Brin MRC is to promote excellence in mathematical sciences. The Brin MRC is home to educational and research activities in all areas of mathematics. The Brin MRC provides opportunities to the global mathematical community to interact with researchers at the University of Maryland. The center allows the University of Maryland to expand and showcase its mathematics and statistics research excellence nationally and internationally.

List of Participants

Andrew Ballard, DeepMind, UK
Andrew Ferguson, University of Chicago, USA
Benedict Leimkuhler, University of Edinburgh, UK
Benoit Roux, University of Chicago, USA
Bettina Keller, Free University Berlin, Germany
Bodhi Vani, University of Maryland, USA
Carsten Hartmann , BTU Cottbus-Senftenberg, Germany
Christopher Moakler, University of Maryland, USA
David Wales, Cambridge University, UK
Dedi Wang, University of Maryland, USA
Doron Levy, University of Maryland/Director, Brin MRC
Eric Beyerle, University of Maryland, USA
Eric Vanden-Eijnden, New York University, USA
Gabriel Stoltz, Ecole des Ponts Paris, France
Grant Rotskoff, Stanford University, USA
Greg Bowman, University of Pennsylvania, USA
Hugo Touchette, Stellenbosch University, South Africa
Katie Newhall , University of North Carolina, USA
Konstantinos Spiliopoulos, Boston University, USA
Lukas Herron, University of Maryland, USA
Luke Evans, University of Maryland, USA
Margot Yuan, University of Maryland, USA
Maria Cameron, University of Maryland, USA
Omar Valsson, University of North Texas, USA
Peter Koltai, University of Bayreuth, Germany
Pilar Cossio, Flatiron Institute, NY, USA
Pratyush Tiwary, University of Maryland
Ron Elber, University of Texas, Austin, USA
Sapna Sarupria, University of Minnesota, USA
Shams Mehdi, University of Maryland, USA
Shashank Sule, University of Maryland, USA
Todd Gingrich, Northwestern University, USA
Tony Lelievre, Ecole des Ponts Paris, France
Weiqing Ren, National University of Singapore, Singapore
Xuhui Huang, University of Wisconsin-Madison, USA

Yannis Kevrekidis, Johns Hopkins University, USA
Zachary Bezemek, Boston University, USA
Zachary Smith, University of Maryland, USA