Abstract

At first glance, the two topics addressed in each Part of the present PhD thesis may seem orthogonal: indeed, generative models and linear response theory appear to have little overlap. However, many interesting research topic are linked by common themes after a deep analysis. In the present case, the central narrative thread revolves around Nonequilibrium Statistical Physics. The first fundamental tool from that field that has been used in the present work is Jarzynski identity (C. Jarzynski, 1997); it provides a connection between microscale and macroscale, relating microscopic work along trajectories and free energy, respectively. On the other hand, Onsager reciprocal relations (ORR) represent a milestone in that area (L. Onsager, 1931): they serve as a bridge between a microscopic property (time reversal symmetry) and a macroscopic one (response tensors).

In the first Part, we show how *recent* theoretical results in Statistical Physics can be very instrumental in state-of-the-art applications; generative model represent a substantial research challenge since they are already used in everyday life, even if we are far from having a complete theoretical picture about them. In a nutshell, we propose a novel training algorithm for Energy-Based Models (EBMs), which is a class of diffusion generative models strongly inspired by Statistical Physics, namely by Boltzmann-Gibbs ensemble; in light of this relation, a key strength of EBMs compared to other models is their interpretability. Standard procedures, such as those based on Constrastive Divergence, heavily relies on approximations of the real loss objective already in an ideal setup. Because of that, the practical implementation of such methods usually requires a lot of empirical tricks, often not theoretically justified. In contrast, our proposal is exact; furthermore, no extra bias is introduced by discretization in time and the algorithm provides for free additional information on the trained EBM (i.e. the normalization constant of the trained probabilistic model). Our contribution is based on Jarzynski identity in continuous time and Annealed Importance Sampling in discrete time.

To provide insights into the structure, this section is organized into four chapters. The first chapter offers a historical introduction to generative models, focusing on Energy-Based Models (EBMs) in relation to Statistical Physics and Data Science. The second chapter covers essential technical preliminaries necessary for contextualizing our work. This includes defining EBMs and exploring their purpose, as well as their relationship with other state-of-the-art generative models such as Variational Auto-Encoders, Generative Adversarial Net-

works, Diffusion-Based Models, and Normalizing Flows. We aim for a unifying approach to highlight similarities and differences between these unsupervised models. The third chapter delves into the relationship between EBMs and the sampling problem. Given that EBM training relies on the ability to sample from a Boltzmann-Gibbs ensemble, we discuss key sampling routines such as the Metropolis-Hastings Algorithm, Unadjusted Langevin Algorithms (ULA), and Metropolis Adjusted Langevin Algorithms (MALA). In the final section, we emphasize the connection between EBMs and Statistical Physics. This serves to justify the adoption of the Boltzmann-Gibbs ensemble and provides important context for utilizing the Jarzynski identity in the main result of this thesis.
The third chapter contains the main novel theoretical result we propose about EBM training. The core idea is the use of nonequilibrium sampling, that is sequential Monte Carlo in discrete time, to efficiently compute the gradient of cross-entropy. Such quantity is necessary to perform KL divergence minimization, or equivalently maximization of log-likelihood, which is the standard approach in statistical learning. We present continuous and discrete time versions of our algorithm, as well as algorithmic aspects having particular relevancy in practical applications. In the last chapter we present experimental result to validate our theoretical findings; we investigate our training routine as opposed to standard procedures like Contrastive Divergence (CD) algorithm. We show as already for Gaussian Mixture Model, our proposal evidently outperforms CD. Similar results are obtained for real image datasets as MNIST and CIFAR-10.
In the second Part, we show that *established* theoretical findings in Statistical Physics can be still object of refinements. ORRs basically provide information on the structure response tensors; the main request for such relations to hold is canonical time reversal symmetry, i.e. the invariance of the equations of motion under the inversion of velocities. Our work demonstrates how we can relax this condition by expanding upon the definition of time reversal symmetry. This expansion enables us to prove that the set of symmetries leading to time reversal invariance is broader. The experimental validity of ORRs has been proven in many contexts where canonical time reversal seems to not hold. Thus, our result contributes to explain some of these examples. Regarding the organization of the treatment, we present the two published papers on the topic, being the second a substantial extension of the first preliminary work.