Abstracts of talks (alphabetical order of speakers)

D. Aristoff, Optimizing stratified steady-state resampling methods

We show how to optimize steady-state resampling methods for Markov processes based on stratification. Here, a low dimensional projection of state space – usually corresponding to a few reaction coordinates – has been divided into a set of strata, or bins. Traditionally, resampling is based on an ad hoc rule: maintaining about the same number of replicas in each bin. We derive, using first principles, a strategy for optimizing the replica allocation among a pre-selected set of bins. We also propose a variance estimation that may be used to help choose parameters, including resampling times and the collection of bins. Potential applications include efficient computation of reaction rates or mean first passage times. (Joint work with Daniel M. Zuckerman.)

A. Arnold, Large-time behavior in (hypo)coercive ODE-systems and kinetic models

In this talk we discuss the convergence to equilibrium in conservative-dissipative ODE-systems, kinetic relaxation models (of BGK-type), and Fokker-Planck equation. This will include symmetric, non-symmetric and hypocoercive evolution equations. A main focus will be on deriving sharp decay rates.

We shall start with hypocoercivity in ODE systems, with the "hypocoercivity index" characterizing its structural complexity.

BGK equations are kinetic transport equations with a relaxation operator that drives the phase space distribution towards the spatially local equilibrium, a Gaussian with the same macroscopic parameters. Due to the absence of dissipation w.r.t. the spatial direction, convergence to the global equilibrium is only possible thanks to the transport term that mixes various positions. Hence, such models are hypocoercive.

We shall prove exponential convergence towards the equilibrium with explicit rates for several linear, space periodic BGK-models in dimension 1 and 2. Their BGK-operators differ by the number of conserved macroscopic quantities (like mass, momentum, energy), and hence their hypocoercivity index. Our discussion includes also discrete velocity models, and the local exponential stability of a nonlinear BGK-model.

The third part of the talk is concerned with the entropy method for (non)symmetric Fokker-Planck equations, which is a powerful tool to analyze the rate of convergence to the equilibrium (in relative entropy and hence in L1). The essence of the method is to first derive a differential inequality between the first and second time derivative of the relative entropy, and then between the entropy dissipation and the entropy. For hypocoercive Fokker-Planck equations, i.e. degenerate parabolic equations (with drift terms that are linear in the spatial variable) we modify the classical entropy method by introducing an auxiliary functional (of entropy dissipation type) to prove exponential decay of the solution towards the steady state

in relative entropy. The obtained rate is indeed sharp (both for the logarithmic and quadratic entropy). Finally, we extend the method to the kinetic Fokker-Planck equation (with non-quadratic potential).

References:
F. Achleitner, A. Arnold, E.A. Carlen: On linear hypocoercive BGK models; in Springer Proceedings in Mathematics & Statistics, Vol. 126, 2016; p. 1-37.
F. Achleitner, A. Arnold, E.A. Carlen: On multi-dimensional hypocoercive BGK models; KRM 11, 2018; p. 953-1009.
F. Achleitner, A. Arnold, B. Signorello: On optimal decay estimates for ODEs and PDEs with modal decomposition; to appear 2018.
F. Achleitner, A. Arnold and D. Strzer: Large-time behavior in non-symmetric Fokker-Planck equations; Rivista di Matematica della Universit di Parma 6 (2015) 1-68.
A. Arnold, J. Erb: Sharp entropy decay for hypocoercive and non-symmetric Fokker-Planck equations with linear drift; preprint

M. Benaïm, Processes with reinforcement and approximation of Quasi-Stationary Distributions

I will discuss a class of recursive algorithms for the approximation of the quasi-stationary distribution of a general Markov chain on a compact metric space killed in finite time. The idea is to run the process until extinction and then to bring it back to life at a position randomly chosen according to the empirical occupation measure of its past positions. The original idea of the algorithm goes back to, Aldous, Flannery, and Palacios in the late 80s. I will explain here how the long term behavior of the process can be well understood and analyzed by combining the ODE method (developed in the late 90s for processes with reinforcement) with recent results by Champagnat and Villemonais. This is based on recent works in collaboration with B. Cloez and F. Panloup. (Annals of Applied Probability, 2018, Vol. 28, No 4, 2370-2416.) Continuous time version of the algorithm (for diffusions with soft killing) have been very recently investigated by A. Wang, G. Roberts and D. Steinsaltz (arXiv:1808.07086).

N. Berglund, Trace process and metastability

The trace process associated with a subset A of the space of a Markov chain is the process monitored only when visiting A. I will show on a number of simple examples how this process is useful to uncover the metastable dynamics of chains in discrete or continuous space; in particular, how it allows to determine spectral properties, without having to assume reversibility.

Joint work with Manon Baudel (CERMICS, Ecole des Ponts).

J. Bierkens, Reflections on the bouncy particle sampler and Zig-Zag sampler

In recent years piecewise deterministic Markov processes (PDMPs) have emerged as a promising alternative to classical MCMC algorithms. In particular these PDMP based algorithms have good convergence properties and allow for efficient subsampling. Although many different PDMP based algorithms can be designed, two algorithms play fundamental roles: the Bouncy Particle sampler and the Zig-Zag sampler. In this talk both algorithms will be introduced and a comparison of properties of these algorithms will be presented, including

recent results on ergodicity and on scaling with respect to dimension. This is joint work with Pierre-Andr Zitt, Kengo Kamatani and Gareth Roberts.

F. Bouchet, Rare events in complex dynamical systems: the examples of the climate and the solar system dynamics – Instantons in complex dynamical systems ?

Do rare events have similar phenomenologies in complex deterministic dynamical systems similar compared to simple diffusions? This question will be discussed mainly through two interesting examples: instantons for the destabilisation of the solar system through Mercury eccentricity growth, or the probability of extreme heat waves in climate models. How to compute those rare events in complex dynamical systems is another key question. I will address the conditions for current rare event algorithms to be useful.

C.-E. Bréhier, Sampling invariant distributions of SPDEs

We consider the question of sampling (efficiently) the invariant distribution of some ergodic (semilinear parabolic SPDEs) of the type

$$du(t) = Au(t)\,dt + F(u(t))\,dt + dW(t),$$

driven by space-time white noise. A standard integrator is the linear implicit Euler scheme, however its weak order of convergence is only equal to 1/2 (contrary to 1 in a finite dimensional SDE context). We will see several techniques which may help reduce the complexity of this infinite dimensional sampling problem. This will be supported by theoretical analysis and numerical simulations.

F. Cazals, Randomized algorithms for volume/density of states calculations in high-dimensional spaces

Randomized algorithms are at the heart of numerical methods computing volumes (in a geometric setting) or density of states (is a statistical physics setting). This talk will review the main ideas common to these methods, and elaborate on them. In the first part, we will review the strategies used in volume calculation methods [1,2] and density of states calculations methods [3]. The emphasis will be placed on recent developments improving practical performances (convergence tests, annealing schedule, rounding). In the second part, we will present a novel random walk based upon billiard walk [4] and Hamiltonian Monte Carlo. Robustness issues to make the method effective will be discussed. Tests for bodies of various dimensions presented, stressing the range of dimensions best suited for the method.

[1] I. Emiris an V. Fisikopoulos. Efficient random-walk methods for approximating polytope volume. In Proceedings of the thirtieth annual symposium on Computational geometry, page 318. ACM, 2014.
[2] B. Cousins and S. Vempala. A practical volume algorithm. Mathematical Programming Computation, 8(2):133-160, 2016.
[3] F. Wang and D.P. Landau. Efficient, multiple-range random walk algorithm to calculate the density of states. Physical review letters, 86(10):2050, 2001.
[4] B. Polyak and E.N. Gryazina. Billiard walk-a new sampling algorithm for control and optimization. IFAC Proceedings Volumes, 47(3):6123-6128, 2014.

A. Duncan, A general Stein based methodology for tuning sampling algorithms in molecular dynamics simulations

As increasingly sophisticated enhanced sampling methods are developed, a challenge that arises is how to tune the simulation parameters appropriately to achieve good sampler performance. This is particularly pertinent for biased sampling methods, which trade off asymptotic exactness for computational speed. While a reduction in variance due to more rapid sampling can outweigh any bias introduced, the inexactness creates new challenges for hyperparameter tuning. The natural metric in which to quantify this discrepancy is the Wasserstein or Kantorovich metric. However, the computational difficulties in computing this quantity has typically dissuaded practitioners. To address this, we introduce a new computable quality measure using Stein operators constructed from diffusions which quantify the maximum discrepancy between sample and target expectations over a large class of test functions. We will also illustrate applications to hyper-parameter selection, convergence rate assessment, and quantifying bias-variance tradeoffs in asymptotically biased sampling. A preprint can be found here: arxiv.org/abs/1611.06972

A. Durmus, The promises and pitfalls of stochastic gradient langevin dynamics

Stochastic Gradient Langevin Dynamics (SGLD) has emerged as a key MCMC algorithm for Bayesian learning from large scale datasets. While SGLD with decreasing step sizes converges weakly to the posterior distribution, the algorithm is often used with a constant step size in practice and has demonstrated spectacular successes in machine learning tasks.

The current practice is to set the step size inversely proportional to N where N is the number of training samples. As N becomes large, we show that the SGLD algorithm has an invariant probability measure which significantly departs from the target posterior and behaves like as Stochastic Gradient Descent (SGD). This difference is inherently due to the high variance of the stochastic gradients.

Several strategies have been suggested to reduce this effect; among them, SGLD Fixed Point (SGLDFP) uses carefully designed control variates to reduce the variance of the stochastic gradients. We show that SGLDFP gives approximate samples from the posterior distribution, with an accuracy comparable to the Langevin Monte Carlo (LMC) algorithm for a computational cost sublinear in the number of data points.

We provide a detailed analysis of the Wasserstein distances between LMC, SGLD, SGLDFP and SGD and explicit expressions of the means and covariance matrices of their invariant distributions. Our findings are supported by limited numerical experiments.

A. Eberle, Couplings and error bounds for MCMC methods

Many Markov Chain Monte Carlo methods on Euclidean state spaces can be seen as discretizations of diffusion processes. For example this is the case for Random Walk Metropolis, the unadjusted and the Metropolis adjusted Langevin Algorithm, and Kinetic Langevin Monte Carlo. Other methods such as Hamiltonian Monte Carlo share at least structural similarities with related diffusion processes. Therefore, a natural way to derive error bounds is to carry over techniques from the diffusion case. Here, often rather precise bounds can be derived by combining couplings (e.g. synchronous, reflection or sticky couplings) with contraction prop-

erties in appropriate Wasserstein distances. In this overview talk, I will explain how to derive well-behaved bounds in the diffusion case and how to carry them over to MCMC methods.

I. Gentil, A short review on the Schrödinger problem via Otto calculus

We propose a review on the Schrödinger problem with an analytic point of view. We explain how the Schrödinger problem approach the Wasserstein distance and how it is linked to the flow gradients.

A. Guillin, uniform in time propagation of chaos in non convex potential

We will use here simple coupling techniques to provide the first uniform in time propagation of chaos results, when the confining potential is convex at infinity (and lower bounded curvature) and the interaction potential is not too "strong". It is a joint work with A. Durmus and A. Eberle.

A. Guyader, Rare Event Simulation for Molecular Dynamics

This talk is devoted to the presentation of algorithms for simulating rare events in a molecular dynamics context, e.g., the simulation of reactive paths. We will consider $\mathbb{R}^d$ as the space of configurations for a given system, where the probability of a specific configuration is given by a Gibbs measure depending on a temperature parameter. The dynamics of the system is given by an overdamped Langevin (or gradient) equation. The problem is to find how the system can evolve from a local minimum of the potential to another, following the above dynamics. After a brief overview of classical Monte Carlo methods, we will expose recent results on adaptive multilevel splitting techniques.

C. Hartmann, A least squares Monte-Carlo approach to rare events simulation

The talk will be devoted to the question of computing the optimal change of measure for certain classes of rare event simulation problems that appear in statistical mechanics, e.g. in molecular dynamics. The method is based on a representation of the rare event sampling problem as an equivalent (or: dual) stochastic optimal control problem, whose value function characterizes the optimal (i.e. minimum variance) change of measure. The specific duality behind the problem is then used to devise numerical algorithms for computing the optimal change of measure. I will describe one approach in detail that is built on a semi-parametric representation of the value function as the solution of a forward-backward stochastic differential equation that can be solved by least-squares Monte Carlo. I will discuss the general approach, with a particular focus on the choice of the ansatz functions and the solution of high-dimensional problems, and illustrate the numerical method with simple toy examples. (This is joint work wirh Omar Kebiri and Lara Neureither.)

J. Hénin, Adaptive sampling based on collective variables in biomolecular simulations

K. Kremer, Multiscale Modeling for Soft Matter - Perspectives and Challenges

Material properties of soft matter are governed by a delicate interplay of energetic and entropic contributions. In other words, generic universal aspects are as relevant as local chemistry specific properties. Thus many different time and length scales are intimately coupled, which often makes a clear separation of scales difficult. This introductory lecture will review recent advances in multiscale modeling of soft matter. This includes different approaches of sequential and concurrent coupling. Furthermore problems of representability and transferability will be addressed as well as the question of scaling of time upon coarse graining. Finally some new developments related to data driven methods will be shortly mentioned.

D. Le Peutrec, Most probable exit points for the overdamped Langevin dynamics

(Joint work with Giacomo Di Ges, Tony Lelivre, and Boris Nectoux)

We consider in this talk, in the small temperature regime $h \to 0^+$, the overdamped Langevin process associated with some Morse potential f on some bounded domain Ω. The infinitesimal generator of this process is given by the differential operator $L = \nabla f \cdot \nabla - \frac{h}{2}\Delta$.

When Ω is a confining well of the potential f with one unique local minimum, it is well known that the trajectories of the process, starting from any $x \in \Omega$, exit Ω in a neighborhood of $\mathrm{argmin}_{\partial\Omega} f$ with a probability tending to 1 when $h \to 0^+$.

We will try in this talk to generalize these results to more general domains Ω when the process is initially distributed according to the quasi-stationary distribution in Ω. Using "leveling" properties, our results also extend to some particular deterministic initial conditions.

J. Mattingly, Approximate MCMC

The Markov Chain Monte Carlo method is the dominant paradigm for posterior computation in Bayesian analysis. It is common to control computation time by making approximations to the Markov transition kernel. Comparatively little attention has been paid to computational optimality in these approximating Markov Chains, or when such approximations are justified relative to obtaining shorter paths from the exact kernel. We give simple, sharp bounds for uniform approximations of uniformly mixing Markov chains. We then suggest a notion of optimality that incorporates computation time and approximation error, and use our bounds to make generalizations about properties of good approximations in the uniformly mixing setting. The relevance of these properties is demonstrated in applications to a minibatching-based approximate MCMC algorithm for large data logistic regression and low-rank approximations for Gaussian processes.

L. Michel, Semiclassical methods for Langevin dynamics at low temperature

Semiclassical methods have shown to be very efficient to get quantitative description of metastability of Langevin dynamics. In this talk we try to explain the main ideas of this approach in both reversible and non-reversible cases.

M. Michel, From cluster algorithms to PDMP algorithms: a Monte Carlo story of symmetry exploitation

During this talk, I will present how the development of non-reversible algorithms by piecewise deterministic Markov processes (PDMP) was first motivated by the impressive successes of cluster algorithms for the simulation of lattice spin systems. I will especially stress how the spin involution symmetry crucial to the cluster schemes was replaced by the exploitation of more general symmetry, in particular thanks to the factorization of the energy function.

L. Miclo, Construction of set-valued dual processes via random mappings

The strong stationary times introduced by Aldous and Diaconis [1986] provide a probabilistic approach to quantitative convergence to equilibrium. They are often obtained as the absorption times of intertwining dual processes, following a method due to Diaconis and Fill [1990]. We will see how explicit constructions can be deduced from certain random mappings, related to the coupling-from-the-past algorithm of Propp and Wilson [1996] and to the evolving sets of Morris and Peres [2005]. This approachis very flexible and can be adapted, via random flows, to diffusions to extend Pitman's theorem [1975] on the intertwining relation between the Brownian motion and the Bessel-3 process.

P. Monmarché, Adaptive Biasing Force method and tensor approximation

The Adaptive Biasing Force (ABF) method aims to sample metastable Gibbs measure by learning on-the-fly the derivative of the free energy associated to some reaction coordinates and then using the latter as a biasing force. The motivation is that, after biasing, at equilibrium, the reaction coordinates are uniformly distributed on a chosen compact set, which means that the energy barriers have been flattened and that all the values of the reaction coordinates have been explored. Nevertheless, this requires to keep in memory a function of the reaction coordinates, which restrain the classical method to a small number of those. This can be circumvented by keeping in memory, instead of the free energy, an approximation of it as a sum of tensor products of one-dimensional functions. This is a joint work with Virginie Ehrlacher and Tony Lelivre.

M. Ottobre, Diffusion theory in absence of the Hörmander condition

In 1968 Hörmander introduced a sufficient condition to ensure hypoellipticity of second order partial differential operators. As is well known, this seminal work of Hörmander had deep repercussions in PDE and probability theory as well as in numerical and applied analysis. In this talk we will first review the Hörmander condition by an analytical, probabilistic and geometric perspective. We then present a condition which is weaker than the Hoermander condition, the so-called UFG condition. Kusuoka and Strook showed that it is still possible to build a solid PDE theory for diffusion semigroups even in absence of the Hormander condition. We will explain the significance of the UFG condition by several points of view and present new results (the first of this type) on the geometry and long time behaviour of diffusion semigroups that do not satisfy the Hormander condition. We will emphasize that UFG-diffusions constitute a large class of SDEs which i) exhibit multiple equilibria, and one is nonetheless able to determine the basin of attraction of each invariant measure; ii) has a

potential to be exploited to improve the performance of simulations. This is a joint work with T. Cass, D. Crisan and P. Dobson.

D. Perez, Parallel Approaches to Long-time Atomistic Simulations: challenges and opportunities at large computational scales

Molecular Dynamics (MD) is a workhorse of computational materials science. Indeed, MD can in principle be used to obtain any thermodynamic or kinetic quantity for a given interatomic potential. This enviable quality however comes at a steep computational price, limiting the system sizes and simulation times that can be achieved in practice. While the size limitation can be efficiently addressed using spatial decomposition strategies, the same approach cannot extend the timescales much beyond microseconds. This is a significant issue, as this implies that the constant increase in the computing power delivered by leadership-scale machines cannot be leveraged to make long-time predictions.

In this talk, we discuss additional parallelization strategies that can be used to address this timescale limitation for systems that evolve through rare transitions, concentrating on the Parallel Trajectory Splicing (ParSplice) method. We show how, taken together, these ideas can significantly extend the simulation space accessible to MD. We also discuss challenges that have to be addressed in order to fully exploit the current and future generations of massively-parallel hardware.

P. Plechac, Approximation of quantum observables by ab initio molecular dynamics

The standard results show that the ab initio molecular dynamics on the electron ground state approximates quantum observables in the canonical ensemble when the temperature is low compared to the first electron eigenvalue gap. In this talk we discuss connection between ab initio molecular dynamics and quantum observables using the tools of semi-classical analysis and Weyl quantization. We present error analysis of such approximations for quantum observables in nuclei-electrons systems. We derive a certain weighted average of different dynamics that approximates quantum observables at any temperature provided the electron eigenvalue surfaces do not cross. (joint work with A. Kammonen, M. Sandberg, A. Szepessy)

M. Prapotnik, Open Boundary Molecular Dynamics a DNA molecule in a hybrid explicit/implicit salt solution

The electrolyte concentration of the aqueous bathing environment can profoundly affect the behavior of biomolecules. Nevertheless, due to computational limitations, molecular simulations of biophysical systems are usually performed either at nominally zero salt concentration or at excessive salt concentrations. In this talk, I will present an efficient molecular simulation approach for an atomistic DNA molecule at realistic physiological ionic conditions. The simulations are performed by employing the open boundary molecular dynamics (OBMD) method that allows for simulation of open systems that can exchange mass, momentum, and energy with the environment. In our approach, the computational burden is drastically alleviated by embedding the DNA molecule in a mixed explicit/implicit salt-bathing solution. In the explicit domain, the water molecules and ions are both overtly present in the system, whereas

in the implicit water domain, only the ions are explicitly present and the water is described as a continuous dielectric medium. Water molecules are inserted and deleted into/from the system in the intermediate buffer domain that acts as a water reservoir to the explicit domain, with both water molecules and ions free to enter or leave the explicit domain. The presented approach is general and allows for efficient molecular simulations of biomolecules solvated in bathing salt solutions at any ionic strength condition.

J. Reygner, Central Limit Theorem for stationary Fleming-Viot particle systems in finite spaces

We consider the Fleming-Viot particle system associated with a continuous-time Markov chain in a finite space. Assuming irreducibility, it is known that the particle system possesses a unique stationary distribution, under which its empirical measure converges to the quasi-stationary distribution of the Markov chain. We complement this Law of Large Numbers with a Central Limit Theorem. Our proof essentially relies on elementary computations on the infinitesimal generator of the Fleming-Viot particle system, and involves the so-called π-return process in the expression of the asymptotic variance. Our work can be seen as an infinite-time version, in the setting of finite space Markov chains, of recent results by Cérou, Delyon, Guyader and Rousset.

This is a joint work with Tony Lelivre and Loucas Pillaud-Vivien.

M. Rousset, Central Limit Theorem for general Fleming-Viot particle system

In this talk, we will detail the proof of the central limit theorem (CLT) for general Fleming-Viot particle systems. The latter is based on CLT for martingale, and a careful stochastic analysis of the classical martingales arising from Fleming-Viot systems. The result includes the case of particles killed by hard obstacles, and can be applied to various splitting Monte Carlo algorithms such as Adaptive Multilevel Splitting that can be recast as a Fleming-Viot system.

The talk will be complementary to the talk of Arnaud Guyader.

G. Simpson, Sampling on Rough Energy Landscapes

I will discuss some challenges and results for sampling on rough energy landscapes. This will include numerical simulation results along with analytical results in model problems.

K. Spiliopoulos, Mean Field Analysis of Neural Networks

Machine learning, and in particular neural network models, have revolutionized fields such as image, text, and speech recognition. Today, many important real-world applications in these areas are driven by neural networks. There are also growing applications in engineering, robotics, medicine, finance and protein folding. Despite their immense success in practice, there is limited mathematical understanding of neural networks. In this talk, I will illustrate how neural networks can be studied via stochastic analysis, and develop approaches for addressing some of the technical challenges which arise. We analyze one-layer neural networks in the asymptotic regime of simultaneously (A) large network sizes and (B) large numbers of

stochastic gradient descent training iterations. We rigorously prove that the empirical distribution of the neural network parameters converges to the solution of a nonlinear partial differential equation. In addition, a consequence of our analysis is that the trained parameters of the neural network asymptotically become independent, a property which is commonly called "propagation of chaos". Our theoretical results are applied to study the evolution of the parameters for the well known MNIST dataset where an interesting bimodal distribution arises.

T. Swinburne, Uncertainty-driven construction of Markov models from accelerated molecular dynamics

A common way of representing the long-time dynamics of materials is in terms of a Markov chain that specifies the transition rates for transitions between metastable states. This chain can either be used to generate trajectories using kinetic Monte Carlo, or analyzed directly, e.g., in terms of first passage times between distant states. While a number of approaches have been proposed to infer such a representation from direct molecular dynamics (MD) simulations, challenges remain. For example, as chains inferred from a finite amount of MD will in general be incomplete, quantifying their completeness is extremely desirable. Second, making the construction of the chain as computationally affordable as possible is paramount. In this work [1], we simultaneously address these two questions. We first quantify the local completeness of the chain in terms of Bayesian estimators of the yet-unobserved rate, and its global completeness in terms of the residence time of trajectories within the explored subspace. We then systematically reduce the cost of creating the chain by maximizing the increase in residence time against the distribution of states in which additional MD is carried out and the temperature at which these are respectively carried out. Using as example the behavior of vacancy and interstitial clusters in materials, we demonstrate that this is an efficient, fully automated, and massively-parallel scheme to efficiently explore the long-time behavior of materials. We also show how accommodation of exchange, rotation, reflection and translation symmetries can massively enhance sampling efficiency.

[1] TD Swinburne and D Perez, Self-optimized construction of transition rate matrices from accelerated atomistic simulations with Bayesian uncertainty quantification, Physical Review Materials 2018

E. Vanden-Eijnden, Machine learning, particle systems, and scientific computing

The methods and models of machine learning (ML) are rapidly becoming de facto tools for the analysis and interpretation of large data sets. The ability to synthesize and simplify high-dimensional data raises the possibility that (deep) neural networks may also find applications in scientific computing: indeed, ML techniques are increasingly being tested e.g. in the context of quantum physics, molecular dynamics simulation, PDE solving, etc. These studies rely on the assumption that neural networks are capable to accurately represent high dimensional functions: The performance of neural networks on high-dimensional data distributions suggests that it may be possible to parameterize a representation of a given highdimensional function with controllably small errors, potentially outperforming standard interpolation methods that are limited by the curse of dimensionality. In this talk I will present rigorous results about the representation error and trainability of neural networks,

obtained by mapping the parameters of the neural network to a system of particles relaxing with an interaction potential determined by the loss function and analyzing the empirical distribution of these parameters/particles when the network size is large. I will also discuss what these results imply for applications in scientific computing.

T. van Erp, Transition path sampling approaches for the study of rare events

Since the late nineties, Transition Path Sampling (TPS) emerged as an interesting approach to increase the time scale of molecular dynamics (MD) [1]. In the essence, it combines Monte Carlo and MD by generating many short MD trajectories via a detailed-balance Metropolis-Hastings algorithm. By requiring certain conditions on the sampled paths (e.g. start- and end-conditions), defining the 'path ensemble', a large number of rare transitions can be generated in a relatively short simulation time. Examples of such transitions are chemical reactions, protein folding and nucleation. The advantage of TPS compared to most other rare event methods is that it provides information on the spontaneous process without altering the the underlying dynamics. The approach can also be used to compute quantitative properties like rates and activation energies by combining the results of a series of path ensemble simulations. Transition Interface Sampling (TIS) and Replica Exchange TIS (RETIS) have perfectionized this approach [2]. In this talk, I will give an overview of TPS, TIS, RETIS, and related algorithms and provide an example study on water dissociation [3].

[1] C. Dellago, P. G. Bolhuis, F. S. Csajka, and D. Chandler, Transition path sampling and the calculation of rate constants, J. Chem. Phys. 108, 1964 (1998)
[2] R. Cabriolu, K. M. Skjelbred, P. G. Bolhuis, and T. S. van Erp, Foundations and latest advances in replica exchange transition interface sampling, J. Chem. Phys. 147, 152722 (2017)
[3] M. Moqadam, A. Lervik, E. Riccardi, V. Venkatraman, B. K. Alsberg, and T. S. van Erp, Local initiation conditions for water autoionization, Proc. Natl. Acad. Sci. USA, 115, E4569-E4576, (2018).

D. Wales, Energy Landscapes: from molecules and nanodevices to machine learning

The potential energy landscape provides a conceptual and computational framework for investigating structure, dynamics and thermodynamics in atomic and molecular science. This talk will summarise new approaches for global optimisation, quantum dynamics, the thermodynamic properties of systems exhibiting broken ergodicity, and rare event dynamics. Applications will be presented that range from prediction and analysis of high-resolution spectra, to coarse-grained models and design principles for self-assembly of mesoscopic structures, with recent results for machine learning landscapes.

Selected publications

[1] Perspective: Energy Landscapes for Machine Learning, PCCP, 19, 12585-12603, 2017.
[2] Feature Article: Exploring Biomolecular Energy Landscapes, Chem. Commun., in press, DOI:10.1039/c7cc02413d
[3] Perspective: Insight Into Reaction Coordinates and Dynamics From the Potential Energy Landscape, JCP, 142, 130901, 2015.
[4] Energy Landscapes: Some New Horizons, Curr. Op. Struct. Biol., 20, 3-10, 2010.
[5] Energy Landscapes, Cambridge University Press, Cambridge, 2003

W. Zhang, Properties and error analysis of the effective dynamics for diffusion processes

Projecting a high-dimensional stochastic dynamics on the reaction coordinate space has attracted considerable attentions in the literature. In this poster, I will present some recent work on the study of the effective dynamics, which provides a low-dimensional representation of the original diffusion process along a given reaction coordinate. I will discuss the properties inherited by the effective dynamics, as well as some error approximation results. Numerical algorithms will be also briefly mentioned.

This presentation is based on joint work with Carsten Hartmann, Christof Schütte, and Tony Lelièvre.

[1] W. Zhang and Ch. Schütte, Reliable approximation of long relaxation timescales in molecular dynamics, Entropy, 2017.
[2] W. Zhang and C. Hartmann and Ch. Schütte, Effective dynamics along given reaction coordinates and reaction rate theory, Faraday Discuss., 195, 365-394, 2016.
[3] T. Lelièvre and W. Zhang, Pathwise estimates for effective dynamics: The case of nonlinear vectorial reaction coordinates, arXiv preprint 1805.01928

P.-A. Zitt, Ergodicity of the zigzag process

The zigzag process is a Piecewise Deterministic Markov Process which can be used in a MCMC framework to sample from a given target distribution. Using the 'Meyn-Tweedie' approach, we derive convergence results and a CLT for empirical means, even when the noise is 'minimal' in an appropriate sense. The main difficulty is to prove irreducibility; we will give the key ideas leading to this proof.

Abstracts of posters (alphabetical order of presenters)

M. Baudel, Time error estimation for metastable Markov processes

Markov processes with metastable sets are of interest to study chemical reactions and thermally activated processes. The adaptive multilevel splitting algorithm enables to compute reactive trajectories i.e. equilibrium trajectories leaving a metastable state and ending in another one. An estimator based on this algorithm was presented by Frdric Crou, Arnaud Guyader, Tony Lelivre and David Pommier to compute transition times from metastable state to another one. The aim of this work is to discuss and control the error between the mean transition time starting from the reactive entrance distribution and the quasi-stationary distribution. Work in progress with Arnaud Guyader and Tony Lelivre.

A. Chevallier, Efficient random walk for Wang-Landau algorithm in high-dimensional spaces

The Wang-Landau (WL) algorithm is a recently developed stochastic algorithm computing densities of states of a physical system [1, 2]. Since its inception, it has been used on a variety of (bio-)physical systems [3, 4, 5], and its convergence has been proved under reasonable assumptions [6]. The convergence speed of the algorithm is tightly tied to the connectivity properties of the underlying random walk. As such, we propose an efficient random walk that uses geometrical information to improve connectivity. Furthermore, we investigate the adverse effects of high dimensionality on the connectivity and propose a way to alleviate the problem. We show our results on selected toy systems, but also on the usual 60 dimensional benchmark system, dialanine.

[1] F. Wang and D.P. Landau. Efficient, multiple-range random walk algorithm to calculate the density of states. Physical review letters, 86(10):2050, 2001.
[2] D.P Landau, S-H. Tsai, and M. Exler. A new approach to monte carlo simulations in statistical physics: Wang-landau sampling. American Journal of Physics, 72(10):1294-1302, 2004.
[3] P. Poulain, F. Calvo, R. Antoine, M. Broyer, and P. Dugourd. Performances of wang-landau algorithms for continuous systems. Physical Review E, 73(5):056704, 2006.
[4] Pedro Ojeda-May and Martin E Garcia. Electric field-driven disruption of a native beta-sheet protein conformation and generation of a helix-structure. Biophysical journal, 99(2):595-599, 2010.
[5] A. Swetnam and M. Allen. Improving the Wang-Landau algorithm for polymers and proteins. Journal of computational chemistry, 32(5):816-821, 2011.
[6] G. Fort, B. Jourdain, E. Kuhn, T. Lelivre, and G. Stoltz. Convergence of the wang-landau algorithm. Mathematics of Computation, 84(295):2297-2327, 2015.

Q. Du, Variance estimation for Adaptive Multilevel Splitting

Adaptive Multilevel Splitting (AMS), also called Subset Simulation, is a generic Monte Carlo method to simulate realizations of a rare event. It is already known that the mean field particle system estimator is consistent and satisfies a Central Limit Theorem. The goal of this presentation is to propose and prove the consistency of two estimators of the asymptotic variance that appears in the latter CLT, one of which is the same as the variance estimator proposed by Lee and Whiteley for Particle Filters. These estimators are constructed by tracing the genealogical informations of the associated particle system, which only requires a single simulation. The analysis is heavily based on the study of some coalescent trees-typed measures of the Feynman-Kac Interacting Particle Systems (IPS), that can also be easily adapted to some more general frameworks like Adaptive Sequential Monte Carlo (ASMC). This is a joint work with Arnaud Guyader.

G. Ferré, Adaptive sampling of large deviations

Computing the large deviation functions associated to additive functionals of Markov processes is an important problem in statistical physics. Indeed, such functions help estimating probabilities and pathways of rare events, and generalize the concept of free energy and entropy to non-equilibrium systems. They are however difficult to estimate in practical situations. We propose here an algorithm inspired by the works of Borkar and collaborators that learns on the fly a driven process realizing a given fluctuation, and provides a direct estimator of the large deviation functions.

F. Hedin, The Generalized Parallel Replica dynamics for the long time simulation of metastable biochemical systems

Metastability is one of the major encountered obstacle when performing long molecular dynamics simulations, and many methods were developed to address this challenge. The "Parallel Replica" (ParRep) dynamics is known for allowing to simulate very long trajectories of metastable Langevin dynamics in the materials science community, but it relies on assumptions that can hardly be transposed to the world of biochemical simulations. The later developed "Generalized ParRep" variant solves those issues, but it was not applied to significant systems of interest so far.

In this work, we present the first publicly available implementation of the Generalized Parallel Replica method, targeting frequently encountered metastable biochemical systems, such as conformational equilibria or dissociation of protein-ligand complexes. It will be shown that the resulting C++ implementation exhibits a strong linear scalability, providing up to 70% of the maximum possible speedup on several hundreds of CPUs.

L. Lopes, Simulating rare events in molecular dynamics with the Adaptive Multilevel Splitting

Simulation of rare events has been an important field of research in biophysics for nearly two and a half decades now. A usual quantity of interest is the transition rate, or equivalently its inverse, the transition time. The Adaptive Multilevel Splitting (AMS) is a powerful and versatile method to estimate rare events probabilities [1]. The idea of the algorithm is to split the phase space into cells and calculate the probability to pass from one cell to another.

The positions of the intermediate interfaces, which are used to split reactive trajectories, are adapted on the fly in order to reduce the variance of the estimation of the probability of the rare event of interest [2,3]. In this study we apply the AMS method to two distinguished unbinding cases. The first with the -cyclodextrin and two different ligands for which there are published experimental transition rates to compare to [4]. The second is a with the HSP90 protein and a drug candidate, part of a collaboration with Sanofi.

[1] F. Cérou and A. Guyader, Stoch. Anal. Appl. 25, 417 (2007)
[2] C.-E. Bréhier, M. Gazeau, L. Goudenège, T. Lelièvre and M. Rousset, Ann. Appl. Probab. 26, 3559 (2016)
[3] C.-E. Bréhier, T. Lelièvre and M. Rousset, ESAIM: PS 19, 361 (2015)
[4] Zhang, X., Gramlich, G., Wang, X., & Nau, W. M., JACS 124(2), 735 (2002).

B. Nectoux, Exit point distribution of the overdamped Langevin process in the small temperature regime

We present recent results on the exit point distribution of the overdamped Langevin process in the small temperature regime. More precisely, motivated by the desire to analyse the metastability of the exit event from a bounded domain Ω, we investigate the more probable places of exit of Ω as well as their relative probabilities when the process is initially distributed according to the quasi-stationary distribution of the process in Ω (the case of deterministic initial conditions within Ω is also investigated). Moreover, in the setting we consider, the potential function f has several critical points in Ω and their energies can be higher than the minimum value of f on the boundary of Ω. Such critical points can lead to significant changes in the concentration of the exit point distribution.

The proof is based on semi-classical analysis. The results obtained extend, as far as the reversible case is concerned, the results obtained by Kamin, Perthame, Day, and Wentzell-Freidlin, to the case when f has several critical points in Ω.

This is a joint work with Giacomo Di Gesú, Tony Lelièvre, and Dorian Le Peutrec.

L. Neureither, Averaging and conditional expectations: Some aspects of a comparison

Averaging and conditional expectations (also known as effective dynamics) are well-known methods which yield low dimensional approximations for the dynamics of interest (which are usually given by the dynamics of the slow degrees of freedom). In this work, we analyse differences between these approaches, which arise due to irreversibility of the underlying dynamics. Further, we give convergence results for the effective dynamics in the setting in which averaging applies and formulate conditions under which the two approaches agree in the timescale separation limit.

F. Nüske, Spectral Properties of Projected Dynamics

Diffusion processes are widely used to model complex phenomena, e.g. in finance, climate research, or in molecular dynamics simulations of biomolecules. Simulating these systems at full resolution is often prohibitively costly and results in huge amounts of data that are difficult to analyze. Recently, a systematic way of obtaining effective dynamics on reduced coordinates has been proposed [1,2]. The coordinates can be any smooth transformation of the original

state space, and the dynamics are obtained as averages over level sets of the transformation. Here, we present theoretical and numerical results on the approximation quality of the effective dynamics for reversible diffusions [3]. We also discuss parameter estimation for the effective dynamics from simulation data of the original process, and show how physical interpretability of the parameters can be enforced using additional sparsity constraints [4].

[1] Legoll and Lelièvre, Nonlinearity (2010)
[2] Zhang, Hartmann, and Schütte, Faraday Discuss. (2016)
[3] Nüske, Koltai, Boninsegna, and Clementi (submitted to J. Nonlinear Sci.)
[4] Boninsegna, Nüske, and Clementi, J. Chem. Phys. (2018)

N. Nüsken, Constructing sampling schemes via coupling: Markov semigroups and optimal transport

We develop a general framework for constructing and analysing coupled Markov chain Monte Carlo samplers, allowing for both (possibly degenerate) diffusion and piecewise deterministic Markov processes. For many performance criteria of interest, including the asymptotic variance, the task of finding efficient couplings can be phrased in terms of problems related to optimal transport theory. We investigate general structural properties, proving a singularity theorem that has both geometric and probabilistic interpretations. Moreover, we show that those problems can often be solved approximately and support our findings with numerical experiments. For the particular objective of estimating the variance of a Bayesian posterior, our analysis suggests using novel techniques in the spirit of antithetic variates. This is joint work with Greg Pavliotis.

L. Richter, Optimal importance sampling using stochastic control

We investigate importance sampling of stochastic processes and outline its connections to optimal control and large deviations theory. One can formally identify an optimal change of measure, however, especially in high dimensions clever numerical strategies are necessary to reduce variance of Monte Carlo estimates effectively. We review some iterative and approximate approaches and outline how the problem can be seen from different perspectives.

J. Rolland, From models of wall flows to three dimensional Couette flow in the study of extremely rare collapse and build up of transitional wall turbulence

Wall flows transit to turbulence in a peculiar manner when the Reynolds number (the ratio of advection over viscosity) is increased: the laminar baseflow remains linearly stable up to large Reynolds numbers (actually for all Reynolds numbers in many configurations), so that turbulent flow is a distinct state which has to be triggered by finite amplitude perturbations. Once it is triggered, turbulence can sustain itself without forcing by extracting energy from the baseflow for durations strongly dependent on control parameters. As a consequence, one finds multistability between the laminar baseflow and several configuration of wall turbulence, with exponentially decaying transition rates.

The multistability events (reactive trajectories) and their rate of occurrence are calculated numerically in a hierarchy of systems (a two degrees of freedom model by Dauchot & Manneville [1], a one dimensional model by D. Barkley [2] and in three dimensional plane

Couette flow) using a dedicated method, Adaptive Multilevel Splitting, which exponentially accelerates the sampling [3]. Among other things, this corroborates and greatly extends a scaling law (see [4]) in Reynolds number R and size L for the mean first passage time before collapse of turbulence $T \asymp \exp(L(AR - B))$ [5] (with A and B two positive constants). This scaling law can be derived in the Dauchot & Manneville model using non equlibrium statistical physics approaches. The derivation points toward the mechanisms of turbulence collapse. This also stresses on the fruitful description of wall turbulence in terms of large deviations in the infinite size limit (for extreme fluctuations as well as typical pdfs). Noise induced paths going all the way from laminar to turbulent flow in three dimensional plane Couette flow are calculated for the first time. The study of these paths indicates very peculiar reactive trajectory structure and duration. These are most likely caused by the maze of saddle points and heteroclinic connections lying between laminar and turbulent flow in phase space.

[1] O. Dauchot, P. Manneville, Local versus global concepts in hydrodynamic stability theory, J. Phys. 2, 371-389 (1997).
[2] D. Barkey, Theorical perspective on the route to turbulence n a pipe, J. Fluid Mech. 803 (2016).
[3] F. Cérou, A. Guayder, Adaptive multilevel splitting for rare events analysis, Stochastic Analysis and Application 25, 417-443 (2007).
[4] T. Schneider, B. Eckhardt, Lifetime statistics in transitional pipe flow, Phys. Rev. E 78 046310 (2008).
[5] J. Rolland, Extremely rare collapse and build up of turbulence in stochastic models of transitional wall flows, Phys. Rev. E 97, 023109 (2018).

J. Roussel, A Perturbative Approach to Control Variates in Molecular Dynamics

We propose in [1] a general variance reduction strategy for diffusion processes. Our approach does not require the knowledge of the measure that is sampled, which may indeed be unknown as for nonequilibrium dynamics in statistical physics. We show by a perturbative argument that a control variate computed for a simplified version of the model can provide an efficient control variate for the actual problem at hand. We illustrate our method with numerical experiments and show how the control variate is built in practice.

[1] J. Roussel and G. Stoltz. A perturbative approach to control variates. arXiv:1712.08022, 2017.

U. Sharma, An inequality connecting entropy distance, Fisher Information and large deviations

The connection between relative entropy and Fisher Information in the context of Markov processes is classical the Fisher Information is defined as the time derivative of the relative entropy of two solutions to the Markov process under consideration. In this work, we present a new inequality which quantifies the evolution of relative entropy, when the two measures being compared do not solve the same equation. This is joint work with Bastian Hilder, Mark Peletier and Oliver Tse.

Z. Trstanova, Sampling strategies and diffusion maps

The main challenge for sampling Boltzmann-Gibbs distributions comes from the high dimensionality of the system and complicated (metastable) energies. I will focus on Langevin dynamics and diffusion maps. Diffusion maps are a dimension reduction technique that can provide an approximation of the generator of Langevin dynamics. This approximation can serve as an automatic tool for exploration of local geometry of the underlying manifold. I will explain how this strategy can accelerate sampling and highlight these ideas by numerical simulations.

T. Wöhrer, Sharp decay estimates in defective evolution equations with uncertainty

We look at the Lyapunov functional method for linear ODEs and give an explicit construction of such functionals that yields sharp decay estimates, including defective ODE systems. As an application, we consider two kinetic evolution equations on the torus, namely the two velocity BGK model and the linear convection-diffusion equation. Adding an uncertainty parameter to these equations and analyzing its linear sensitivity leads to defective ODE systems in Fourier space. By applying the Lyapunov functional construction, we prove sharp long time behavior of exponential order multiplied by a polynom in time. The appearance of the uncertainty parameter in the two applications makes it important to have decay estimates that are uniform in the non-defective limit.