Summary

Probabilistic graphical models provide an unified framework to analyze physical systems of many interacting degrees of freedom, by combining elements of graph and probability theory. Interacting physical systems display a variety of collective phenomena, depending on the mutual dependencies between units and the consequent topological structure of their interactions, whose first analysis led to the foundation of statistical mechanics and the study of phase transitions. A paradigmatic example is given by classic spin models defined on graphs or lattices, historically developed to describe the microscopic origin of magnetism; at the same time, they provide a general description of a wide class of phenomena in several research fields, from biology, neuroscience, computer science and econophysics. In this perspective, the formalism of graphical models provides a common framework and set of methodologies to analyze interacting systems in virtually any field of pure and applied science. In general, the difficulty of analyzing high-dimensional system is that the presence of interactions makes any computation unfeasible in practice, as the volume of the configuration space grows exponentially with the system size. In this sense, the term *approximate inference* in the manuscript's title refers to the generic problem of estimating relevant features of a probabilistic graphical model, such as its marginal distributions.

The main goal of this thesis is to introduce a novel class of approximation schemes to estimate marginal probabilities on discrete (spin) models, called Density Consistency. This method shares similarities with other message-passing schemes commonly employed in statistical physics and inference, such as Belief Propagation and Expectation Propagation. The novelty introduced by Density Consistency relies on a simple way to encode approximate loop corrections coming from all the cycles in the graph, by exploiting a refined Gaussian approximation.

The structure of the manuscript resembles the research path I carried out during my three years of PhD, and it is divided in two parts. After a brief introduction presented in the first chapter, Part I focuses on the Density Consistency method: in particular, I discuss a generic derivation on probabilistic graphical models of binary degrees of freedom, using the factor graph representation. Its properties and its relations to other advanced mean field methods are discussed, and its performances are evaluated on finite size systems. Furthermore, I present an analytic theory for the ferromagnetic Ising model in the thermodynamic limit, providing a closed form expression for the critical temperature. Finally, Density Consistency is applied to the Inverse Ising problem in statistical inference and its performances are compared to other state-of-the-art techniques.

Part II of the manuscript contains a standalone chapter, in which I discuss another project I contributed to during my last year of PhD, somehow prompted by the COVID-19 pandemic. By using message-passing techniques in a Bayesian inference framework, we developed an on-line epidemic mitigation protocol in order to detect the individuals with the highest risk to be infected, starting from the knowledge about their contacts, that can be registered using digital contact tracing applications. Probabilistic inference provide a criterion to selectively isolate individuals with the highest risk, so that an effective epidemic suppression can be achieved while avoiding global containment measures, with consequent and well-known economic and social drawbacks.

This project has been carried out jointly with another group of researchers based in Paris (France) and Lausanne (Switzerland).

I summarize below the contents of each chapter and the list of co-authored papers, all of them being covered in the manuscript.

Thesis Outline

Chapter 1: **Introduction**
I present some basic notions of equilibrium statistical physics, graph theory and statistical inference, in order to provide a common background knowledge and notation used through the whole manuscript.

Part I **Loop corrections in spin models through Density Consistency**

Chapter 2: **Approximate methods in statistical physics**
I discuss several state-of-the-art approximations in statistical physics and high-dimensional inference, with a special focus on message-passing techniques such as Belief Propagation. An additional detailed derivation of Expectation Propagation is presented, that is required to better understand Density Consistency.

Chapter 3: **Density Consistency**
This is the core chapter of the manuscript, where Density Consistency is derived on graphical models of binary spins and its main properties are discussed. The contents of this chapter are included in Paper A.

Chapter 4: **Results: forward problem**
I evaluate the performances of Density Consistency in comparison to other approximations. In particular, the first part presents a series of results on finite size systems, while in the second DC is used to derive a quasi-analytic solution for the ferromagnetic Ising model in the thermodynamic limit. The contents of this chapter are included in Paper A.

Chapter 5: **The Inverse Ising problem**
This chapter focuses on the Inverse Ising Problem in statistical physics. After presenting some background motivations and formulating the problem in a Bayesian framework, I derive an approximate solution for the maximum likelihood estimator using Density Consistency, whose performances on synthetic data are compared with other state-of-the-art techniques. The contents of this chapter are included in Paper B.

Chapter 6: **Conclusions and future perspectives**
I summarize the main findings obtained so far on Density Consistency. In addition, I describe some future directions and possible applications of Density Consistency to be yet investigated.

Part II **Bayesian inference approaches for epidemic mitigation**

Chapter 7: **Bayesian inference approaches for epidemic mitigation**
This chapter describes a Bayesian-inference guided mitigation protocol for epidemic spreading processes from contact tracing data. The main method used to perform the approximate inference is Belief Propopagation; results are validated on top of a compartmental model designed to describe a realistic spreading of SARS-CoV-2 in a population. The contents of this chapter are related to Paper C.

Coauthored papers and pre-prints

Paper A Alfredo Braunstein, Giovanni Catania and Luca Dall'Asta,
Loop corrections in spin models through density consistency,
Physical Review Letters 123, 020604 (2019)
arXiv:1810.10602.

Paper B Alfredo Braunstein, Giovanni Catania, Luca Dall'Asta and Anna Paola Muntoni,
A Density Consistency approach to the inverse Ising problem,
Journal of Statistical Mechanics: theory and experiment, 033416 (2021)
arXiv:2010.13746

Paper C Antoine Baker, Indaco Biazzo, Alfredo Braunstein, Giovanni Catania, Luca Dall'Asta, Alessandro Ingrosso, Florent Krzakala, Fabio Mazza, Marc Mézard, Anna Paola Muntoni, Maria Refinetti, Stefano Sarao Mannelli and Lenka Zdeborová,
Epidemic mitigation by statistical inference from contact tracing data,
arXiv:2009.09422, accepted on Proceedings of the National Academy of Sciences