Understanding the mechanistic details underlying DNA compaction is key to the study of biological processes. In order to fit in the cell nucleus, the DNA, which, if unrolled, measures 2 meters in length, rolls itself around proteins, the histones, to form nucleosomes, which are the building blocks of chromatin. While the first level of compaction, the nucleosome, has been studied in atomic resolution, the topology of the chromatin fibre remains debated. Therefore, there is a growing need for studies featuring a synergy between theoretical and experimental approaches, and the cross-validation of respective results. We propose a mesoscale bottom-up model: starting from the all-atom human nucleosome crystal structure, we aim to build a coarse-grained approach where nucleosomes and linker DNA are each represented as a limited number of interacting centres. We parameterize our model with all-atom Molecular Dynamics (MD) and energetic estimates derived from the non-linear Poisson-Boltzmann equation. We combine a map of nucleosome core particle (NCP) and linker DNA non-bonded energies relative to different positioning and orientations with bonding forces obtained from all-atom MD simulations on NCPs, linker DNA, and NCP with linker DNA. Our mesoscale force-field envisions three main kinds of interaction: mechanical, desolvation and electrostatic. This simplified model allows the study of larger scales of chromatin organization, beyond the oligonucleosome level. We optimize our mesoscale force-field through a multi-objective optimization algorithm, using both energies and forces as inputs. The algorithm optimizes parameters employing statistical learning through linear ridge regression and differential evolution. We perform accompanying and complementary experiments, focusing on non-invasive methods that do not require aggressive sample preparation that could perturb the topology. We validate intermediate results comparing them to existing oligonucleosome models in literature and to our own experiments, aimed at providing structural and thermodynamic information. Overall, we propose a cohesive mesoscale bottom-up model, combining simulations and experiments, thus testing different hypotheses in order to shed light on the determinants of chromatin conformation.
**Panagiota Birmpa**

*Quantification of mesoscopic and macroscopic Fluctuations in interacting particle systems*

We study the most probable way an interface moves on a macroscopic scale from an initial to a final position within a fixed time in the context of large deviations for a stochastic microscopic lattice system of Ising spins with Kac interaction evolving in time according to Glauber dynamics. Such interfaces separate two stable phases of a ferromagnetic system and in the macroscopic scale are represented by sharp transitions. It is natural to ask for the corresponding large deviations result, i.e., for the probability of macroscopic interfaces evolving differently from the deterministic limit law. This is particularly interesting when studying metastable phenomena of transitions from one local equilibrium to another as one needs to quantify such large deviations which cannot be captured by the deterministic evolution. We derive quantitative estimates for the upper and the lower bound of the cost functional that penalizes all possible deviations and obtain explicit error terms which are valid also in the macroscopic scale. Furthermore, by finding the minimizers of this cost functional for the macroscopic motion of the interface in a fixed time, we prove that the probability of such events can concentrate on nucleations should the transition happen fast enough.

**Grégoire Ferre**

*Adaptive algorithm for sampling large deviation functions*

Since its introduction in the 70’s, large deviation theory has become a prominent tool in the study of scaling problems in physical systems. The two main functions of the theory, the cumulant function an the rate function, respectively generalize free energy and entropy for non-equilibrium systems. Information on these functions provide insight on the system at hand. In particular, the functions associated to fluctuations in time of ergodic averages, which characterize a form of phase transition, are interesting in applications. Such functions are in general very difficult to compute by direct Monte Carlo simulation, since an efficient sampling requires simulating an exponential number of trajectories with respect to time. This led to designing specific algorithms for this problem. I present in this poster an importance sampling strategy for reducing the variance of standard estimators. The originality of this work, realized with Hugo Touchette, resides in building on the fly estimators of an optimal drift, the cumulant function, its derivative, and the rate function at once. We provide relevant applications on equilibrium and non-equilibrium toy examples from statistical physics.

**Hamid Ghasemi**

*Predicting failure of graphene in biaxial loading using machine learning*

Artificial neural networks (ANN) and support vector machines (SVM) are used to develop models that predict whether a defect-free sheet of graphene will fail under a given set of transverse stresses and at a given temperature. The data required to develop the predictive models are produced by carrying out a large set of molecular dynamics (MD) simulations at five temperatures of 100 K, 300 K, 500 K, 1000 K, and 1500 K. Stresses along the principal edge configurations of graphene (i.e. armchair and zigzag directions) are first normalized to the corresponding uniaxial ultimate strength values at a given temperature. The combinations of normalized stresses resulting in the failure of graphene are then used to train the ANN and SVM models. The results show that the failure of graphene is considerably influenced by the
temperature applied. They also indicate that the ANN model exhibits either the same or slightly better accuracy than the SVM model in accurately predicting the failure of graphene.

**Anton Martinsson**

*Simulated Tempering Method in the Infinite Switch Limit with Adaptive Weight Learning*

We investigate the theoretical foundations of the simulated tempering method and use our findings to design efficient algorithms. We demonstrate that the most efficient approach to simulated tempering is to vary the temperature infinitely rapidly. In this limit, we can replace the equations of motion for the temperature and physical variables by averaged equations for the latter alone, with the forces rescaled according to a position-dependent function defined in terms of temperature weights.

**Frank Pinski**

*More may be different, but is infinitely more inherently different?*

We consider the Brownian dynamics of a particle moving in an external potential. The particle path is constrained at the beginning of the path to be in one well, and after a time $T$, to be in a different well. The probability of such doubly constrained paths is described by the Onsager-Machlup (OM) functional. In a one-dimensional example and for thermodynamically-allowed paths, we show, using a finite number ($N_t$) of time mesh points, that the OM functional takes on the same range of values. Although the OM functional is not constant, it is flat in that the mean ($\frac{1}{2}N_t$) and variance ($\frac{1}{2}N_t$) are the same, independent of the constraints. We then consider the Ito-Girsanov expression, which is the continuous-time limit of the OM functional. Unlike the OM functional, this expression seems to indicate that some paths are more probable than others. As we show, this suggestion is not correct. Looking at the limiting behavior as one approaches the continuous-time limit: being proportional to $N_t$, not only does the value of the OM functional diverge but so does its variance. We conclude that the Ito-Girsanov expression only describes Brownian paths in an infinite-dimension path space. When used in computer algorithms where a finite representation of the path is used, the expression can produce manifestly unphysical results.

**Tom 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. I will talk about some recent work [1] to 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

**Tiffany Vlaar**

*Introducing sparsity in neural networks by enhanced exploration*

Neural networks are well-known for their tendency to overfit, and for being overparametrised. They can often be sparsified significantly without losing accuracy. Machine learning practitioners have therefore developed various techniques to compress neural networks, such as neural network pruning or student-teacher learning. This saves computational resources and increases their models' ability to generalise, by limiting overfitting. We propose a sampling technique inspired by methods from molecular dynamics that introduces sparsity in neural networks during training.