Statistical Physics of Sparse and Dense Models in Optimization and Inference

Hinnerk Christian Schmidt

NNT : 2018SACLS366

Thèse de doctorat

Statistical Physics of Sparse and Dense Models in Optimization and Inference

Thèse de doctorat de l'Université Paris-Saclay
préparée à l'Université Paris-Sud

Ecole doctorale n°564 École Doctorale Physique en Île-de-France (PIF)
Spécialité de doctorat: Physique

Thèse présentée et soutenue à Gif-sur-Yvette, le 18 Octobre 2018, par

Hinnerk Christian Schmidt

Composition du Jury :

Alberto Rosso Directeur de Recherche, Université Paris Sud (Laboratoire de Physique Théorique et Modèles Statistiques)	Président
Galen Reeves Professeur assistant, Duke University (Department of Electrical and Computer Engineering and Department of Statistical Science)	Rapporteur
Pan Zhang Associate Professeur, Chinese Academy of Sciences (Institute of Theoretical Physics)	Rapporteur
Marc Lelarge Directeur de Recherche, INRIA-ENS (Département d'Informatique)	Examinateur
Guilhem Semerjian Maître de Conférences, Ecole Normale Supérieure (Laboratoire de Physique Théorique)	Examinateur
Lenka Zdeborová Chargé de Recherche, CEA, Université Paris Saclay, CNRS (Institut de Physique Théorique)	Directeur de thèse
Florent Krzakala Professeur, Ecole Normale Supérieure (Laboratoire de Physique Statistique)	Invité

«Woran arbeiten Sie?» wurde Herr K. gefragt.
Herr K. antwortete: «Ich habe viel Mühe,
ich bereite meinen nächsten Irrtum vor.»

BERTHOLT BRECHT, MÜHSAL DER BESTEN

Acknowledgments

I express my deep gratitude to Christian Osmer, who has shaped and supported me in so many ways.

I am most thankful to my family that has encouraged and comforted me when necessary! Non of this would have been possible without the support of my parents and siblings. I would like to thank my father for having lifted my scientific curiosity and my mother for having provided me the liberty to explore with no boarders.

Since my arrival in Paris I have drawn much energy from the support of family Ammar that have provided me a surrogate family and made me feel home in a strange place. Thank you Rebecca! For accompanying me, for your understanding and care, for much distraction and laughter, your advice and for reminding me that there is so much else to care about.

I am very grateful to Lenka, who has accepted me as her PhD student besides my non existing prior knowledge on disordered systems. Thank you, Lenka, for letting me work in my own paste and way, with few barriers. You have accompanied me with much patience through my scientific explorations and pushed me back onto the right path when I got lost in the details. You and Florent have not only shared your professional life with us but also you private time; for all this many thanks. Further more special thanks to all the members of our group; those in IPhT and those in ENS!

I would like to thank my old friends Martin Thordsen, Maximilian Persch, Niklaus Olsen and Tomasz Jadamus for good times and distraction. Lukas Ophaus, Niels Linnemann and Francesco Massa for their friendship and company during my undergraduate studies and beyond.

I am happy to have made many new friends during the cause of this work. I am most grateful to Andre Manoel for very valuable scientific help and advice and for being a good friend. Specially I thank Riccardo Ravasio, Pierre Heidmann, Séverin Charbonnier, Santiago Migliaccio, Luca Rizzo and Niall Robertson for their companionship! Thibault Lesieur, Christophe Schülke, Fabrizio Antenucci, Valerio Volpati, Pierfrancesco Urbani, Eric Tramel, Alejandro Lage Castellanos, Laura Foini, Sebastian Goldt, Benjamin Aubin and Stefano Sarao for good times in Paris and Saclay as well as scientific and non-scientific advice.

Many teachers and tutors have accompanied and supported me. I would like to thank Irene D'Amico, Philip Walther, Caslav Brukner and Svetlana Gurevich for their recommendations. Giulio Biroli and Marco Schiro for their tutorship during the PhD. My contributions would

not have been possible without the many scientific discussions over the years. My deep thanks to Pierfrancesco Urbani who has taught me much and always taken time to answer my questions. I would further like to acknowledge Florent Krzakala, Guilhem Semerjian and Yan Fyodorov for discussions and help. Henry Pfister and Galen Reeves for their enlightening explanations and helpful advice. Many thanks to Pan Zhang and Galen Reeves, who have taken the time to review the thesis and provided valuable comments and questions. I am grateful to Alberto Rosso, Guilhem Semerjian and Marc Lelarge for having accepted the invitation as members of the jury and taken the time to examine it.

I would like to acknowledge the help of Sylvie Zaffanella and the rest of the staff of IPhT that have made my time at IPhT very simple. Further I acknowledge the financial support of the CEA CFR scholarship and the "IDI2015" project funded by the IDEX Paris-Saclay.

Contents

— 0 —

Introduction

Foreword

At first sight the subjects treated in this thesis do not seem to be physics problems. Among them are a constraint satisfaction problem, a dynamical process on a graph and two inference problems. However, a closer look reveals that all these problems are linked through the very fact that they contain *disorder* in one way or another and that they can be formulated in terms of a *probabilistic graphical model*. This connection reveals the strong tie to the statistical physics of disordered systems which can be exploited to examine the problems through a statistical physics approach. Such an approach is the appropriate way in which to study systems that contain too many constituents, to complex interactions or are for some other reason too complicated to be described in a meaningful microscopic fashion.

This link is easiest outlined for inference problems in which one is provided some data and tries to infer something about the underlying data generating process. Typically such data-sets are noisy and/or big; both are strong arguments for a statistical physics approach. The data plays the role of the disorder and Bayes theorem provides us the formal link to statistical physics. Denote the data $\boldsymbol{D}$ and let the generative process be described by some set of variables $\boldsymbol{G}$, then Bayes theorem provides the connection between the *posterior probability* for $\boldsymbol{G}$, given $\boldsymbol{D}$, written $P(\boldsymbol{G} \mid \boldsymbol{D})$, and the *partition function*, $Z(\boldsymbol{Y})$:

$$P(\boldsymbol{G} \mid \boldsymbol{D}) = \frac{P(\boldsymbol{D} \mid \boldsymbol{G})P(\boldsymbol{G})}{Z(\boldsymbol{D})} \,.$$

The inference problems considered will be set in the *Bayes optimal* framework, where one knows the functional form of $P(\boldsymbol{D} \mid \boldsymbol{G})$ and $P(\boldsymbol{G})$.

Throughout this thesis there are two recurring questions: (I) what is the optimal achievable performance? And (II) what are the algorithmic strategies that lead to such an optimal performance. The averaged logarithm of the partition function, $\mathbb{E} \ln Z(\boldsymbol{D})$, the free energy, is the generating function that provides the physical properties of the system. It is linked to both the above questions. The *replica method* facilitates the computation of the free energy in terms of some macroscopic order parameters that are linked to the best achievable error in the inference of $\boldsymbol{G}$ and therefore answers the first question. The *cavity method*

permits to formulate a microscopic free energy, the minimization of which (with respect to the microscopic variables), leads to a message passing algorithm that can be exploited to approach the second question. The inference problems considered are both related to *matrix factorization (or estimation)*. First, the dense Dawid-Skene model is studied, which leads to a bipartite graphical model that can be mapped onto *low-rank matrix factorization*. And secondly, a particular *symmetric extensive-rank matrix estimation* problem is studied for which it is necessary to combine the replica method with tools borrowed from *random matrix theory*.

The *constraint satisfaction problems* fall into a very similar mathematical framework. However, the problem considered is not set in the Bayes optimal framework. The physical implications are that *glassy* states may be present. The space of solutions can exhibit a difficult structure that causes the simple replica symmetric solution of the replica free energy to be no longer correct. Instead, the Parisi replica symmetry breaking scheme must be applied for a complete treatment of these problems. This also has immediate consequences for the algorithmic approach that needs to be extended as well.

The deep link to statistical physics carries beyond the analogy to inference and constraint satisfaction problems. Threshold models are common approach to model collective dynamical processes on networks and are another such example that is closely connected to the bootstrap percolation process in statistical physics. A question of major importance concerns the minimal contagious sets in such dynamical processes. This problem can be considered an optimization problem over initial conditions of the dynamical process. Or, equally, a spin-glass problem with particular interactions. Although quite different from the previous two models, it can be shown to be accessible by the very same methods [GS15]. However, the algorithmic aspects of this problem are still not well elaborated and leave open questions (a) after good local algorithms and (b) their performance.

Contributions and Organization of the Manuscript

In part I the necessary background is introduced and the previously mentioned methods are outlined. In chapter 1 an introduction to probabilistic graphical models is given. Thereafter, in chapter 2, the concepts of disordered systems are introduced, necessary to develop the connection between inference problems and statistical physics. Finally, in chapter 3, some elements from random matrix theory are collected that are necessary to resolve the extensive-rank matrix estimation problem. Part II of this thesis contains my contributions that are outlined in order of appearance in the following.

Circular Coloring

The circular coloring problem is a constraint satisfaction problem that belongs to a larger class of problems that generalizes the canonical graph coloring problem. Whereas in the canonical coloring two nodes of a graph that are connected by an edge are required to have different colors, in circular coloring the colors are ordered into a circle, and two adjacent nodes are required to have two adjacent colors.

My main contribution comprises in the analysis of the problem in the replica symmetric and one-step replica symmetry breaking framework. It revealed that circular coloring exhibits several interesting features, not common to other discrete constraint satisfaction problems. Further more it establishes non-rigorously a conjecture by Nešetřil about the 5-circular colorability of sub-cubic graphs, for random graphs from a statistical physics point of view.

This work was published in the *Journal of Statistical Mechanics: Theory and Experiment, Volume 2016, August 2016* [SGZ16]:

- Christian Schmidt, Nils-Eric Guenther and Lenka Zdeborová, "Circular coloring of random graphs: statistical physics investigation"

Dense Dawid-Skene Model (Crowdsourcing)

Consider the categorization of M data points through the contribution of N individuals and assume that each individual can be characterized by a single reliability parameter (the probability that a correct label is returned). This is the so-called Dawid-Skene model. In the DS model two regimes can be distinguish. The sparse regime, in which each individual is assigned only a sub-extensive number of questions. And the dense regime, in which each individual is assigned an extensive number of questions. Most of the previous theoretical studies of crowdsourcing have focused on the sparse limit of the DS model.

In contrast, my contribution is the closed form analysis of the DS model in the dense limit that reveals the phase diagram of the problem. This is done by showing that the dense DS model belongs to a larger class of low-rank matrix estimation problems for which it is possible to express the asymptotic, Bayes-optimal, performance in a simple closed form. Additionally, numerical results are obtained on how these results translate into the sparse regime and experiments on a real world dataset are performed.

The work was submitted to the *IEEE Transactions on Information Theory*

- Christian Schmidt, Lenka Zdeborová, "Dense limit of the Dawid-Skene model for crowdsourcing and regions of sub-optimality of message passing algorithms"

and can also be found on the arxiv [SZ18].

Contagious Sets in Random Graphs

Consider a model in which the nodes of the graph can be either in an active or inactive state and in which a node v changes state from inactive to active if more than $t_v = d_v - k$ of its neighbors are active. The number t_v is called the threshold for node v. Destroying the k-core is equivalent to the activation of the whole graph under the above threshold dynamics. The smallest such set is the minimal contagious set.

My main contribution is show that the COREHD algorithm (a local greedy strategy) translates to a random process on the degree distribution of the graph that can be tracked exactly by the derivation of the continuous limit of the random process. This leads to new rigorous upper bounds on the size of the minimal contagious sets for

random graphs with bounded degree, drawn from the configuration model. I also proposed some new heuristics that attack the problem algorithmically and outperform previously known local state-of-the-art algorithms.

The contributions have been submitted as

- Christian Schmidt, Henry D. Pfister and Lenka Zdeborová, "On Minimal Sets to Destroy the k-Core in Random Networks"

to *Physical Review E* and an eprint can be found under [SPZ18].

Extensive-Rank Matrix Estimation

Consider the simplest possible problem that falls into the category of extensive-rank matrix factorization problems. The symmetric version $\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{X}^\intercal + \boldsymbol{W}$ with a Gaußian prior on $\boldsymbol{X}$ under symmetric Gaußian noise, $\boldsymbol{W}$. If $\boldsymbol{X} \in \mathbb{R}^{N\times R}$ and $N/R = O(1)$, in the limit where $N \to \infty$, the matrix is said to have extensive rank.

I compute the Bayes-optimal error on $\boldsymbol{X}$. In the course of this computation I show that it is not possible to reduce the replica free energy to a single scalar order parameter in the form previously assumed in other works. Instead, it is necessary to keep the spectral density of a whole overlap matrix and the free energy becomes a functional of this density. In order to resolve the free energy it is necessary to solve a matrix saddle point equation that leads to a singular integral equation in the large N limit. I solve the equations and express the asymptotic error on $\boldsymbol{X}$ in terms of this solution.

This part of my work is currently in preparation to be submitted.

Notation

Abbreviations

Fig.	Figure
Chap.	Chapter
Sec.	Section or Subsection
Alg.	Algorithm
Tab.	Table
w.r.t.	with respect to
l.h.s. / r.h.s.	left/right hand side
s.t.	such that
i.i.d.	independent and identically distributed
w.h.p.	with high probability
RV	Random Variable
PD	Probability Distribution
JPD	Joint probability distribution
CLT	Central Limit Theorem
MRF	Markovian Random Field
RG	Random Graph
CSP	Constraint Satisfaction Problem
(M)MSE	(Minimum) Mean Square Error
MAP	Maximum A Posteriori
MBE	Minimum Bitwise Error
(r)BP	(relaxed) Belief Propagation
SP	Survey Propagation
AMP	Approximate Message Passing
SE	State Evolution
RS	Replica Symmetric
RSB	Replica Symmetry Breaking/Broken
CCOL	circular coloring
MC	Monte Carlo
SA	Simulated Annealing

GB	Gibbs-Boltzmann
TD	Thermodynamic

Mathematical Operators and Functions

$:=$	equal by definition
$\simeq$	asymptotically equal
$\stackrel{!}{=}$	equal by condition
x	A scalar
$\boldsymbol{x}$	A vector (x_i)
$\boldsymbol{X}$	A matrix (x_{ij})
$\mathbf{1}$	The identity matrix
$[\boldsymbol{X}]_{ij}$	x_{ij}
$\hat{\mathbf{e}}_i$	unit vector
$\|\boldsymbol{x}\|_p$	p-norm of x.
$\|\bullet\|$	The 1-norm, if $\bullet$ is a vector and the cardinality, if $\bullet$ is a set.
$\mathrm{d}\boldsymbol{x}$	$\prod_i \mathrm{d}x_i$
$\mathrm{d}\boldsymbol{X}$	$\prod_{i,j} \mathrm{d}x_{ij}$
$\mathbf{M}^\intercal$	Transpose of $\mathbf{M}$
$\mathrm{D}x$	$P_x(x)\mathrm{d}x$ and similarly for a vector/matrix
$\mathcal{G}(\mathcal{V}, \mathcal{E})$	A graph with nodes $\mathcal{V}$ and edges $\mathcal{E}$
∂a	The set of neighbors of node a in a graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$
$\Pr(\arg)$	Probability that $\arg$ is true
$\mathcal{N}(x; \mu, \sigma^2)$	$\frac{1}{\sqrt{2\pi\sigma^2}} \mathrm{e}^{-\frac{1}{2\sigma^2}(x-\mu)^2}$
$\mathbb{I}(\bullet)$	indicator function that evaluates to
$\mathbb{E}_x[\bullet]$ and $\mathbb{E}[\bullet]$	expectation w.r.t. x and expectation over all randomness
$\langle . \rangle$	average w.r.t. the GB measure
$\mathrm{Tr}\boldsymbol{M}$	Trace of the square matrix $\boldsymbol{M}$: $\sum_i m_{ii}$
$\mathrm{Re}z$ and $\mathrm{Im}z$	The real and imaginary part of z
$\fint$	Principal Value integral

Part I

Concepts and Tools

— 1 —

Probabilistic Graphical Models

The purpose of this chapter is to introduce the mathematical concepts, necessary and/or useful to understand and work with probabilistic models. The focus will then be put on Markovian random fields, which lead to a formulation of constraint satisfaction problems and dynamical processes on random graphs. Belief propagation will appear naturally as a tool to derive the average case properties of constraint satisfaction problems. These concepts can be extended to fully connected graphical models, where they can be used to study (and solve) inference problems.

Probability is an approach to access and formalize uncertainty. A probabilistic model, in that sense, is a model of an uncertain situation. To understand the concept of probabilities it is useful to consider an *experiment*. The set of all possible outcomes of the experiment is denoted the *sample space*. The elements of the sample space are referred to as *outcomes*, *states* or *realizations*. A subset of the sample space is called an *event*. The set of all *observed* outcomes is the *data*. If the experimental outcome is uncertain, then a *probabilistic model* can be introduced as an appropriate mathematical formulation of the situation.

Different questions of interest may arrive, as illustrate in Fig. 1.1. In this thesis we will be dealing with both questions depicted and this chapter aims to provide the mathematical background to deal with them. Sections 1.1 and 1.2 introduce the theory of probabilistic models and provide a background in the theory of random graphs that will be useful to describe a large class of probabilistic models. Finally, Sec. 1.3 will introduce the concepts of statistical inference.

Although *probability* is a very functional and successful approach to uncertainty, its interpretation is, ironically, much less certain. The *frequentists* approach to capture the uncertainty in our experiment is to interpret probability in terms of the frequencies of occurrences of the outcomes. This does, however, implicate the necessity of (infinite) repeatability of the experiment. On the other hand, the *Bayesian* approach adapts the point of view, that probabilities are to be interpreted as subjective beliefs about the outcome of an experiment. Subjectivity might, in that context, be understood as the lack of complete information. This eliminates the necessity of repeatability of the experiment.

Both approaches carry advantages and disadvantages. Whether one or the other is favor-

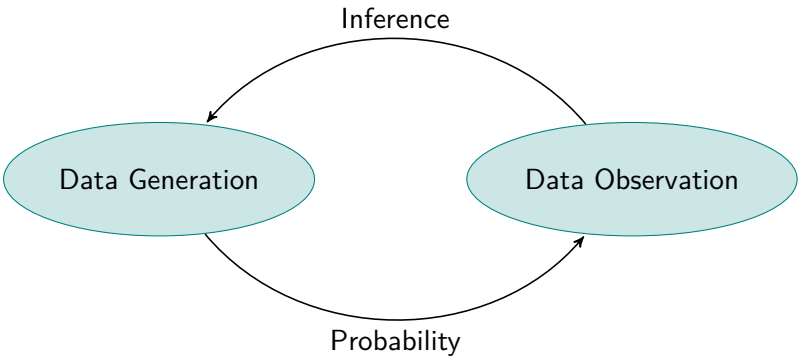

Figure 1.1: In this thesis we will investigate questions of two types. (1) Given a generative process, what are the probabilities that specific events occur. In other words, what is the likelihood that a specific dataset is observed. In the scheme we denote this direction with "probability". (2) The inverse problem deals with the "inference" of the generating process from the observed data. Adapted from Fig. 1 in [Was13].

able mostly depends on the concrete problem at hand. The frequentists approach may be applied independent of any underlying model. The Bayesian approach permits us to model the uncertainty. Utilized wisely this can be a powerful advantage that we shall resort to in this thesis.

1.1 Probabilistic Models

In order to introduce probabilistic models it is necessary to review some elementary statistics. The sample space, together with a probability distribution define a probabilistic model. Both these concepts are introduced in Sec. 1.1.1, together with some notations. Building on this we introduce Markovian random fields in Sec. 1.1.2, that will be very useful tool to represent a broad class of probability distributions. We will use them to formulate constraint satisfaction problems in a probabilistic fashion. We will see that any Markovian random field can be represented as a factor graph; This will lead to the definition of graph ensembles. Our focus will lay on pairwise Markovian random fields. In this context the the circular coloring problem is introduced, that will be studied in detail in Chap. 4 of the thesis. The pairwise Markovian random fields will lead to the definition of some particular ensembles of random graphs that will be studied. It will be shown how dynamical processes can be defined and studied on these graph ensembles. This will be useful in Chap. 6, where a dynamical process is studied that will lead to an upper bound of the size of the minimal contagious set on some class of random graphs. Belief propagation, a message passing algorithm that leads to the marginal probabilities over the nodes of a (tree-like) graph, is introduced as a tool to first study constraint satisfaction, and later inference problems. In this context, crowdsourcing is introduced as an inference problem. The problem will be analyzed in detail in Chap. 5. In order to solve it, a simplification of belief propagation is introduced: approximate message

passing, which will yield the Bayes optimal estimates and can be analyzed by the state evolution equations that are derived in the end of this chapter.

1.1.1 Statistics in a Nutshell

In order to specify a probabilistic model it is necessary to fix the sample space, Ω, of the problem. The sample space might be continuous or discrete, depending on the nature of the elements that compose the sample space, i.e., the outcomes. It only remains to define the probability density function, or simply **probability distribution**, $P(A)$ that assigns a non-negative real number to the event $A \subseteq \Omega$. Naturally, $P(\Omega) = 1$.

A convenient approach to work with probabilities is to introduce **random variables**. A random variable, x, is a mapping from the sample space, Ω, to the real numbers, $x : \Omega \to \mathbb{R}$ for each $s \in \Omega$, i.e. $s \mapsto x(s)$. If x is distributed according to $P_x(x)$ we write

$$x \sim P_x(x)\,. \tag{1.1}$$

To deal with more than one random variable we introduce the random vector $\boldsymbol{x} = (x_1, \ldots, x_N)^\intercal$. We will often refer to the "random vector" simply as another "random variable". The **joint probability distribution** (JPD) captures the behavior of $\boldsymbol{x}$

$$\boldsymbol{x} \sim P_{\mathbf{x}}\left(x_1, \ldots, x_N\right)\,. \tag{1.2}$$

Indeed we will neglect the "joint" in the "joint distribution" and just refer to it as a "distribution". The **marginal distribution** of x_i can be recovered by integrating out $\boldsymbol{x}_{\setminus i} = (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_N)^\intercal$

$$P_{x_i}\left(x_i\right) = \int P_{\mathbf{x}}\left(x_1, \ldots, x_N\right) \mathrm{d}\boldsymbol{x}_{\setminus i}\,. \tag{1.3}$$

If the random variables $x_1, \ldots, x_N$ are **independent**, then their probability distribution factorizes

$$P_{\boldsymbol{x}}\left(x_1, \ldots, x_N\right) = \prod_{i=1}^{N} P_{x_i}\left(x_i\right)\,. \tag{1.4}$$

If the P_{x_i} are additionally identical for all i, we say that the x_i are independent and identically distributed.

Assume that the set of random variables $\boldsymbol{x}$ is partitioned as $\boldsymbol{x} = \boldsymbol{x}_A \cup \boldsymbol{x}_B$, with $A = i_1^a, i_2^a, \ldots, i_{N_A}^a$ and similarly for B. In the following we write $\boldsymbol{x}_A, \boldsymbol{x}_B$ for $\boldsymbol{x}_A \cup \boldsymbol{x}_B$. The **conditional probability distribution** of $\boldsymbol{x}_A$ given $\boldsymbol{x}_B$ is

$$P_{\boldsymbol{x}_A|\boldsymbol{x}_B}\left(\boldsymbol{x}_A \mid \boldsymbol{x}_B\right) = \frac{P_{\boldsymbol{x}_A,\boldsymbol{x}_B}\left(\boldsymbol{x}_A, \boldsymbol{x}_B\right)}{P_{\boldsymbol{x}_B}\left(\boldsymbol{x}_B\right)}\,. \tag{1.5}$$

The subset $\boldsymbol{x}_B = (x_{i_1^b}, x_{i_2^b}, \ldots)^\intercal$ of random variables can be understood as *observed*, or fixed, random variables. Conditional probabilities behave exactly like regular probabilities

and consequently one can extend the concepts of independence and marginal distributions to conditional distributions. The conditional marginals of $P_{\boldsymbol{x}_A|\boldsymbol{x}_B}(\boldsymbol{x}_A \mid \boldsymbol{x}_B)$ are

$$P_{x_{i^a}|\boldsymbol{x}_B}\left(x_{i^a} \mid \boldsymbol{x}_B\right) = \int P_{\boldsymbol{x}_A|\boldsymbol{x}_B}\left(\boldsymbol{x}_A \mid \boldsymbol{x}_B\right) \mathrm{d}\boldsymbol{x}_{A\setminus i^a} \tag{1.6}$$

and we say that the two subsets $\boldsymbol{x}_A$ and $\boldsymbol{x}_B$ are **conditionally independent** of $\boldsymbol{x}_S$, if

$$P_{\boldsymbol{x}_A,\boldsymbol{x}_B|\boldsymbol{x}_S}\left(\boldsymbol{x}_A, \boldsymbol{x}_B \mid \boldsymbol{x}_S\right) = P_{\boldsymbol{x}_A|\boldsymbol{x}_S}\left(\boldsymbol{x}_A \mid \boldsymbol{x}_S\right) P_{\boldsymbol{x}_B|\boldsymbol{x}_S}\left(\boldsymbol{x}_B \mid \boldsymbol{x}_S\right) . \tag{1.7}$$

From (1.5) the Bayes theorem follows

$$P_{\boldsymbol{x}_A|\boldsymbol{x}_B}\left(\boldsymbol{x}_A \mid \boldsymbol{x}_B\right) = \frac{P_{\boldsymbol{x}_B|\boldsymbol{x}_A}\left(\boldsymbol{x}_B \mid \boldsymbol{x}_A\right)}{P_{\boldsymbol{x}_B}\left(\boldsymbol{x}_B\right)} P_{\boldsymbol{x}_A}\left(\boldsymbol{x}_A\right) . \tag{1.8}$$

It will be our fundamental tool tackle inference problems, as will be discussed later.

An alternative specification of the probability distribution is given by its **moments**. The k-th moment of $P_x(x)$is defined as

$$\mathbb{E}_x[x^k] \coloneqq \int \mathrm{d}x\, x^k\, P_x(x)\,, \tag{1.9}$$

It follows the moment generating function as the Laplace transform of x as $\mathbb{E}_x[\mathrm{e}^{tx}]$, which contains the same information as $P_x(x)$. In many situations it is not necessary to know the entire distribution, i.e. all moments, instead it often suffices to only know the first few moments of x. One particularly important example is the **Normal distribution**

$$\mathcal{N}(x; \mu, \sigma^2) \coloneqq \frac{1}{\sqrt{2\pi\sigma^2}} \mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{1.10}$$

for which it suffices to know the first two moments. All higher moments either follow from the mean and the (co)variance, or vanish. In the following we denote the **mean** by

$$\mu = \mathbb{E}_x[x] \tag{1.11}$$

and the **variance** by

$$\sigma^2 = \mathbb{E}_x[(x-\mu)^2] \tag{1.12}$$

respectively.

The normal distribution is important because it appears as the asymptotic distribution of the sample average. Let $\{X_1, \ldots, X_N\}$ denote N independent **sample** from $P_x(x)$. The **central limit theorem** (CLT) states that the **sample average**

$$\bar{x}_N \coloneqq \frac{1}{N} \sum_{i=1}^{N} X_i\,, \tag{1.13}$$

Figure 1.2: A graphical representation of a Markov chain of order one (left) and of order two (right).

is asymptotically normal distributed, with mean μ and variance σ^2/N. It follows the **law of large numbers**

$$\lim_{N\to\infty} \bar{x}_N = \mu\,. \tag{1.14}$$

The CLT is typically more formally expressed in terms of convergence in probability of the partial sum

$$S_N \coloneqq \sum_{i=1}^{N} X_i\,, \tag{1.15}$$

for which the CLT states that

$$\lim_{N\to\infty} \Pr\left(\frac{S_N - N\mu}{\sigma\sqrt{N}} \leq t\right) = \frac{1}{2\pi}\int_{-\infty}^{t} \mathrm{d}s\, \mathrm{e}^{-s^2/2}\,. \tag{1.16}$$

A notational remark: often the subscript of the distribution is an unnecessary repetition. However, sometimes it is useful to distinguish that two distributions are different. When the latter is not clear from the context, we will include the subscript, otherwise, if there is no ambiguity, we will drop it to lighten the notation.

1.1.2 Markovian Random Fields

We start by the introduction of Markov chains. These are suitable objects in many applications and Markovian random fields can be regarded as a generalizations of them.

Box 1.1: Markov Chains

Consider the random vector $\boldsymbol{x} = (x_1, x_2, x_3, \dots)^\intercal$ and let the state space of x_i be $\Omega = \{s_1, s_2, s_3, \dots\}$. Then $\boldsymbol{x}$ is called a **(discrete) Markov chain** if

$$P\left(x_{t+1} = s_{j_{t+1}} \mid x_t = s_{j_t}, x_{t-1} = s_{j_{t-1}}, \dots, x_1 = s_{j_1}\right) = P\left(x_{t+1} = s_{j_{t+1}} \mid x_t = s_{j_t}\right)\,. \tag{1.17}$$

That is: the transition probabilities only depend on the current state of x_t, but not on the remaining states of $(x_i)_{i<t}$. The index t here is chose consciously because Markov chains are often utilized to model probabilistic temporal processes where the x_t are random variables that represent probabilistic changes of a state in time. We will come back to this interpretation in chapter 6, but for now let us ignore this aspect.

What are the implication of $\boldsymbol{x}$ being a Markov chain? In other words, what are the restrictions that a Markov chain implies for the joint distribution? To that end we consider an example.

Assume that $\boldsymbol{x} = (x_1, x_2, x_3, x_4)^\intercal$ is a Markov chain and consider its distribution. As a consequence of (1.6) we may write

$$P(x_1, x_2, x_3, x_4) = P(x_4 \mid x_3, x_2, x_1)\, P(x_3, x_2, x_1) \tag{1.18}$$

and then use the Markov chain property (1.17) to further simplify the r.h.s.

$$P(x_4 \mid x_3, x_2, x_1)\, P(x_3, x_2, x_1) = P(x_4 \mid x_3)\, P(x_3, x_2, x_1)\ . \tag{1.19}$$

Repeated application of this argument to the r.h.s. then yields

$$P(x_1, x_2, x_3, x_4) = P(x_4 \mid x_3)\, P(x_3 \mid x_2)\, P(x_2 \mid x_1)\, P(x_1)\ . \tag{1.20}$$

The joint probability distribution factorizes over the pairs (x_1, x_2), (x_2, x_3) and (x_3, x_4). Since only variables within the pairs are interacting we may write

$$P(x_1, x_2, x_3, x_4) \propto \psi_1(x_1, x_2)\psi_2(x_2, x_3)\psi_3(x_3, x_4)\,, \tag{1.21}$$

with ψ_a some interaction function, not necessarily a PD.
We can represent (1.21) graphically. For each variable x_i we draw a round note and for each function ψ_a we draw a (black) square node. Then variable node i is connected to function node a by an edge if x_i appears in the argument of ψ_a. For (1.21) this is done in the left of Fig. 1.2. It becomes apparent from the figure that x_4 is conditionally independent of x_2 and x_1, when we condition on x_3. Graphically this is because all paths that lead to x_4 go thorough x_3, i.e. x_3 separates the two sets $\{x_1, x_2\}$ and $\{x_4\}$. This may be deducted from (1.19) by virtue of the Markov property $P(x_4 \mid\ x_3)P(x_3, x_2, x_1) = P(x_4 \mid x_3)P(x_1, x_2 \mid x_3)P(x_3)$.
These ideas extend to Markov chains of the order r, for which

$$P\left(x_{t+1} = s_{j_{t+1}} \mid x_t = s_{j_t}, x_{t-1} = s_{j_{t-1}}, \ldots, x_0 = s_{j_0}\right) = P\left(x_{t+1} = s_{j_{t+1}} \mid x_t = s_{j_t}, \ldots, x_t = s_{j_{t-r+1}}\right)\ . \tag{1.22}$$

How does such a Markov chain factorize? Let us consider $P(x_1, x_2, x_3, x_4)$ to be a Markov chain of order two. We can actually already draw the graphical representation because each node x_{t+1} only interacts with the nodes x_t and x_{t-1}; we do so in Fig. 1.2. At the same time, we can write $P(x_1, x_2, x_3, x_4) = P(x_1, x_4 \mid x_2, x_3)P(x_2, x_3)$. By virtue of the Markov property it is then straight forward to see that the r.h.s. can be further simplified to $P(x_1, x_4 \mid x_2, x_3)P(x_2, x_3) = P(x_1 \mid x_2, x_3)P(x_4 \mid x_2, x_3)P(x_2, x_3)$. Thus we have $P(x_1, x_2, x_3, x_4) = \psi_1(x_1, x_2, x_3)\psi_2(x_2, x_3, x_4)$.

Motivated by the observations in Box 1.1 we draw our attention to JPD of the form

$$P(\boldsymbol{x}) = \frac{1}{Z} \prod_{a=1}^{M} \psi_a\left(\boldsymbol{x}_{\partial a}\right)\ . \tag{1.23}$$

We introduced the normalization Z, such that the functions ψ_a need not be probabilities. However, they must be *positive* everywhere. The random vector may be drawn from a

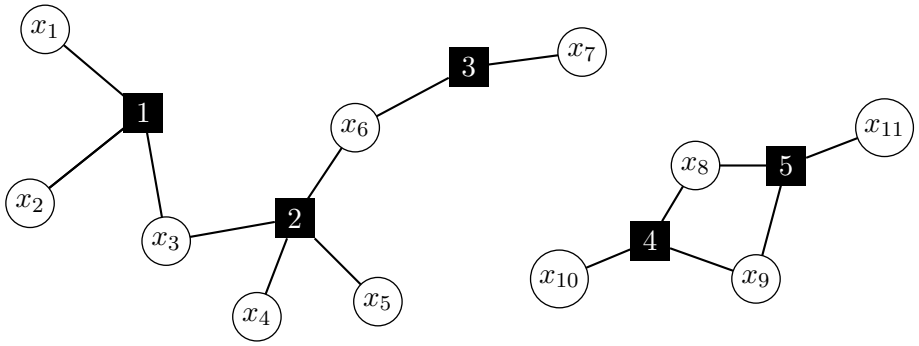

Figure 1.3: The graphical presentation of (1.26).

discrete or continuous state space $\boldsymbol{x} \in \Omega^N$. We have seen that such distributions can be represented graphically. We draw a function node for each a and $\partial a \subseteq \{1, \ldots, N\}$ are all the nodes that are connected it. Such a graphical representation is also known as **(factor) graph** and the function nodes are also referred to as factor nodes.

Let us call a subset of variable nodes *a region*. Then, for the three disjoint subsets A, B, S, we say that a region S separates the regions A and B if there is no path that leads from a node in A to a node in B, without passing through S. It is now natural to observe that if S separates A from B, then the two regions A and B are independent, once we condition on S, and their JPD factorizes. The **global Markov property** holds

$$P(\boldsymbol{x}_A, \boldsymbol{x}_B \mid \boldsymbol{x}_S) = P(\boldsymbol{x}_A \mid \boldsymbol{x}_S) P(\boldsymbol{x}_B \mid \boldsymbol{x}_S) . \tag{1.24}$$

In other words

$$P(\boldsymbol{x}_A \mid \boldsymbol{x}_S, \boldsymbol{x}_B) = P(\boldsymbol{x}_A \mid \boldsymbol{x}_S) . \tag{1.25}$$

Distributions of the type (1.23) imply that the global Markov property holds. They are known as **Markovian random fields** (MRF).

As an example consider the following MRF

$$P(\boldsymbol{x}) = \psi_1(x_1, x_2, x_3) \psi_2(x_3, x_4, x_5, x_6) \psi_3(x_6, x_7) \psi_4(x_{10}, x_8, x_9) \psi_5(x_8, x_9) , \tag{1.26}$$

How does the associated graphical model (factor graph) look like? Following the rules, introduced in Box 1.1, the graphical representation is found by drawing M function nodes and N variable nodes and connecting those variable nodes that appear in the argument of ψ_a to the function node a. It is depicted in Fig. 1.3.

Given a factorized JPD it is easy drawing the associated factor graph. But how about the converse: given a graphical model in form of a factor graph, is it possible to write the JPD as a product of the contributions from each function node as in (1.23)? It turns out that it is not possible and naïvely inversing the rules can lead to mistakes However, owing to Hammersley and Clifford, there is a constructive way to obtain the factorization, once the factor graph is specified. To that end we must introduce the concept of cliques. A **clique** is

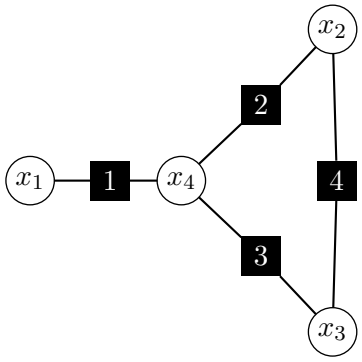

Figure 1.4: Attempting to factorize the JPD of this graphical model according to (1.23) yields an errornous factorization.

a subgraph such that, for any pair of variable nodes in the subgraph, there is a function node adjacent to both of them. The *completion* of a factor graph is then understood as adding one factor node for each clique that connects to all variable nodes in that clique.

Following [MM09] we consider the graphical model in Fig. 1.4 as an example. Inverting our rules would lead to a JPD of the form $\psi_1(x_1, x_4)\psi_2(x_2, x_4)\psi_3(x_3, x_4)\psi_4(x_2, x_3)$. This cannot be true because there is only one region that separates the graph into two disjoint sets, that is $\{x_4\}$. By virtue of the global Markov property the JPD thus factorizes as $P(x_1, x_2, x_3, x_4) = \psi_1(x_1, x_4)\psi_2(x_2, x_3, x_4)$. Indeed, after completion of the graph, the latter result is recovered from the factor graph. The **Hammersley-Clifford theorem** [HC71] guarantees that if $P(\boldsymbol{x}) > 0$ and $P(\boldsymbol{x})$ satisfies the global Markov property w.r.t. a factor graph, then $P(\boldsymbol{x})$ can always be written in the form (1.23) w.r.t. the *completed graph*.

Of particular interest shall be **pairwise MRFs** in which the interactions are restricted to be no higher than two-body

$$P(\boldsymbol{x}) = \frac{1}{Z} \prod_{(ij)\in\mathcal{E}} \psi_{ij}(x_i, x_j) \prod_{i=1}^{N} \phi_i(x_i)\,. \tag{1.27}$$

Here we denoted with $\mathcal{E}$ the pairs of variable nodes (x_i, x_j) that are connected through a single function node. The ϕ_i may be interpreted as local fields, acting on the variables x_i. Graphically such a model can take different forms; the previously considered Markov chains of order one fall into this category (Fig. 1.2). Further examples will follow in the next section that addresses pairwise MRFs.

The graphically representation of pairwise MRFs as factor graphs is found by drawing a round variable node for each variable x_i and a square function/factor node for each function ψ_{ij} and ϕ_i. For pairwise models it is not essential to draw all the factor nodes because they are merely connecting two nodes. One might disregard them and draw directly a line between any two variable nodes that interact. Consequently each edge (ij) between the variable nodes x_i, x_j carries an interaction $\psi_{ij}(x_i, x_j)$. The contribution from $\phi_i(x_i)$ can be represented in different ways. One is to represented it by another edge that leads from x_i to

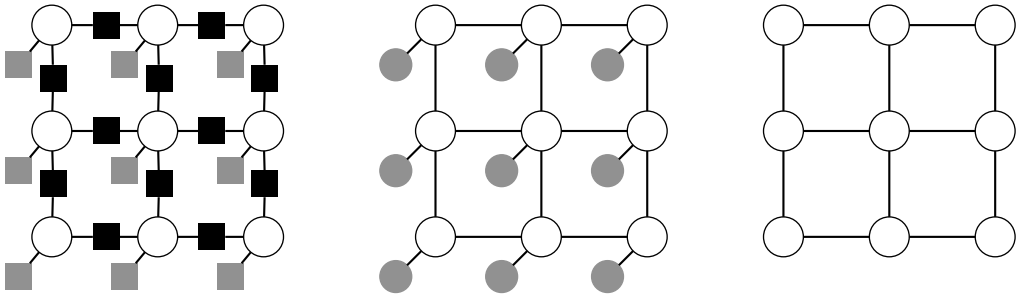

Figure 1.5: A pairwise MRF and its different representations, as explained in the text.

an auxiliary variable, represented by a gray node. Another way is to absorb its contribution into the the variable node directly.

After all, these are just representations of JPDs of the form (1.27) and we may chose any representation that suits our needs best. In this section on random graphs we will be working with the latter representation that only contains the variable nodes and some lines that are connecting them.

The attention will from now on lay on graphical models that are themselves subject to randomness. In other words, the elements $(ij) \in \mathcal{E}$ in (1.27) will be random variables.

1.2 Random Graphs

The graphical models motivates the definition of a **Graph** as the pair $\mathcal{G}(\mathcal{V}, \mathcal{E})$, with $\mathcal{V}$ the set of *nodes* and $\mathcal{E}$ the set of *edges*. The nodes may also be referred to as *vertices* and the edges as *links*. **Random graphs** (RG) are graphs that were drawn at random from some ensemble of graphs. Here we limit ourselves to the ensemble of pairwise MRF that take the form (1.27). In the whole thesis we will only consider undirected graphs (these are graphs for which the edges do not point into a specific direction). Throughout this thesis we will be interested in the properties of random graphs that hold with probability tending to 1 when the number of nodes goes to infinity, $N \to \infty$. We say they hold with high probability.

1.2.1 Random Graph Ensembles

There are plenty of random graph models out there. Different models may result in graphs with different structures. Some of them are analytically tractable, others are not. Most importantly for us will be the locally tree-like structure of these graphs. We will work with ensembles that "guarantee" (in the probabilistic sense) this desired property.

Erdős-Rény Graphs

There are two variants of Erdős-Rény graphs.

- The $\mathscr{G}(N, M)$ model consists of all graphs that contain N nodes and M edges and each graph is equally likely.
- The $\mathscr{G}(N, p)$ model consists of all graphs with N nodes and edges chosen independently with probability p.

The difference between these two models disappear when $N \to \infty$ and they become interchangeable (in most aspects) if we set $M = p\binom{N}{2}$. The major difference between the two arises from the fact that the number of edges is a random variable in the $\mathscr{G}(N, p)$ model. There are $N(N-1)/2$ possible pairings for the N nodes; each of them obtains an edge with probability p. Thus we expect $pN(N-1)/2$ edges, as compared to M in the $\mathscr{G}(N, M)$ model. As $N \to \infty$ the fluctuations around the expected number of edges become neglectable. For a formal treatment of this statement cf. Theorem 2.2 in [Bol01].

We will be interested in the limit where $N \to \infty$ and therefore pick $\mathscr{G}(N, p)$ as our model of choice for the analytical treatments in this section.

The **degree** of a node in an undirected graph is the number of links that it has to other nodes. We denote d_i the degree of node i, i.e., the size of the set of its neighbors, $d_i = |\partial i|$. The degree distribution encapsulates important information about the graph. Before deriving it we first note that

$$M = \frac{1}{2} \sum_i d_i \tag{1.28}$$

and combine this observation with the expected number of edges, $pN(N-1)/2$ in the $\mathscr{G}(N, p)$ model we obtain that $p = c/(N-1)$ which, for large N, we replace by

$$p = \frac{c}{N}\,, \tag{1.29}$$

where we introduced the **average degree**

$$c = \frac{1}{N} \sum_i d_i\,. \tag{1.30}$$

For the $\mathscr{G}(N, p)$ model the probability that node i has degree d_i is a binomial process. We fix one node and compute the probability that we draw d edges in $N-1$ trials

$$P(d_i = d) = \binom{N-1}{d} p^d (1-p)^{N-d}\,. \tag{1.31}$$

In the limit $N \to \infty$ the binomial distribution can be approximated by the Poisson distribution

$$\lim_{N\to\infty} P(d_i = d) = \frac{c^d \mathrm{e}^{-c}}{d!}\,. \tag{1.32}$$

Apparently the degree distribution only depends on a single parameter, the average degree c. When c is varied typical graphs from the ensemble exhibit different structures [ER60]. For $c < 1$ the graph $\mathcal{G}$ consists of small, disconnected, components that are w.h.p. trees. At

$c = 1$ a phase transition takes place and the previously disconnected components suddenly connect. When $c > 1$ they form a giant connected component, that contains $\Theta(N)$ nodes. The *local structure* remains **tree-like**.

Clearly, when c grows beyond unity, the tree-like structure must break down at some scale. Equation (1.28) implies that w.h.p. number of edges is

$$M = \frac{c}{2} N \,. \tag{1.33}$$

Since there are only N nodes and $\Theta(N)$ edges, adding further edges will necessarily lead to the closure of a **loop**, when $c > 1$. With high probability this happens on a scale of $\log(N)$. Let us outline a hand-waving derivation. We pick a node i uniformly at random among all nodes. First we reveal its direct neighbors, w.h.p. there are c of them. Similarly, the number of neighbors at distance 2 is w.h.p. c^2. At distance r we expect c^r neighbors. A loop must close when the revealed neighbors contain all nodes in the graph, i.e., when $c^r = N$. Thus we expect that loops typically appear at a distance $r = \log(N)/\log(c)$. Obviously short(er) loops are present, but they are subleading. The crude estimate does indeed provide the correct *typical length* of the **shortest loop** in an Erdős-Rény graph

$$l = \frac{\log(N)}{\log(c)} + o(1) \,. \tag{1.34}$$

The main difficulty is to control the errors that are accumulating in each of the r steps. A rigorous treatment can be found in [Bol01] (Theorem 10.19).

Random Regular Graphs

A close relative of the Erdős-Rény graphs are the *random regular graphs*. We denoted by $\mathscr{G}(N, d)$ the model of ***d*-regular graphs** that consists of all graphs that contain N nodes, all of them of degree d. In some respects d-regular graphs are much simpler to work with. For instance, their degree distribution is trivially

$$P(d_i = d) = \delta(d_i - d) \,. \tag{1.35}$$

Consequently $c = d$ and the number of edges is $M = dN/2$. In other respects they are more difficult to treat. One major difficulty is to sample uniformly from $\mathscr{G}(N, d)$. Even evaluating how many d-regular graphs there are, $|\mathscr{G}(N, d)|$, turns out to be very difficult, even when $N \to \infty$. It will be necessary to introduce another model that is asymptotically equivalent to $\mathscr{G}(N, d)$.

The **configuration model**, also *pairing model*, has the required property and can be defined as follows. Consider dN *half-edges*, partitioned into N cells, $1, \ldots, N$, with d half-edges in each cell. The configuration model consists of all graphs that are *perfect matchings* of the half-edges into $dN/2$ pairs (ij) that do not contain self-loops (pairs of the form (ii)) or multiple edges (repetition of the same pairing). Each pairing corresponds to an edge and each cell to a node.

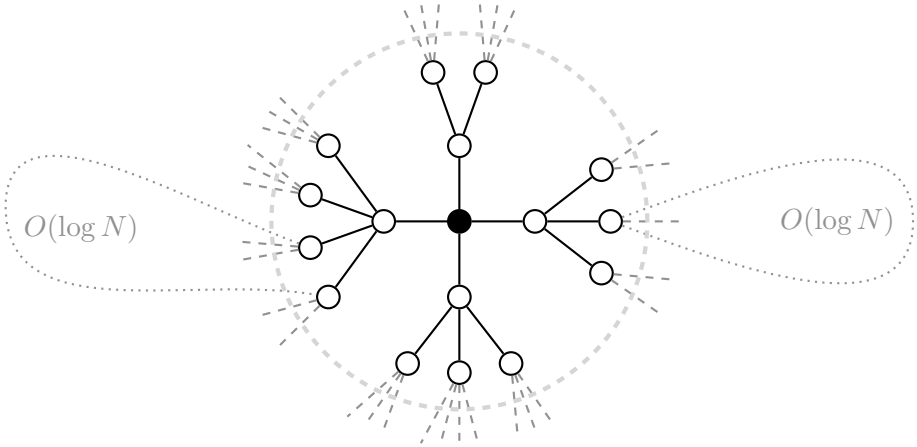

Figure 1.6: An exemplary branching process in the neighborhood of a node (black) in an Erdős-Rény graph. The dashed circle denotes the boarder from the second to the third generation. Erdős-Rény graphs, as well as regular random graphs, are typically tree-like up to a distance of the order of $\log N$ (the number of nodes that the dotted loops contain).

The configuration model can be generalized to arbitrary degree distributions by adapting the elements in the cells to the desired **degree sequence** (the number of nodes of degree d). Therefore it is also applicable to Erdős-Rény graphs. Similarly to the Erdős-Rény graphs, also random regular graphs exhibit a local tree-like structure and the typical loop-size is

$$l = \frac{\log(N)}{\log(d-1)} + O(\log(N)) \ . \tag{1.36}$$

A rigorous proof is given in [Bol01] (Theorem 10.15).

The vanilla implementation of the above procedure consists in producing pairings uniformly at random and rejecting a sample if it contains self-loops or repeated edges. The expected number of steps is of the order of $Nde^{\frac{d^2-1}{4}}$, if $d = O(n^{1/3})$. For $d = O(\log(N))$ this is polynomial, but can soon become impractical in applications. There is an alternative method due to Steger and Wormald [SW99] that samples *asymptotically* uniformly from the configuration model with expected running time of $O(nd^2)$.

1.2.2 Sparse and Dense Graphs

There are several parameters in the previously introduced models. The number of nodes, N, shall be our reference. Since we will be considering $N \to \infty$ it is important to determine how the other parameters scale with respect to N. For the random d-regular graphs this reduces to determine how d scales with N. For Erdős-Rény graphs it breaks down to c in the limit where the two models $\mathscr{G}(N, M)$ and $\mathscr{G}(N, p)$ are interchangeable. Since the average degree is equal to d for d-regular graphs, we shall use c as our common reference point.

We distinguish between sparse graphs, with $c = \Theta(1)$, and dense graphs that have $c = \Theta(N)$. Genuinely, the number of edges in sparse graphs is of the same order as the number of nodes, $M = \Theta(N)$. Another implication of the sparse regime is that the typical loop size is $\Theta(\log(N))$, while they are $\Theta(1)$ in the dense regime.

The $\Theta(\log(N))$ loop size in the sparse regime will be crucial in Sec. 1.3.2 where we introduce message passing as an algorithm to efficiently compute JPDs on *trees*. We might hope that an algorithm, designed for trees, still provides a good approximation if deviations from the tree structure are small over large local scale, i.e., when the typical loops size is large. Concurrently, the large connectivity in the dense regime, $c = \Theta(N)$ often erodes individual contributions from neighboring nodes which can significantly simplify the analytical treatment by reducing the many individual contribution, to an effective Gaußian contribution, by virtue of the CLT.

It is, however, important to note that these are just some tendencies that do not necessarily apply in general and many characteristics not only depend on whether a graph is sparse or not, but also the concrete graph model and application at hand, as we shall experience in Chap. 4 and 5.

1.2.3 Dynamical Processes on Graphs

The previously encountered graph models can all be understood as prescriptions for the construction of random graphs from that particular ensemble. There might be different prescription that lead to the same result (in terms of the properties of the resulting graph, not the specific sample). Consider the two encountered Erdős-Rény graph models as an example. One prescription for the construction of graphs from the $\mathscr{G}(N, M)$ model consists in throwing M edges uniformly at random at the $\binom{N}{2}$ available pairs of nodes. For graphs from the $\mathscr{G}(N, p)$ model we could proceed very similarly. Now a possible prescription is to go through the $\binom{N}{2}$ distinct pairings and for each pair add an edge with probability p. Another example is the prescription to obtain uniform samples from the ensemble of random regular graphs by means for the configuration model.

All these prescriptions can be considered examples of graph processes. A **graph process** is a process that starts with some initial graph $\mathcal{G}_0$ and adds (or removes) edges according to some stochastic rule. This process defines a sequence of random graphs $\mathcal{G}_0, \mathcal{G}_1, \mathcal{G}_2, \ldots$. Depending on the rules, this process typically ends at some iteration, i_s, at which some desired criterion is reached. Throughout this subsection we employ the subscript i to denote the iteration of the considered dynamical process. In practice it is often not necessary to follow the evolution of the dynamical process $\mathcal{G}_i$ directly. Instead it usually suffices to follow the evolution of some characteristic observables, $\boldsymbol{A}_i$. This can be done, provided that the evolution of these observables contains enough information for a closed form description of the dynamical process of $\boldsymbol{A}_i$. That is, if we can find $\boldsymbol{A}_{i+1} \sim P(\boldsymbol{A}_1, \ldots, \boldsymbol{A}_i)$.

Let us assume the process is Markovian, such that $\boldsymbol{A}_{i+1} \sim P(\boldsymbol{A}_i)$. In many situations this can be recast as

$$\boldsymbol{A}_{i+1} = \boldsymbol{A}_i + \boldsymbol{T}\boldsymbol{z}\,, \tag{1.37}$$

with $\boldsymbol{z}$ some appropriate indicator variable and $\boldsymbol{T}$ a transition matrix. Then the idea (hope) is that when $N \to \infty$ the discrete random process can be approximated well by the expected change

$$\mathbb{E}\left[\boldsymbol{A}_{i+1} - \boldsymbol{A}_i\right] = T(\frac{\boldsymbol{A}_i}{N}) \tag{1.38}$$

and that the rescaled process

$$\boldsymbol{\eta}(t) = \lim_{N\to\infty} \frac{1}{N}\boldsymbol{A}_{\lfloor tN \rfloor}\,. \tag{1.39}$$

is well described by a set of deterministic differential equations, such that

$$\dot{\boldsymbol{\eta}} = T\left(\boldsymbol{\eta}\right) \tag{1.40}$$

A general purpose theorem due to Wormald [Wor99a] that we outline in Appendix A.1 often allows to prove concentration of the random process around the deterministic approximation. The ideas are best understood by the study of the simple example in the box below.

Box 1.2: The Standard Random Graph Process

We consider the $\mathscr{G}(N, M)$ model as a simple example and we follow the previously introduced prescription. We denote by $v_{d,i}$ our observable: the total number of nodes of degree d at iteration i. We start with G_0 a graph with N vertices and zero edges, i.e. we start with $v_{0,0} = N$ and $v_{d,0} = 0 \;\forall\; d > 0$. Each time an edge is thrown it hits two distinct nodes. When a node of degree d is hit at iteration i its degree changes from $d \to d+1$, thus $v_{d,i} \to v_{d,i+1} - 1$ and $v_{d+1,i} \to v_{d+1,i+1} + 1$. If we collect $\boldsymbol{v}_i = (v_{0,i}, v_{1,i}, v_{2,i}, \dots)^\intercal$ we may rewrite the stochastic process as

$$\boldsymbol{v}_{i+1} = \boldsymbol{v}_i + \boldsymbol{T}\boldsymbol{z}\,, \tag{1.41}$$

with the transition matrix

$$T_{mn} = -\delta_{m\,m} + \delta_{m+1\,m} \tag{1.42}$$

and $\boldsymbol{z}$ a random vector $\boldsymbol{z} = \hat{\mathbf{e}}_{d_1} + \hat{\mathbf{e}}_{d_2}$ with $\hat{\mathbf{e}}_{d_k}$ the unit vector, pointing in direction d_k. Thus $\boldsymbol{z}$ indicates the degree of the two nodes at the end of a thrown edge. For example, $\boldsymbol{z} = (1, 1, \dots)^\intercal$ indicates that the thrown edge landed between a node of degree zero and another of degree one. So the pair d_1, d_2 should be drawn from the degree distribution of possible pairings that do not contain multiple edges or self-loops.

When N is large this can be approximated by drawing d_1 and d_2 from the renormalized degree sequence at iteration i

$$d_1,\, d_2 \sim \frac{\boldsymbol{v}_i}{\|\boldsymbol{v}_i\|_1}\,. \tag{1.43}$$

Further the degree sequence only changes by a maximum of $O(N^{-1})$ in every iteration we may replace $\boldsymbol{z} = 2\hat{\mathbf{e}}_d$ and we obtain

$$\boldsymbol{v}_{i+1} = \boldsymbol{v}_i + 2\boldsymbol{T}\hat{\mathbf{e}}_d + O\left(\frac{1}{N}\right)\,. \tag{1.44}$$

Next we compute the expected change of the stochastic process $\boldsymbol{v}_i$ w.r.t. $\hat{\mathbf{e}}_d$. This amounts to replacing the RV $\hat{\mathbf{e}}_d$ by $\boldsymbol{v}_i/\|\boldsymbol{v}_i\|_1$ and since $\|\boldsymbol{v}_i\|_1 = N$ we obtain

$$\mathbb{E}\left[\boldsymbol{v}_{i+1} - \boldsymbol{v}_i\right] = 2\,\boldsymbol{T}\frac{\boldsymbol{v}_i}{N}\,. \tag{1.45}$$

Utilizing (1.39) both sides can be replaced with their rescaled equivalents, which yields the following set of differential equations

$$\dot{\boldsymbol{\eta}} = 2\boldsymbol{T}\boldsymbol{\eta}\,. \tag{1.46}$$

Looking at these equations component-wise

$$\dot{\eta}_0 = -2\eta_0 \tag{1.47}$$
$$\dot{\eta}_d = -2\eta_d + 2\eta_d \quad \forall\ d > 0 \tag{1.48}$$

reveals that the dynamical process is indeed a Poisson process. It's solution is $\eta_d(t) = (2t)^d \mathrm{e}^{-2t}/d!$ and if the process is run until all M edges are exhausted, i.e. at $t_\mathrm{s} = M/N = c/2$, and (1.32) is recovered

$$\eta_d(t_\mathrm{s}) = \frac{c^d \mathrm{e}^{-c}}{d!}\,. \tag{1.49}$$

The degree distribution is relevant because many physical properties can be derived from it, once it is determined. It is well known, that the standard graph process undergoes several phase transitions during its evolution. They can all be derived, once the degree distribution is known.

The percolation phenomenon in the standard graph process is a good introductory example. As more and more edges are added to the empty graph, small, isolated components grow (almost all of them are trees). At the **percolation** threshold these components come together and form a "giant" connected component that *percolates* the graph "as a whole". Our notion of a "giant component that percolates the graph as a whole" is that it contains an $\Theta(N)$ fraction of the nodes, not that it obtains all nodes. The point where the graph contains w.h.p. *all* nodes will indicate another transition.

To derive the threshold value from the degree distribution, we consider the branching process that reveals the neighborhood of a random node in the graph. We will be following the k-th generation at the front of the branching process (cf. Fig. 1.6). The contributions from the nodes in the front at each generation are added up to form the number of nodes in the connected component. When the process moves from the k-th to the $k+1$-th generation the number of nodes in the front increases by $d-2$ because there are $d-1$ nodes in the direction of the moving front and one node must be subtracted as it was already counted in the front of the previous generation.

Since all the branches have the same properties it suffices to only consider the expected gain when moving from the k-th to the $k+1$-th generation. This can be expressed in terms of the degree (let us use η_d from the box above, expressed in the natural time scale $t = c/2$). The gain, $\sum_{d\leq 1}(d-2)d\,\eta_d(c/2) \simeq c^2 - c$, must be positive for a giant component to exist, which is only the case for $c > 1$. This argument was proven to hold in more generality in [MR95].

As was already mentioned above, another point of interested is the threshold where the graph contains w.h.p. no more isolated vertices. This can be obtained right away from the degree distribution as the point where $\eta_0 \overset{N\to\infty}{\longrightarrow} 0$, i.e., when $c > \ln(N)$.

Thus we have the following characteristic points in the evolution of graphs from $\mathscr{G}(N, M)$

- When $c < 1$ a graph in $\mathscr{G}(N, M)$ contains w.h.p. no connected components.
- When $c > 1$ a giant component that contains $\Theta(N)$ nodes is w.h.p. present.
- If $c < \ln(N)$ isolated vertices are w.h.p. present.
- If $c > \ln(N)$ a graph in $\mathscr{G}(N, M)$ is w.h.p. connected, i.e. it contains no isolated vertices.

Bootstrap Percolation and Contagious Sets

The concept of percolation is related to the sudden emergence of a global structure in the random graph. Percolation problems are not limited to the appearance of a giant connected component in a graph.

In **bootstrap percolation** [CLR79] we start with a graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$ and successively remove nodes that have degree smaller than k. This procedure is iterated until a stable configuration is reached. In other words, we are interested in the *largest subgraph* with minimal degree at least k, also known as the ***k*-core** of $\mathcal{G}(\mathcal{V}, \mathcal{E})$. For completeness, let us define a **subgraph** as a subset of nodes and edges that contains only those edges that lead back into the selected subset of nodes. If there exists no such subgraph of size $\Theta(N)$, the k-core is said to be (asymptotically) empty.

In order to reveal the k-core, one can find a set of differential equations that approximate the graph process of deletion of nodes that have degree smaller than k. It turns out that, depending on k, the nature of the transition from a non-empty to an empty k-core can be of different types. Unlike the giant component in the standard graph process, which appears in a *second order* phase transition, it is possible to show that for bootstrap percolation the transition is *first order* for $k \leq 3$. When $k \leq 2$ the transition is second order and the k-core grows together with with the giant component [PSW96b].

k-cores are objects of interest in the study of collective dynamics in social and biological networks. Examples include, and are not limited to, the spreading of epidemics, failure propagation in networks and adaption and innovation diffusion. In the **threshold model** it is assumed that each variable node, x_v, in the graph can take one of two state, $x_v \in \{0, 1\}$. If $x_v = 0$, node v is said to be inactive, while if $x_v = 1$ it is said to be active. Given some initial condition, the graph is subject to the following dynamical process. Node v switches its state from inactive to active if τ_v or more of its neighbors are active. Once activated a node remains in this state forever.

The result of running such a dynamical process crucially depends on the initial state. The question of major importance concerns the identification of the smallest set, that leads to an activation of the whole graph. The significance of the k-core of a graph can be laid out by consideration of the complementary problem. Nodes that remain inactive, after the dynamical process has converged, must have less than τ_v active neighbors. Or, equivalently, they have $\kappa_v = d_v - \tau_v + 1$ or more inactive neighbors. Since the dynamics are irreversible,

activating a node with at least τ_v active neighbors may be seen as removing an inactive node with fewer than $d_v - \tau_v$ nodes from the graph. Thus we are back to *bootstrap percolation as a particular version of the threshold model* with the choice $\tau_v = d_v - k + 1$. In that sense, destroying the k-core is equivalent to the activation of the whole graph. In other words, the set of nodes that leads to the destruction of the k-core is the same set of nodes, that lead, once activated, to the activation of the whole graph (subject to the above threshold dynamics). Following the literature [Rei12; CFKR15; GS15] the smallest such set is referred to as the **minimal contagious set**.

The construction of minimal contagious sets has been subject to extensive studies in recent years. In a general graph this decision problem is known to be NP-hard [DR09]. However, even evaluating its size poses major difficulties. To review briefly the theoretical works most related to our contribution we start with the special case of $k = 2$ that has been studied more thoroughly than $k \geq 3$. The choice of $k = 2$ leads to the removal of the 2-core, i.e. removal of all cycles, and is therefore referred to as the decycling problems. A series of rigorous works analyzes algorithms that are leading to a the previously best known rigorous bounds on the size of the decycling set in random regular graphs [BWZ02; HW08]. Another series of works deals with the cavity method and estimated values for the decycling number that are expected to be exact or extremely close to exact [Zho13; GS15; BDSZ16]. The case of contagious sets with $k \geq 3$ is studied less broadly, but the state-of-the-art is similar to the decycling problem. Rigorous upper bounds stem from analysis of greedy algorithms [CFKR15].

In Chap. 6 we will show that the questions can be attacked by definition of an appropriate dynamical process.

1.3 Optimization and Statistical Inference

The previous two sections introduced probabilistic models and have shown that they are most conveniently expressed in terms of graphical models, or graphs. When the structure of these graphs is subject to randomness several essential properties emerge. Some concrete applications in this section will provide a further justification for the indispensability of random graphs. Referring to Fig. 1.1 we have, so far, mainly dealt with questions that fall into the category of "probability". This section is devoted to the other branch in that schema: "inference".

In this section we shall define a broad class of optimization problems on (random) graphs, the so-called *constraint satisfaction problems*. Combining from the last sections the tree-like structure of the sparse random graph ensembles, with the properties of pairwise Markovian random fields we will derive a low complexity optimization algorithm for constraint satisfaction problems on sparse random graphs, known as *belief propagation*, or simply *message passing*. Following this we shall discover that *inference problems* can be treated in a very similar manner.

1.3.1 Constraint Satisfaction Problems

Formally, a **constraint satisfaction problem** (CSP) consists of a set of variables $\{x_1, \ldots, x_N\}$ and a set of constraints $\{\psi_1, \ldots, \psi_M\}$. Each constraint $a \in \{1, \ldots, M\}$ concerns k_a variables $\partial a \subseteq \{1, \ldots, N\}$. We consider variables that take values in some finite, discrete sample space $x_i \in \mathcal{X}$, with $|\mathcal{X}| = q$. The constraints are functions of $\boldsymbol{x}_{\partial a} \in \mathcal{X}^{k_a}$, such that

$$\psi_a : \begin{cases} \mathcal{X}^{k_a} \to \{0, 1\} \\ \boldsymbol{x}_{\partial a} \mapsto \psi_a(\boldsymbol{x}_{\partial a}) \end{cases} . \tag{1.50}$$

If ψ_a evaluates to 1 the constraint is said to be satisfied, if it evaluates to 0 it is said to be violated. A *solution* to a CSP is an assignment that satisfies all the constraints, i.e., when

$$\prod_{a=1}^{M} \psi_a \left(\boldsymbol{x}_{\partial a}\right) = \begin{cases} 1 & \Rightarrow \text{ solution} \\ 0 & \Rightarrow \text{ no solution} \end{cases} \tag{1.51}$$

Comparing this with (1.23) reveals that the normalization constant of the associated JPD just counts the total number of solutions to a CSP

$$Z = \#\text{solutions} \,. \tag{1.52}$$

The principal question is that after the existence of solutions to the CSP of interest. (Q1) does there exist any $\boldsymbol{x}$, such that (1.51) evaluates to one? (Q2) If yes, how many such configurations exist? However, even if the answer to this question is negative, one might be interested in knowing what is the best achievable configuration, such that as many clauses as possible are satisfied? To that end, we introduce a **cost function** that counts the total number of violations in the CSP

$$\mathcal{H}\left(\boldsymbol{x}\right) = \sum_{a} \left[1 - \psi_a \left(\boldsymbol{x}_{\partial a}\right)\right] . \tag{1.53}$$

We rephrase the last question. (Q3) What is the smallest achievable $\mathcal{H}(\boldsymbol{x})$? All these questions can be rephrased in one question: how many solutions are there to the following equation

$$\mathcal{H}\left(\boldsymbol{x}\right) = E \,. \tag{1.54}$$

While this question is difficult to answer in general, methods from statistical physics give us a handle on how to answer this question w.r.t. some ensemble of problems.

From the definition it is clear, that all such CSPs can be represented graphically as a factor graph. A function node is drawn for each constraint and connected to the k_a variable nodes that enter in the argument of ψ_a. Thus the answer will depend on the characteristic of the factor graph at hand. In the spirit of statistical physics, we will consider the situation in which a *random instance* of a CSP is created by drawing this factor graph from some ensemble.

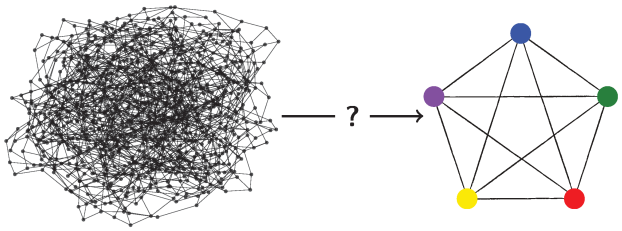

Figure 1.7: A q-coloring of $\mathcal{G}$ is equivalent to a homomorphism of $\mathcal{G}$ to the complete graph K_q. Here $q = 5$.

We will then investigate how the answer to the above question depends on the parameters of the consulted ensemble. We have not treated random factor graphs explicitly, but they are very similar to the random graphs considered in the previous section. The details won't be of importance in this thesis as all the novel contributions and most examples concern either pairwise models or can be understood without specific knowledge on the properties of random factor graph ensembles.

Another question concerns the construction of solution. This question depends very strongly on the specific CSP. We will address this issue for the particular case of circular coloring (see below) later in Chap. 4.

Example: K-SAT

In K-SAT every clause involves K variables and each variable takes boolean values. The functions ψ_a are logical disjunctions of the variables. For instance for $K = 3$ a clause could look like this: $(x_1 \vee \neg x_2 \vee x_5)$. The underlying factor graph model can be of different forms. For instance one might consider the Erdős-Rény graph analogue: one collects the M function nodes on one side and the N variable nodes on the other side and then throws from each function node K edges at random. Another choice would be the random regular graph analogue where we fix the degree of the variable nodes to be D (which only works for $DN = KM$). As in the configuration model, one may then find random pairings between the DN stubs on one side and the KM stubs on the other side.

Example: q-coloring

In q-coloring the variables take one of q colors, $x_i \in \{1, \dots, q\}$ and the interaction is pairwise, such that two adjacent nodes violate a clause if they carry the same color,

$$\mathcal{H}(x_i, x_j) = \delta_{x_i x_j}\,; \quad \text{therefore} \quad \mathcal{H}(\boldsymbol{x}) = \sum_{(ij)\in\mathcal{E}} \delta_{x_i x_j}\,. \tag{1.55}$$

A q-coloring is then an assignment of colors to the nodes of the graph, such that $\mathcal{H}(\boldsymbol{x}) = 0$.

The question of whether a q-coloring exists for a given graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$ can be rephrased. Since each color may be adjacent to any other color, but itself, finding a q-coloring is equivalent to finding a mapping, that preserves the nodes and edges of the graph, onto the

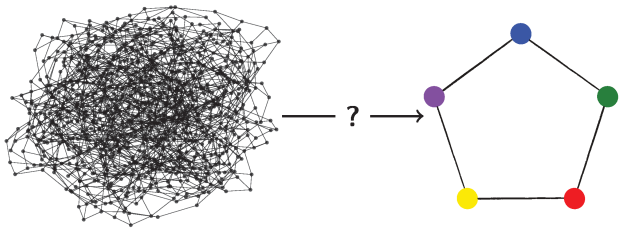

Figure 1.8: A q-circular-coloring of $\mathcal{G}$ is equivalent to a homomorphism of $\mathcal{G}$ to the cycle $\mathcal{C}_q$. Here $q = 5$.

complete graph K_q. Such a mappings are known as homomorphism. Formally, a *homomorphism* from $\mathcal{G}(\mathcal{V}, \mathcal{E})$ to $\mathcal{G}'(\mathcal{V}', \mathcal{E}')$ is a mapping $f : \mathcal{V} \to \mathcal{V}'$, i.e., $i \mapsto f(i)$ for $i \in \mathcal{V}$, that satisfies that if $(ij) \in \mathcal{E}$ then $(f(i)f(j)) \in \mathcal{E}'$. Therefore a homomorphism from $\mathcal{G}(\mathcal{V}, \mathcal{E})$ to K_q is equivalent to a q coloring. Compare also Fig. 1.7.

Example: circular-coloring

Circular coloring is a stronger constraint "coloring problem". Again, nodes can attain the values $x_i \in \{1, \dots, q\}$. Then the graph is q-circular colorable if and only if there exists an assignment of q colors to the nodes such that, if a node $i \in \mathcal{V}$ is of color x_i, then all nodes $j \in \partial i$ are of color $x_j \in \{x_i - 1, x_i + 1\}$ modulo q. Thus

$$\mathcal{H}(x_i, x_j) = 1 - \delta_{x_i\, x_j - 1} - \delta_{x_i\, x_j + 1}\,. \tag{1.56}$$

and

$$\mathcal{H}(\boldsymbol{x}) = \sum_{(ij)\in\mathcal{E}} \left[1 - \delta_{x_i\, x_j - 1} - \delta_{x_i\, x_j + 1}\right]\,. \tag{1.57}$$

In analogy to the q-coloring, the q-circular coloring can be rephrased in terms of a graph homomorphism: it is equivalent to a homomorphism onto a cycle of q colors, $\mathcal{C}_q$, as is depicted in Fig. 1.8

1.3.2 Message Passing

We remain in the setting of the previous section. We have discovered that CSPs are a special form of MRFs

$$P(\boldsymbol{x}) = \frac{1}{Z} \underbrace{\prod_a \psi_a(\boldsymbol{x}_{\partial a})}_{\psi(\boldsymbol{x})}\,. \tag{1.58}$$

with the particular interactions of the form (1.51). The normalization constant Z counts the number of solutions to the CSP and is therefore of major importance. A first approach to evaluate Z just follows from the definition

$$Z = \sum_{\boldsymbol{x}} \psi(\boldsymbol{x}) \tag{1.59}$$

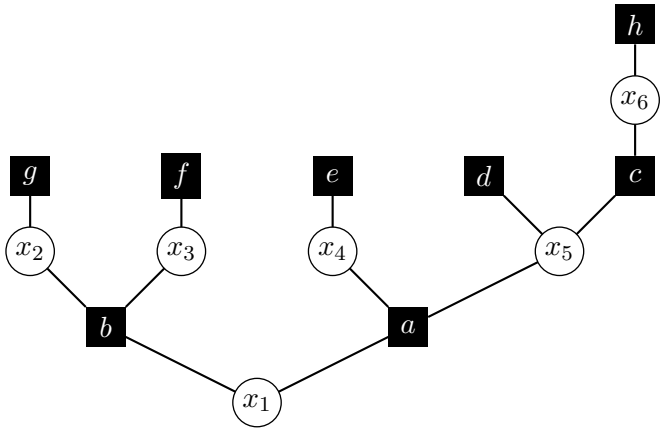

Figure 1.9: Marginals can be evaluated efficiently on sparse factor graphs that are trees.

and would naïvely cost $|\mathcal{X}|^N$ operations. If, however, the underlying factor graph is a *tree*, the sum can be carried out more efficiently.

We consider the example of Fig. 1.9 for which

$$\psi(\boldsymbol{x}) = \psi_a(x_1, x_4, x_5)\,\psi_b(x_1, x_2, x_3)\,\psi_c(x_5, x_6)\,\psi_d(x_5)\,\psi_e(x_4)\,\psi_f(x_3)\,\psi_g(x_2)\,\psi_h(x_6)$$

and evaluate the normalization constant

$$\sum_{x_1,\ldots,x_6} \psi_a(x_1, x_4, x_5)\,\psi_b(x_1, x_2, x_3)\,\psi_c(x_5, x_6)\,\underbrace{\psi_d(x_5)}_{\hat{m}_{d\to5}(x_5)}\underbrace{\psi_e(x_4)}_{\hat{m}_{e\to4}(x_4)}\underbrace{\psi_f(x_3)}_{\hat{m}_{f\to3}(x_3)}\underbrace{\psi_g(x_2)}_{\hat{m}_{g\to2}(x_2)}\underbrace{\underbrace{\psi_h(x_6)}_{\hat{m}_{h\to6}(x_6)}}_{\hat{m}_{h\to6}(x_6)}$$

$$= \sum_{x_1,\ldots,x_5} \psi_a(x_1, x_4, x_5)\,\psi_b(x_1, x_2, x_3)\,\hat{m}_{d\to5}(x_5)\hat{m}_{e\to4}(x_4)\hat{m}_{f\to3}(x_3)\hat{m}_{g\to2}(x_2)\underbrace{\sum_{x_6}\psi_c(x_5, x_6)\,m_{6\to c}(x_6)}_{\hat{m}_{c\to5}(x_5)}$$

$$= \sum_{x_1,\ldots,x_3} \psi_b(x_1, x_2, x_3)\,\underbrace{\hat{m}_{f\to3}(x_3)}_{m_{3\to b}(x_3)}\underbrace{\hat{m}_{g\to2}(x_2)}_{m_{2\to g}(x_2)}\underbrace{\sum_{x_4,x_5}\underbrace{\hat{m}_{e\to4}(x_4)}_{m_{4\to a}(x_4)}\psi_a(x_1, x_4, x_5)\,\underbrace{\hat{m}_{d\to5}(x_5)\hat{m}_{c\to5}(x_5)}_{m_{5\to a}(x_5)}}_{\hat{m}_{a\to1}(x_1)}$$

$$= \sum_{x_1}\hat{m}_{a\to1}(x_1)\underbrace{\sum_{x_2,x_3}\left[\psi_b(x_1, x_2, x_3)\,m_{3\to b}(x_3)m_{2\to g}(x_2)\right]}_{\hat{m}_{b\to1}(x_1)}$$

$$= \sum_{x_1}\hat{m}_{a\to1}(x_1)\hat{m}_{b\to1}(x_1)\,.$$

Note that each time we push a sum to the end, the cost of that sum is equal to the cardinality of $\mathcal{X}$ to the power of the number of arguments that we are summing over plus

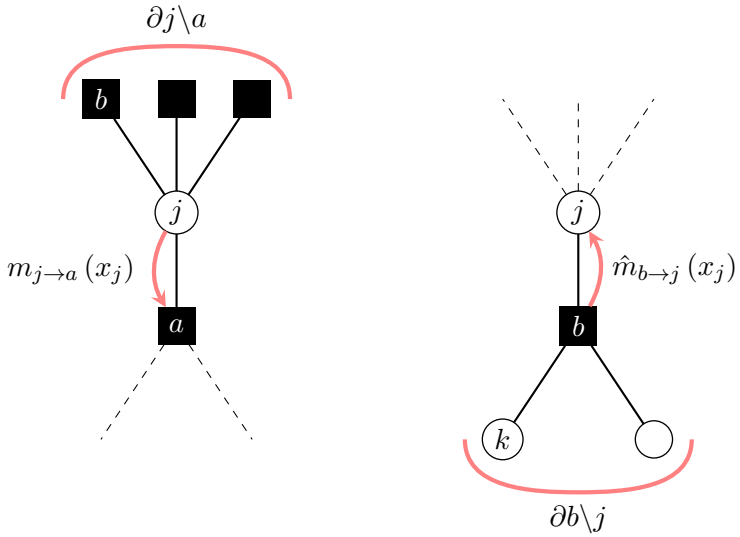

Figure 1.10: The pushed sums can be interpreted as directed messages that are passed from variable nodes to function nodes (left) and vice versa (right).

one for the variable that appears in the sum, but is not summed over, i.e., the degree of the associated function node. The leading order contribution is thus typically $|\mathcal{X}|^{d_{\max}}$, with $d_{\max}$ the maximum degree of the function nodes. Further more it is not only possible to obtain Z from pushing the sums forward, but also all the marginals. One can pick any variable node x_i and obtain its marginal $\sum_{\boldsymbol{x}\backslash i} \psi(\boldsymbol{x})$ by successive summation over the leafs of the tree that is rooted in x_i.

Apparently, the following pattern of **belief propagation** (BP) holds[1]

$$\begin{aligned} m_{j\to a}(x_j) &= \frac{1}{Z^{j\to a}} \prod_{b\in\partial j\backslash a} \hat{m}_{b\to j}(x_j) \\ \tilde{m}_{b\to j}(x_j) &= \frac{1}{Z^{b\to j}} \sum_{\boldsymbol{x}_{\partial b\backslash j}} \psi_b(\boldsymbol{x}_{\partial b}) \prod_{k\in\partial b\backslash j} m_{k\to b}(x_k) \end{aligned} \tag{1.60}$$

and the marginals are obtained from

$$m_j(x_j) = \frac{1}{Z^j} \prod_{b\in\partial j} \hat{m}_{b\to j}(x_j)\ . \tag{1.61}$$

[1] To find a solution to these equations, they must be iterated. In this case the equations are furnished with iteration indices as follows (with $\mathcal{F}_1$ and $\mathcal{F}_2$ represent the r.h.s. of (1.60))

$$\begin{aligned} m^{t+1}_{j\to a}(x_j) &= \mathcal{F}_1(\{\hat{m}^t_{b\to j}(x_j)\}) \\ \tilde{m}^t_{b\to j}(x_j) &= \mathcal{F}_2(\{m^t_{k\to b}(x_k)\})\,. \end{aligned}$$

Here we have additionally introduced normalizations, such that the above beliefs are probability distributions. The **messages** or **beliefs**, $m_{j\to a}(x_j)$ and $\tilde{m}_{a\to j}(x_j)$, can be obtained by induction from consideration of the two elementary building blocks in Fig. 1.10. For a tree the procedure is simple: one starts at the leaves and propagates the beliefs down towards the root. Therefore different messages, incoming to the same node, are **conditionally independent**. The equations (1.60) are exact and provide the correct marginals and can be used to compute Z.

If the underlying factor graph is not a tree the equations in (1.60) do no longer hold. If, however, it is *locally tree-like* we may expect that the BP equations still provide a good approximative solution. If the equations (1.60) are applied on graphs that are no trees one does sometimes speak of **loopy belief propagation**. In such a case, one typically initializes the messages according to some distribution and iterates (1.60) until a fixed point is reached. Different initializations and iterative schemes may be used. If the graph is not a tree different schemes/initializations may lead to different results. This point will be addressed in detail in Chap. 4 and 5.

Box 1.3: Belief Propagation Equations for Circular Coloring

Recalling that

$$\mathcal{H}(\boldsymbol{x}) = \sum_{(ij)\in\mathcal{E}} \left[1 - \delta_{x_i\, x_j - 1} - \delta_{x_i\, x_j + 1}\right] , \tag{1.62}$$

and using that

$$\Psi_{jk}(x_j, x_k) = \lim_{\beta\to\infty} \mathrm{e}^{-\beta\mathcal{H}(x_i, x_j)} , \tag{1.63}$$

we can compute the message update rule from its definition (1.69). Since $x_i \in \{1, \ldots, q\}$ and to save the eyes, we write

$$m_{x_i}^{i\to j} = \frac{1}{Z^{j\to i}} \prod_{k\in\partial j\setminus i} \sum_{x_k} \psi_{jk}(x_j, x_k)\, m_{x_k}^{k\to j} \tag{1.64}$$

$$= \frac{1}{Z^{i\to j}} \prod_{k\in\partial i\setminus j} \sum_{x_k} \mathrm{e}^{-\beta(1-\delta_{x_i,\, x_k - 1} - \delta_{x_i,\, x_k + 1})} m_{x_k}^{k\to i} \tag{1.65}$$

$$= \frac{1}{Z^{i\to j}} \prod_{k\in\partial i\setminus j} \left[m_{x_i+1}^{k\to i} + m_{x_i-1}^{k\to i} + \mathrm{e}^{-\beta} \sum_{x_k\setminus\{x_k\pm 1\}} m_{x_k}^{k\to i} \right] \tag{1.66}$$

$$= \frac{1}{Z^{i\to j}} \prod_{k\in\partial i\setminus j} \left[\mathrm{e}^{-\beta} + (1 - \mathrm{e}^{-\beta})(m_{x_i+1}^{k\to i} + m_{x_i-1}^{k\to i})\right] . \tag{1.67}$$

In the zero temperature limit one obtains

$$m_{x_i}^{i\to j} = \frac{1}{Z^{i\to j}} \prod_{k\in\partial i\setminus j} \left[m_{x_i+1}^{k\to i} + m_{x_i-1}^{k\to i}\right] , \tag{1.68}$$

as it should be.

Density Evolution for Pairwise Models

For pairwise models the first equation in (1.60) is trivially just passing on the message. We may fully eliminate the intermediate step and write

$$m_{j\to i}(x_j) = \frac{1}{Z^{j\to i}} \prod_{k\in\partial j\setminus i} \sum_{x_k} \psi_{jk}(x_j, x_k)\, m_{k\to j}(x_k) =: \mathcal{F}\left(\{m_{k\to j}(x_k)\}\right) . \tag{1.69}$$

These equations are to be iterated on a single graph until convergence. The resulting fixed point (if it exists) depends on the graph instance and different graphs from the same ensemble may lead to different fixed points; Together they form a distribution over the set of fixed points $\mathcal{P}(m_{j\to i})$. In order to to give an adequate description w.r.t. the graph ensemble it is necessary to consider the full distribution over the fixed points.

Once we consider the full ensemble, each edge carries the same stochastic properties. In other words, the distribution $\mathcal{P}_{j\to i}(m_{j\to i})$ does not depend on the particular edge $j \to i$, but only its typical properties. In analogy to (1.69) one solves the following functional fixed point equation, known as **cavity equations** in statistical physics,

$$\mathcal{P}_0(m) = \sum_{l=1}^{\infty} Q(l) \int \Big[\prod_{j}^{l} \mathcal{P}_j(m_j)\mathrm{d}m_j\Big]\, \delta\Big(m - \mathcal{F}(\{m_j\})\Big) . \tag{1.70}$$

Here the subscript denotes some arbitrary edge in the ensemble and l indicates the number of incoming messages, which is distributed according to the **excess degree distribution**

$$Q(l) = \frac{(l+1)P(l+1)}{\sum_{l=0}^{\infty} lP(l)} , \tag{1.71}$$

with $P(d)$ being the degree distribution. This is because the probability that an edge connects to a node of degree d is equal to $d \cdot P(d)$.

Solving the equations (1.70) analytically is in general very difficult. However, they can be solved by sampling: $\mathcal{P}(m)$ is approximated by an empirical distribution of $\mathcal{N}$ messages (population). Then equation (1.70) is solved as follows. From the population one selects an element m_i uniformly at random. Subsequently an integer $l \sim Q(l)$ is drawn and one samples l further messages $\{m_{j_1}, \ldots, m_{j_l}\}$ from the population. Next one replaces $m_i \leftarrow \mathcal{F}(\{m_{j_1}, \ldots, m_{j_l}\})$. This procedure is repeated until convergence.

1.3.3 Bayesian Inference

The inference settings that will be of interest to us consist of the following setting. Consider a set of variables $\boldsymbol{x} = \{x_1, \ldots, x_N\}$ that are observed through some imperfect procedure, resulting in $\boldsymbol{y} = \{y_1, \ldots, y_M\}$. Given $\boldsymbol{y}$ the aim is to recover $\boldsymbol{x}$.

In the **Bayesian approach** to inference it is assumed that the process that brings $\boldsymbol{x}$ to $\boldsymbol{y}$ is best described by a probabilistic model $P(\boldsymbol{y} \mid \boldsymbol{x})$ and that our knowledge/beliefs about

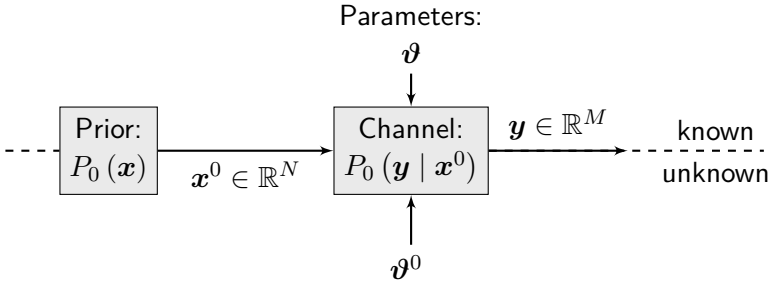

Figure 1.11: The teacher-student scenario.

$\boldsymbol{x}$ is encoded in some JPD $P(\boldsymbol{x})$. It then follows from **Bayes' Theorem** (1.8) that the probability of $\boldsymbol{x}$, given the data (observations) $\boldsymbol{y}$, is

$$P(\boldsymbol{x} \mid \boldsymbol{y}) = \frac{P(\boldsymbol{y} \mid \boldsymbol{x})\, P(\boldsymbol{x})}{P(\boldsymbol{y})}\,. \tag{1.72}$$

Since this formula is at the heart of Bayesian inference each term carries its own name:

- $P(\boldsymbol{x} \mid \boldsymbol{y})$ — Posterior
- $P(\boldsymbol{y} \mid \boldsymbol{x})$ — Channel/Likelihood
- $P(\boldsymbol{x})$ — Prior
- $P(\boldsymbol{y})$ — Partition function/Likelihood

Throughout the thesis we will consider inference settings in the **teacher-student scenario**, which consists of the following framework. The teacher generates a realization of the variables by sampling $\boldsymbol{x}^0 \sim P_0(\boldsymbol{x})$. The sample $\boldsymbol{x}^0$ forms the **ground truth**. From $\boldsymbol{x}^0$ the teacher then constructs the data $\boldsymbol{y}$ according to some probabilistic prescription $P_0(\boldsymbol{y} \mid \boldsymbol{x}^0)$. The data, and some information about the two distributions $P_0(\boldsymbol{x})$ and $P_0(\boldsymbol{y} \mid \boldsymbol{x})$ is then handed to the student. The students goal is to recover $\boldsymbol{x}^0$ from $\boldsymbol{y}$, together with the information it was provided about $P_0(\boldsymbol{x})$ and $P_0(\boldsymbol{y} \mid \boldsymbol{x})$. A neat representation of the procedure can be given in terms of the a flowchart, as in Fig. 1.11.

Since perfect recovery of $\boldsymbol{x}^0$ may be impossible the aim will be to obtain an estimate that is as close as possible to $\boldsymbol{x}^0$. To that end, we introduce an estimator of $\boldsymbol{x}^0$. An **estimator** is a functional prescription of the form $\hat{\boldsymbol{x}} = f(\boldsymbol{y})$. Constructing the function, $f(.)$, is part of the students task. In order to give a definite meaning to "as close as possible" we introduce the Bayesian analogue of a cost function, the **Bayes risk** (or average cost function)

$$\mathcal{R}_0(\hat{\boldsymbol{x}} \mid \boldsymbol{y}) = \mathbb{E}_{\boldsymbol{x}^0|\boldsymbol{y}}\left[\mathcal{H}\left(\boldsymbol{x}^0, \hat{\boldsymbol{x}}\right)\right]\,. \tag{1.73}$$

Where $\mathbb{E}_{\boldsymbol{x}^0|\boldsymbol{y}}$ denotes the average with respect to the *teachers* posterior distribution and $\mathcal{H}(\boldsymbol{x}^0, \hat{\boldsymbol{x}})$ is some cost function. The Bayes risk provides the right notion of "closeness", if (only) the

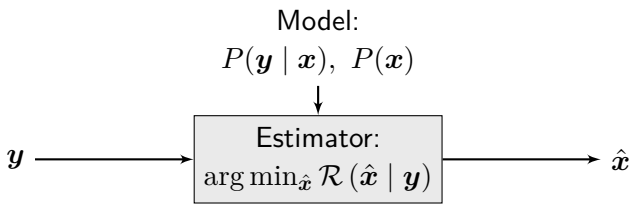

Figure 1.12: In the teacher-student scenario the estimator can be computed, once a model is specified. The model contains the specification of a channel and a prior. These need not match the teachers model.

distributions $P_0(\boldsymbol{x})$ and $P_0(\boldsymbol{y} \mid \boldsymbol{x})$ and the data $\boldsymbol{y}$ are known. The **Bayes optimal estimator** is the one that minimizes the Bayes risk

$$\hat{\boldsymbol{x}}^{\mathcal{H}}\left(\boldsymbol{y}\right) = \arg\min_{\hat{\boldsymbol{x}}} \mathcal{R}_0\left(\hat{\boldsymbol{x}} \mid \boldsymbol{y}\right) . \tag{1.74}$$

The student (inference) aims to reverse the process in Fig. 1.11 by computation of the optimal estimator. However, it might not have full information of the teachers model, $P_0(\boldsymbol{x})$ and $P_0(\boldsymbol{y} \mid \boldsymbol{x})$, and therefore might have to resort to a suboptimal reconstruction scheme. Let us assume that the students best guess of the teachers model is embodied in $P(\boldsymbol{x})$ and $P(\boldsymbol{y} \mid \boldsymbol{x})$; it then attempts to minimize the **mismatched Bayes risk**

$$\mathcal{R}\left(\hat{\boldsymbol{x}} \mid \boldsymbol{y}\right) = \mathbb{E}_{\boldsymbol{x}|\boldsymbol{y}}\left[\mathcal{H}\left(\boldsymbol{x}, \hat{\boldsymbol{x}}\right)\right] \tag{1.75}$$

which will result in a **mismatched estimator**

$$\hat{\boldsymbol{x}}\left(\boldsymbol{y}\right) = \arg\min_{\hat{\boldsymbol{x}}} \mathcal{R}\left(\hat{\boldsymbol{x}} \mid \boldsymbol{y}\right) . \tag{1.76}$$

Since each cost function $\mathcal{H}(\boldsymbol{x}, \hat{\boldsymbol{x}})$ determines a particular estimator, some common cost functions that will be used within the context of this thesis are now presented.

Minimum mean square error (MMSE) The cost function $\mathcal{H}(\boldsymbol{x}^0, \hat{\boldsymbol{x}})$ is the square error between the ground truth and the estimate

$$\mathcal{H}(\boldsymbol{x}^0, \hat{\boldsymbol{x}}) = \sum_{i=1}^{N} (x_i^0 - \hat{x}_i)^2 \tag{1.77}$$

and the associated Bayes risk (1.73) takes the form

$$\begin{aligned} \mathcal{R}_0(\hat{\boldsymbol{x}} \mid \boldsymbol{y}) &= \sum_{i=1}^{N} \int \mathrm{d}\boldsymbol{x}^0\, P(\boldsymbol{x}^0 \mid \boldsymbol{y})\, (x_i^0 - \hat{x}_i)^2 && (1.78) \\ &= \sum_{i=1}^{N} \int \mathrm{d}x_i^0\, P(x_i^0 \mid \boldsymbol{y})\, (x_i^0 - \hat{x}_i)^2 . && (1.79) \end{aligned}$$

With the conditional marginal $P(x_i^0 \mid \boldsymbol{y}) \coloneqq \int \mathrm{d}\boldsymbol{x}_{\setminus i}^0 P(\boldsymbol{x}^0 \mid \boldsymbol{y})$. Differentiation w.r.t. x_i^0 and minimization yields

$$\hat{x}_i = \int \mathrm{d}x_i^0 P(x_i^0 \mid \boldsymbol{y})\, x_i^0 \quad \Leftrightarrow \quad \hat{\boldsymbol{x}}^{\mathrm{MMSE}} = \int \mathrm{d}\boldsymbol{x}^0 P(\boldsymbol{x}^0 \mid \boldsymbol{y})\, \boldsymbol{x}^0 \,. \tag{1.80}$$

The optimal estimator is the mean of the marginal posterior.

Maximum mean overlap (MMO) Consider as a cost function

$$\mathcal{H}(\boldsymbol{x}^0, \hat{\boldsymbol{x}}) = \sum_{i=1}^{N} \Big(1 - \delta(x_i^0 - \hat{x}_i)\Big) \,, \tag{1.81}$$

resulting in

$$\mathcal{R}_0(\hat{\boldsymbol{x}} \mid \boldsymbol{y}) = N - \sum_{i=1}^{N} P(\hat{x}_i \mid \boldsymbol{y}) \,, \tag{1.82}$$

which is minimized by the MMO estimator

$$\hat{x}_i^{\mathrm{MMO}} = \arg\max_{\hat{x}_i} P(\hat{x}_i \mid \boldsymbol{y}) \,. \tag{1.83}$$

The MMO estimator is closely related to the **minimum bitwise error (MBE)**, which is a special case of the MMO estimator, when x_i^0 and $\hat{x}_i$ only attain binary values.

Maximum a-posteriori (MAP) This is another close relative to the MMO estimator. It can be formally derived from $\mathcal{H}(\boldsymbol{x}^0, \hat{\boldsymbol{x}}) = 1 - \delta(\boldsymbol{x}^0 - \hat{\boldsymbol{x}})$; The MAP estimator thus simply maximizes over the a-posteriori probability

$$\hat{\boldsymbol{x}} = \arg\max_{\hat{\boldsymbol{x}}} P\left(\hat{\boldsymbol{x}} \mid \boldsymbol{y}\right) \,. \tag{1.84}$$

High Dimensional Inference

Typically each estimator brings its own pros and cons. The MMSE and MMO estimators both depend upon the computation of the marginal a-posteriori probability distribution, which requires the evaluation of the N-dimensional integral. The MAP estimator requires the solution of an N-dimensional optimization problem.

The difficulty of optimization in high dimensions, i.e. when N is large, depends crucially on the shape of the target function under consideration. For example, if $P\left(\boldsymbol{x} \mid \boldsymbol{y}\right)$ is convex, optimization is usually feasible, even in very high dimensional spaces, because every local minimum is also a global minimum. If, however, the target function is non-convex optimization becomes cumbersome. Many local minima might exist, that are not global minima, which renders gradient based methods prone to failure.

For high-dimensional problems conventional integration methods require a number of discretization points that grows exponentially with N. Unless the integration can be carried out

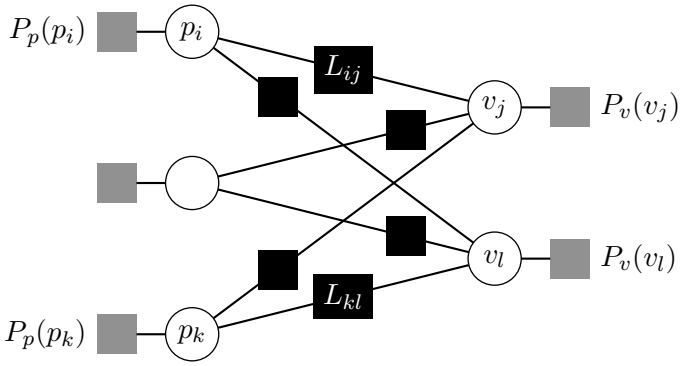

Figure 1.13: Factor graph representation of the dense Dawid-Skene model.

analytically, one must therefore resort to alternative approaches. We have already encountered a setting in which *message passing* provides an alternative approach if the underlying graphical model is sparse. This evokes the question how the underlying graphical model of an inference problem looks like? This question cannot be answered in general terms. However, many interesting settings of inference problems turn out to have dense graphical models associated to them. Although this might seem to present a difficulty (as we loose the desired tree-like structure of the graph), it will prove helpful. In dense models the central limit theorem will lead to an *approximate message passing* algorithm that will admit a low complexity reconstruction of the marginals. Another approach to solve such high dimensional integrals is through Monte Carlo sampling, which will be presented later in Chap. 2.

In the next section we present an inference problem that is interesting to study in both regimes.

1.3.4 Crowdsourcing

Crowdsourcing is a common strategy to categorize data through the contribution of many individuals ("workers"). Such a strategy is often necessary when it is difficult for computers to label data, but of little effort for humans. The development of large-scale crowdsourcing platforms, such as Amazon's MTurk, has popularized crowdsourcing as a simple approach to such problems. Consider N workers, each of them is assigned a subset of a total of M questions/tasks (the data). The workers answer to each of the assigned questions according to their abilities and will. Typically, the received answers are not unambiguous and post-processing has to be performed in order to infer the true information. Now we introduce a model in which each worker can be characterized by a single reliability parameter. It denotes the probability that a worker gives the correct answer to a task. This is the so-called **Dawid-Skene (DS) model** [DS79].

The true label of task j is denoted $v_j^0 \in \{\pm 1\}$. Worker i provides answers to $|\partial i|$ selected tasks, $\partial i = \{j_1, \dots, j_{|\partial i|}\} \subseteq \{1, \dots, M\}$. The answers/labels are denoted L_{ij}, with $L_{ij} \in \{\pm 1\}$ for $j \in \partial i$. The questions for which no answers are provided by worker i,

i.e., the questions in the set $\{1,\dots,M\}\backslash\partial i$, are assigned a null label $L_{ij}=0$ for $j\notin\partial i$. In the DS model each worker is characterized by a single parameter, p_i^0, the probability that the correct answer is provided. Thus the channel takes the following form

$$P_0\left(L_{ij}\mid p_i^0,v_j^0\right)=p_i^0\,\delta(L_{ij}-v_j^0)+(1-p_i^0)\,\delta(L_{ij}+v_j^0)\,,\qquad\text{for } j\in\partial i\,. \tag{1.85}$$

Provided the label matrix L_{ij}, the objective is to infer v_j^0 and p_i^0 in terms of MBE and MMSE respectively. It follows from Sec. 1.3.3 that the optimal estimators require the computation of the marginal distributions of the posterior. We consider the teacher-student scenario in which p_i^0 is sampled from some distribution $P_{p^0}(p_i)$, such that the p_i^0 are i.i.d. samples from a common distribution. Similarly $v_j^0\sim P_{v^0}(v_j)$. From Bayes' Theorem we obtain

$$P\left(\boldsymbol{p},\boldsymbol{v}\mid\boldsymbol{L}\right)=\frac{1}{Z(\boldsymbol{L})}\prod_{1\le j\le M}P_v\left(v_j\right)\prod_{1\le i\le N}P_p\left(p_i\right)\prod_{1\le j\le M,\,1\le i\le N}P\left(L_{ij}\mid p_i,v_j\right)\,. \tag{1.86}$$

If we associate a variable node to each p_i and v_j, it is not difficult to see that the associated graphical model can be drawn as in Fig. 1.13 (we draw the case in which $\partial i=\{1,\dots,M\}$). The factor graph has a **bipartite** structure: each edge connects a node from the set $\{v_j\}$ to a node in the set $\{p_i\}$ (and vice versa), but never directly connects two elements from the same set. If we denote the total number of outgoing connections from the workers (tasks) side as lN (rM) it is clear that $lN=rM$ must hold. The parameter l (r) plays the role of an average degree. If $l\propto M$ ($r\propto N$) we are in the dense regime; if $l,r\propto 1$ we are in the sparse regime.

We shall consider $M,N\to\infty$, but $\alpha=\frac{M}{N}=\Theta(1)$. For each of the two regimes (dense and sparse) there is a particularly interesting scaling region for p_i. We can always write

$$p_i=\frac{1}{2}+\theta_i \tag{1.87}$$

and consider θ_i as our new parameter of interest. In the **sparse regime**, where $|\partial j|=\Theta(1)$ answers are obtained per task and $|\partial j|=\Theta(1)$ answers are provided by each worker, it is intuitive, that the whole region θ_i will be interesting to explore, as a function of the average degrees. If r is small, little information is obtained per task, unless the $\theta_i\approx 1/2$. On the contrary, in the **dense regime**, where one obtains $\Theta(N)$ answers to each task, inference would be trivially easy in this region. Choosing $\theta_i=\Theta(1/\sqrt{N})$ will lead to interesting results, that are neither trivially easy, nor trivially hard (as it would be the case for a $1/N$ scaling).

In analogy to (1.60) the BP equations for this model read

$$\begin{aligned}
\tilde{m}_{ij\to j}(v_j) &= \frac{1}{Z_v^{ij\to j}} \int \mathrm{d}\theta_i \, n_{i\to ij}(\theta_i) P(L_{ij} \mid \theta_i, v_j) \\
n_{i\to ij}(\theta_i) &= \frac{1}{Z_\theta^{i\to ij}} \, P_\theta(\theta_i) \prod_{k\in\partial i\setminus j} \tilde{n}_{ik\to i}(\theta_i) \\
\tilde{n}_{ik\to i}(\theta_i) &= \frac{1}{Z_\theta^{ik\to i}} \sum_{v_k} m_{k\to ik}(v_k) P(L_{ik} \mid \theta_i, v_k) \\
m_{k\to ik}(v_k) &= \frac{1}{Z_v^{k\to ik}} \, P_v(v_k) \prod_{l\in\partial k\setminus i} \tilde{m}_{lk\to k}(v_k) \,.
\end{aligned} \tag{1.88}$$

However, they are quite unfeasible to work with: (a) the variables p_i are in general continuous, which requires the computation of an integral to obtain $\hat{m}_{ij\to j}(v_j)$ and (b) one has to deal with $\Theta(NM)$ messages in the dense case. One possible simplification is obtained by absorbing all $p_{\partial j}$ into a common function node. This is done by integration over the $\boldsymbol{p}$ in (1.86) and is the strategy followed by [LPI12; OOSY16b; OOSY16a]. Owing to the binary labels the resulting BP equations can be solved more efficiently. However, this approach becomes unfeasible in the dense regime because too many messages need to be computed. In the dense regime it is possible to derive a more general algorithm that does not depend upon the integration over $\boldsymbol{p}$ and only requires the computation of $\Theta(N)$ beliefs. This approach is known as approximate message passing and relies on the fact that the messages are effectively Gaußian if the underlying factor graph is densely connected.

1.3.5 Approximate Message Passing (AMP)

The AMP algorithm is closely related to the **Thouless-Anderson-Palmer (TAP) equations** [TAP77] from the theory of spin glasses, with correct time indices [Bol14; ZK16]. AMP can be derived starting from the BP equations (1.88) for the graphical model in Fig. 1.13. The following two simplifications of BP are then made. First, the BP messages are replaced by their means and variances which eradicates the necessity of tracking a whole function for each message. Secondly, each (mean and variance) *message* is replaces by its *marginal* version, reducing the complexity from $O(N^2)$ messages to $O(N)$ marginals.

First we show that the *dense limit* of the DS model belongs to a larger class of low-rank matrix factorization problems that have recently been analyzed in [LKZ17]. To that end we express the channel, $P(L_{ij} \mid p_i, v_j)$, in terms of θ_i, v_j and the fraction of unanswered questions, ρ,

$$\begin{aligned}
P\left(L_{ij} = \pm 1 \mid \theta_i,\ v_j\right) &= (1-\rho) \cdot \frac{1}{2} \cdot \left(1 \pm \sqrt{\frac{\nu}{N}}\, \theta_i v_j\right) . \\
P\left(L_{ij} = 0 \mid \theta_i,\ v_j\right) &= \rho
\end{aligned} \tag{1.89}$$

Here we have extracted the $1/\sqrt{N}$ scaling of θ_i explicitly, such that $\theta_i = \Theta(1)$, but $p_i = 1/2 + \Theta(1/\sqrt{N})$. We further introduced the parameter ν which allows to tune between

different signal to noise ratios. The form of (1.89) assumed that the fraction of un-answered questions, ρ, is independent of (i,j). In the dense regime we are interested in the limit where $N \to \infty$ and all the other parameters $\theta_i, \nu, \alpha, \rho = \Theta(1)$. The link to low-rank matrix factorization becomes clear if one introduces the rank-1 matrix

$$\boldsymbol{w} \coloneqq \frac{\boldsymbol{\theta}\,\boldsymbol{v}^{\intercal}}{\sqrt{N}}\,. \tag{1.90}$$

The derivation of the AMP equations can be done in more general terms and it is not necessary to restrict ourselves to the particular channel (1.89). Instead we will consider the general case of

$$P\left(L_{ij} \mid w_{ij}\right) = \exp\left(g(L_{ij}, w_{ij})\right) \tag{1.91}$$

with $w_{ij} = O(1/\sqrt{N})$. For crowdsourcing we have

$$g(L_{ij}, w_{ij}) = \begin{cases} \log\left(\frac{(1-\rho)}{2}\right) + \log\left(1 \pm \sqrt{\nu} w_{ij}\right) & \text{if } L_{ij} = \pm 1 \\ \log\left(\rho\right) & \text{if } L_{ij} = 0 \end{cases}\,. \tag{1.92}$$

The previous manipulations show that the posterior falls into the same class of problems, considered in [LKZ17]. We will now briefly review the derivation of the AMP equations, first given in [LKZ15], for convenience of the reader. The first step consists in relaxing the BP equations (1.88) such that each message can be re-parametrized in terms of their mean and variance

$$\begin{aligned} \hat{\theta}_{i\to ij} &= \int \mathrm{d}\theta_i n_{i\to ij}(\theta_i)\,\theta_i \\ \sigma_{\theta,\,i\to ij} &= \int \mathrm{d}\theta_i n_{i\to ij}(\theta_i)\,(\theta_i)^2 - \hat{\theta}^2_{i\to ij} \\ \hat{v}_{k\to ik} &= \int \mathrm{d}v_k m_{k\to ik}(v_k)\,v_k \\ \sigma_{v,\,k\to ik} &= \int \mathrm{d}v_k m_{k\to ik}(v_k)\,v_k^2 - \hat{v}^2_{k\to ik}\,. \end{aligned} \tag{1.93}$$

As is shown in Appendix A.4, the messages (1.88) can then be written in a simplified Gaußian form

$$\begin{aligned} n_{i\to ij}(\theta_i) &= \frac{1}{Z_\theta^{i\to ij}}\,P_\theta(\theta_i)\exp\left(B_{\theta,\,i\to ij}\,\theta_i - \frac{1}{2}A_{\theta,\,i\to ij}\,\theta_i^2\right) \\ m_{k\to ik}(v_k) &= \frac{1}{Z_v^{k\to ik}}\,P_v(v_k)\exp\left(B_{v,\,k\to ik}\,v_k - \frac{1}{2}A_{v,\,k\to ik}\,v_k^2\right)\,, \end{aligned} \tag{1.94}$$

where $A_{\theta,\,v}$ and $B_{\theta,\,v}$ can be expressed in terms of the $\hat{\theta}_{i\to ij},\ \sigma_{\theta,\,i\to ij},\ \hat{v}_{k\to ik},\ \sigma_{v,\,k\to ik}$. The

equations can thus be close on the means and variances of the messages (1.93).

$$\begin{aligned}
\hat{\theta}^t_{i\to ij} &= f_\theta\left(A^t_{\theta,\,i\to ij},\ B^t_{\theta,\,i\to ij}\right)\\
\sigma^t_{\theta,\,i\to ij} &= \frac{\partial f_\theta}{\partial B}\left(A^t_{\theta,\,i\to ij},\ B^t_{\theta,\,i\to ij}\right)\\
\hat{v}^{t+1}_{k\to ik} &= f_v\left(A^t_{v,\,k\to ik},\ B^t_{v,\,k\to ik}\right)\\
\sigma^{t+1}_{v,\,k\to ik} &= \frac{\partial f_v}{\partial B}\left(A^t_{v,\,k\to ik},\ B^t_{v,\,k\to ik}\right)\,,
\end{aligned} \tag{1.95}$$

where we have introduced the **input functions**

$$f_x\left(A,\,B\right) = \frac{1}{Z_x(A,B)}\int \mathrm{d}x P_x(x)\mathrm{e}^{-\frac{1}{2}Ax^2+Bx}\,x\,. \tag{1.96}$$

With x indicating either θ or v. Equations (1.95) and (1.96) form the **relaxed BP (rBP)** equations. Note that these equations are solved self-consistently. Such a solution is obtained by iteration, hence we added the time indices.

The rBP equations are a direct consequence of the CLT. Each message is a random variable and since the $O(N)$ incoming messages are only weakly correlated they result in an effective Gaußian field, acting on each variable node. This field is additionally weighted with the prior on each of the sides. Since the outgoing messages only weakly depend on the target node, i.e., $B^t_{x,\,i} = B^t_{x,\,i\to ij} + O(1/\sqrt{N})$ and $A^t_{x,\,i} = A^t_{x,\,i\to ij} + O(1/N)$ (where, again, x stands for θ or v). The marginals can thus be expressed in closed form by **TAPyfication** of the equations in (1.95). The steps are performed in detail in Appendix A.4. We finally obtain the set of **AMP equations** that is independent of the messages and only depends on the marginals

$$\begin{aligned}
B^t_{\theta,\,i} &= \frac{1}{\sqrt{N}}\sum_{k=1}^{M} S_{ik}\hat{v}^t_k - \left(\frac{1}{N}\sum_{k=1}^{M} S^2_{ik}\sigma^t_{v,\,k}\right)\hat{\theta}^{t-1}_i\\
A^t_{\theta,\,i} &= \frac{1}{N}\sum_{k=1}^{M}\left[S^2_{ik}(\hat{v}^t_k)^2 - R_{ik}\left((\hat{v}^t_k)^2+\sigma^t_{v,\,k}\right)\right]\\
\hat{\theta}^t_i &= f_\theta\left(A^t_{\theta,\,i},\,B^t_{\theta,\,i}\right)\\
\sigma^t_{\theta,\,i} &= \frac{\partial f_\theta}{\partial B}\left(A^t_{\theta,\,i},\ B^t_{\theta,\,i}\right)\\
B^t_{v,\,k} &= \frac{1}{\sqrt{N}}\sum_{l=1}^{N} S_{lk}\hat{\theta}^t_l - \left(\frac{1}{N}\sum_{l=1}^{N} S^2_{lk}\sigma^t_{\theta,\,l}\right)\hat{v}^t_k\\
A^t_{\theta,\,k} &= \frac{1}{N}\sum_{l=1}^{N}\left[S^2_{lk}(\hat{\theta}^t_l)^2 - R_{lk}\left((\hat{\theta}^t_l)^2+\sigma^t_{\theta,\,l}\right)\right]\\
\hat{v}^{t+1}_k &= f_v\left(A^t_{v,\,k},\,B^t_{v,\,k}\right)\\
\sigma^{t+1}_{v,\,k} &= \frac{\partial f_v}{\partial B}\left(A^t_{v,\,k},\ B^t_{v,\,k}\right)\,.
\end{aligned} \tag{1.97}$$

Here

$$S_{ij} \coloneqq \left.\frac{\partial \log P_0(L_{ik} \mid w)}{\partial w}\right|^{w=0} \tag{1.98}$$

$$R_{ij} \coloneqq \left(\left.\frac{\partial \log P_0(L_{ik} \mid w)}{\partial w}\right|^{w=0}\right)^2 + \left.\frac{\partial^2 \log P_0(L_{ik} \mid w)}{\partial w^2}\right|^{w=0} . \tag{1.99}$$

The above AMP equations provide a constructive approach to the marginal posterior expectations. Note that the MMO-estimator of the true labels v_j^0 can be extracted from the fixed point of the above equations, $\hat{v}_j^*$, by taking its sign, if the labels are in $\{\pm 1\}$.

Some remarks are appropriate. The additional terms that appear in B_θ and B_v after TAPyfication are **Onsager reaction terms** that correct the mean field contribution from the first sum. The terms S_{ij}^2 and R_{ij} can be replaced by their averages

$$\Delta^{-1} = \mathbb{E}_{P(L_{ij}|w_{ij}=0)}\left[S_{ij}^2\right] \tag{1.100}$$

$$R = \mathbb{E}_{P(L_{ij}|w_{ij}=0)}\left[R_{ij}\right] = 0\,. \tag{1.101}$$

Where the last equality is a consequence of the normalization of the conditional probability: $\int \mathrm{d}L P(L \mid w) = 1 \Rightarrow \int \mathrm{d}L \partial_{w^2} P(L \mid w) = 0$. For the particular case of crowdsourcing we have

$$S_{ij} = L_{ij}\sqrt{\nu} \tag{1.102}$$

$$R_{ij} = L_{ij}^2 \nu - L_{ij}^2 \nu = 0\,. \tag{1.103}$$

In the Bayes optimal setting, where $P_x(x) = P_{x^0}(x)$ (for $x \in \{\theta, v\}$) and $P(L_{ij} \mid w_{ij}) = P_0(L_{ij} \mid w_{ij})$, we have

$$\Delta^{-1} = (1-\rho)\,\nu \tag{1.104}$$

$$R = 0\,. \tag{1.105}$$

1.3.6 State Evolution

In this section we will see how to derive a set of dynamical equations, that tracks the evolution of the AMP equations by means of characteristic order parameters. The AMP equations depend explicitly on the realization of the disorder (for the crowdsourcing this is L_{ij}) through S_{ij} and possibly R_{ij}. Therefore the $B_{\theta,\,i}^t$, $B_{v,\,j}^t$ and $A_{\theta,\,i}^t$, $A_{v,\,j}^t$ are random variables. Recalling that the different messages, incoming to one node are independent by BP assumption we can apply the CLT to the sums on the r.h.s. of the equations for $B_{\theta,\,i}^t$, $B_{v,\,j}^t$. The independence holds only approximately because the underlying graph is not a tree, but the $O(1/\sqrt{N})$ scaling guarantees sufficient independence in the $N \to \infty$ limit. Thus we have

$$B_{\theta,\,i}^t \sim \mathcal{N}\left(\mathbb{E}B_{\theta,\,i}^t,\, \mathbb{E}(B_{\theta,\,i}^t)^2 - (\mathbb{E}B_{\theta,\,i}^t)^2\right) \tag{1.106}$$

$$B_{v,\,j}^t \sim \mathcal{N}\left(\mathbb{E}B_{v,\,j}^t,\, \mathbb{E}(B_{v,\,j}^t)^2 - (\mathbb{E}B_{v,\,j}^t)^2\right) . \tag{1.107}$$

Further more, by the law of large numbers, the r.h.s. of the equation for $A^t_{\theta,\,i}$, $A^t_{v,\,j}$ can be replaced by its average. It remains to compute the first two moments of $B^t_{\theta,\,i}$, $B^t_{v,\,j}$ and the first moment of $A^t_{\theta,\,i}$, $A^t_{v,\,j}$.

It is shown in Appendix A.5 that the first two moments can be expressed in terms of the macroscopic order parameters

$$M^t_\theta = \frac{1}{N}\sum_{i=1}^{N} \hat{\theta}^t_i \theta^0_i \,, \qquad\qquad M^t_v = \frac{1}{M}\sum_{j=1}^{M} \hat{v}^t_j v^0_j \,, \tag{1.108}$$

which leads to

$$\begin{aligned} \mathbb{E}\left[B^t_{\theta,\,i}\right] &= \frac{\alpha}{\Delta} M^t_v \theta^0_i \,, & \mathbb{E}\left[B^t_{v,\,j}\right] &= \frac{1}{\Delta} M^t_u v^0_j \,, \\ \mathbb{E}\left[(B^t_{\theta,\,i})^2\right] &= \frac{\alpha}{\Delta} M^t_v \,, & \mathbb{E}\left[(B^t_{v,\,j})^2\right] &= \frac{1}{\Delta} M^t_\theta v^0_j \,, \\ \mathbb{E}\left[A^t_{\theta,\,i}\right] &= \frac{\alpha}{\Delta} M^t_v \,, & \mathbb{E}\left[A^t_{v,\,j}\right] &= \frac{1}{\Delta} M^t_\theta \,, \end{aligned} \tag{1.109}$$

where

$$\Delta^{-1} = \mathbb{E}_{P_0(L|w^0=0)}\left[\left(\frac{\partial \log P_0(L_{ik} \mid w)}{\partial w}\right)^2\right] . \tag{1.110}$$

Since

$$\hat{\theta}^t_i = f_\theta\left(A^t_{\theta,\,i},\, B^t_{\theta,\,i}\right) \quad \text{and} \quad \hat{v}^{t+1}_k = f_v\left(A^t_{v,\,k},\, B^t_{v,\,k}\right) , \tag{1.111}$$

i.e., the estimates are functions of the order parameters (1.108), the equations can be closed on them

$$M^t_\theta = \mathbb{E}_{\theta^0,W}\left[f_\theta\left(\frac{\alpha\, M^t_v}{\Delta},\, \frac{\alpha\, M^t_v}{\Delta}\theta^0 + \sqrt{\frac{\alpha\, M^t_v}{\Delta}} W\right)\, \theta^0\right] \tag{1.112}$$

$$M^t_v = \mathbb{E}_{v^0,W}\left[f_v\left(\frac{M^t_\theta}{\Delta},\, \frac{M^t_\theta}{\Delta}v^0 + \sqrt{\frac{M^t_\theta}{\Delta}} W\right)\, v^0\right] . \tag{1.113}$$

In the Bayes optimal setting, these equations track the evolution of the AMP equations (1.97) in terms of simple scalar order parameters. This is, in the spirit, reminiscent of what was done in Sec. 1.2.3. The above set of equations will further simplify in the Bayes optimal setting, which will be treated in Chap. 5. We will see in the next chapter that the fixed points of the above equations have a deep physical meaning: the free energy can be expressed in terms of the same order parameters and the minimizers of the free energy correspond to the fixed points of the above equations.

— 2 —

Statistical Physics of Inference Problems

In this chapter some basic concepts from statistical physics are reviewed and disordered systems are introduced. The disorder in the Hamiltonian creates frustration that may lead to non-ergodic behavior of the physical system (below some temperature). It will be shown that inference and constraint satisfaction problems can be considered as disordered systems. The average case behavior of these problems can be extracted from the free energy of the corresponding "physical system".

In classical mechanics the time evolution of a system is the curve that is traced out by the canonical coordinates, $(q_i(t), p_i(t))$. The **phase space** is the set of accessible points $(q_i(t), p_i(t))$. The phase space trajectory is determined by the Hamiltonian equations of motion

$$\dot{q}_i = \frac{\partial \mathcal{H}}{\partial p_i} \,, \quad \dot{p}_i = -\frac{\partial \mathcal{H}}{\partial q_i} \,. \tag{2.1}$$

If the **Hamiltonian**, $\mathcal{H}(q_i(t), p_i(t))$, does not *explicitly* depend on time, the system is said to be in equilibrium and the Hamiltonian corresponds to the total energy of the system.

A statistical approach to a physical system is useful, but not limited (cf. stochastic thermodynamics), when the number of microscopic degrees of freedom is very large. Typically the degrees of freedom grow with the number of constituents, N. In the **thermodynamic limit**, where $N \to \infty$, it is either hopeless, or useless, to keep track of the physical state $\boldsymbol{x}(t) \coloneqq (\boldsymbol{q}(t), \boldsymbol{p}(t))$. Instead it is usually simpler to transition from a microscopic description to a macroscopic one by introduction of a probability distribution over $(\boldsymbol{q}(t), \boldsymbol{p}(t))$ and consideration of averaged quantities with respect to this distribution.

In *thermodynamics* a physical system is described in terms of macroscopic state variables (temperature, energy, volume, ...). The thermodynamic laws, that govern the interplay between these state variables in term of macroscopic state equations, are mostly phenomenological. Statistical physics was born out of the desire to put the laws of thermodynamics on a sound theoretical basis. The challenge of statistical physics is the derivation of the macroscopic laws of thermodynamics from a microscopic consideration of the physical system. In the end of the last chapter, we have already performed such a transition from a microscopic description, in terms of the AMP variables, to a macroscopic description, in terms of the

order parameters. In this chapter we will further elaborate this statistical physics approach to inference and constrained satisfaction problems.

2.1 Equilibrium Statistical Physics

From now on, let us consider the system in equilibrium, with an *instantaneous* microstate denoted as $\boldsymbol{x} = (x_1, \ldots, x_N)$. The fundamental hypothesis in equilibrium statistical physics amounts to the assumption that the system is ergodic, i.e., that it explores all parts of the phase space and that the temporal behavior of the system can be well described by a *distribution* over the instantaneous microstates, $P(\boldsymbol{x})$.

2.1.1 Microcanonical and Canonical Ensembles

In a closed system the total energy

$$E = \mathcal{H}(\boldsymbol{x}) \tag{2.2}$$

is conserved. All $\boldsymbol{x}$ must lay on the hypersurface determined by (2.2). The total number of accessible microstates is

$$\Omega(E) = \int_{E=\mathcal{H}(\boldsymbol{x})} \mathrm{d}\boldsymbol{x}\,. \tag{2.3}$$

This leads to the probability density of the **microcanonical ensemble**

$$P(\boldsymbol{x}) = \begin{cases} \frac{1}{\Omega(E)} & \text{if } \mathcal{H}(\boldsymbol{x}) = E \\ 0 & \text{else .} \end{cases} \tag{2.4}$$

Namely, all microstates of energy E are equally likely. The **entropy** is defined as the log-number of microstates that are tolerated by a given macrostate (here E)

$$S(E) = \ln \Omega(E)\,. \tag{2.5}$$

The above equation is in natural units, where the proportionality constant is set to unity and the additive constant to zero. Note that this can be formally written as the ensemble average

$$S(E) = -\int \mathrm{d}\boldsymbol{x} P(\boldsymbol{x}) \log P(\boldsymbol{x})\,. \tag{2.6}$$

It can be shown that the functional $S[P(\boldsymbol{x})] = -\int \mathrm{d}\boldsymbol{x} P(\boldsymbol{x}) \log P(\boldsymbol{x})$ is the unique choice that fulfills the following requirements that establish that the entropy can be interpreted as a measure of uncertainty:

- It only depends on $P(\boldsymbol{x})$.
- $S[P(\boldsymbol{x}) = \text{const.}] = \max S[P(\boldsymbol{x})]$.

- $S[P(\boldsymbol{x})] = 0$ if $P(\boldsymbol{x}) = \delta(\boldsymbol{x} - \boldsymbol{x}_0)$.
- $S[P(\boldsymbol{x})] = S[P(\sigma(\boldsymbol{x}))]$, where $\sigma(.)$ is some permutation of its argument.
- Additivity: $S[P_A(\boldsymbol{x}) \cup P_B(\boldsymbol{x})] = S[P_A(\boldsymbol{x})] + S[P_B(\boldsymbol{x})]$ if the two systems $A,\ B$ are independent.

The microcanonical ensemble is conceptually nice, but in practice often not applicable to physical systems, as it only applies to situation in which the total energy of the system of interest is fixed. A more "realistic" alternative approach is provided in the **canonical ensemble** in which another control parameter than the energy is constant. In the canonical ensemble the system of interest is in exchange with a *heat bath*, $\mathcal{B}$, an both systems are in thermal equilibrium and share a common temperature, T. The total energy of the combined (closed!) system is $E_0 = E_{\mathcal{B}} + \mathcal{H}(\boldsymbol{x})$. The heat bath has, by definition, much larger energy than our system of interest: $E_{\mathcal{B}} \gg \mathcal{H}(\boldsymbol{x})$. The basic axiom that *all microstates of a common energy hypersurface are equiprobable* also applies in the canonical ensemble and (2.6) still applies.

The Gibbs-Boltzmann measure is the probability measure over microstates $\boldsymbol{x}$ in the canonical ensemble. It can be derived as the measure that maximizes the entropy under the constraint that the average energy of the system is fixed

$$\max - \int \mathrm{d}\boldsymbol{x} P(\boldsymbol{x}) \ln P(\boldsymbol{x}) \tag{2.7}$$

$$\text{s.t. } \mathbb{E}_{\boldsymbol{x}} \left[\mathcal{H}(\boldsymbol{x})\right] = E\,. \tag{2.8}$$

Introducing the Lagrange multiplier β to enforce the average energy and dropping the constant terms, we obtain the following equivalent Lagrange problem

$$\min \left[\int \mathrm{d}\boldsymbol{x} P(\boldsymbol{x}) \log P(\boldsymbol{x}) + \beta \int \mathrm{d}\boldsymbol{x} P(\boldsymbol{x}) \mathcal{H}(\boldsymbol{x}) \right]. \tag{2.9}$$

This can be solved by variation w.r.t. $P(\boldsymbol{x})$, which (after enforcing normalization) yields the **Gibbs-Boltzmann (GB) measure** on the microstates $\boldsymbol{x}$

$$P(\boldsymbol{x}) = \frac{\mathrm{e}^{-\beta \mathcal{H}(\boldsymbol{x})}}{Z(\beta)}\,. \tag{2.10}$$

We see that in the canonical ensemble each microscopic state, $\boldsymbol{x}$, is assigned a certain energy cost in form of the Hamiltonian $\mathcal{H}(\boldsymbol{x})$. Note that β is a Lagrange multiplier that controls the average energy, i.e., $E(\beta)$. As a consequence, low energy states are selected preferably, while high energy states are exponentially more rare. Note that if one imposes the maximum entropy principle, with the average energy constraint, one obtains microstates that do not minimize $\mathcal{H}$ directly, but instead the auxiliary potential $S(E) + \beta(E)E$.

We refer to average w.r.t. the GB measure as

$$\langle \bullet \rangle = \int \mathrm{d}\boldsymbol{x}\ \bullet\ \frac{\mathrm{e}^{-\beta \mathcal{H}(\boldsymbol{x})}}{Z(\beta)}\,. \tag{2.11}$$

2.1.2 Thermodynamic Potentials

General observables $\mathcal{A}(\boldsymbol{x})$ can be expressed in macroscopic form by averaging w.r.t. $P(\boldsymbol{x})$:

$$A(\boldsymbol{\vartheta}) = \langle \mathcal{A}(\boldsymbol{x}) \rangle = \int \mathrm{d}\boldsymbol{x} P(\boldsymbol{x}) \mathcal{A}(\boldsymbol{x}) \,. \tag{2.12}$$

Here we have explicitly expressed A as a function of $\boldsymbol{\vartheta}$ to stress that it may in general depend on some additional parameters (e.g. the parameters of the distribution $P(\boldsymbol{x})$). For example, the average energy is obtained by averaging $\mathcal{H}(\boldsymbol{x})$. Similarly the entropy corresponds to $\mathcal{A}(\boldsymbol{x}) = -\log P(\boldsymbol{x})$. And the partition function formally corresponds to $\mathcal{A}(\boldsymbol{x}) = 1$.

Importantly the **partition function**, that was obtained as the normalization constant of the GB measure,

$$Z(\beta) = \int \mathrm{d}\boldsymbol{x} \mathrm{e}^{-\beta \mathcal{H}(\boldsymbol{x})} \,, \tag{2.13}$$

can now be given a physical meaning. First of all, let us recall that $\beta = 1/T$. In the **high temperature phase** where $\beta \to 0$, it is evident that all states become equiprobable and the partition function counts the number of all microstates. According to the above we are in a phase of maximal entropy with no energy penalization. In the **zero temperature limit**, where $\beta \to \infty$, only those states survive that have smallest accessible energy, E_{gs} and

$$\lim_{\beta \to \infty} \frac{\ln Z(\beta)}{\beta} = -\min_{\boldsymbol{x}} \mathcal{H}\left(\boldsymbol{x}\right) \tag{2.14}$$

This is because the above integral is dominated by its saddle point and so is the average energy $\int \mathrm{d}\boldsymbol{x} \mathcal{H}(\boldsymbol{x}) \mathrm{e}^{-\beta \mathcal{H}(\boldsymbol{x})}$. Therefore the average energy approaches the ground state energy in this case. Note the link to the questions that we posed in section 1.3.1 about CSPs.

What is the meaning of $Z(\beta)$ for generic temperature? The following chain of reasoning reveals that $\log Z(\beta)$ links the average energy with the entropy. To carry out the calculation we assume that we are in the thermodynamic limit, $N \to \infty$. Further we fix β, such that the average energy and the entropy will equilibrate accordingly. And finally it is required that the energy and entropy are extensive quantities, i.e. they are $O(N)$.

$$\begin{aligned} Z\left(\beta\right) &= \int \mathrm{d}E \Omega(E) \mathrm{e}^{-\beta E} \\ &= \int \mathrm{d}E \mathrm{e}^{S(E) - \beta E} \\ &\approx \int \mathrm{d}\epsilon \mathrm{e}^{N(s(\epsilon) - \beta \epsilon)} \\ &\backsimeq \mathrm{e}^{-N \min_{\epsilon}(\beta \epsilon - s(\epsilon))} \\ &= \mathrm{e}^{-N(\beta \epsilon^* - s(\epsilon^*))} \\ &=: \mathrm{e}^{-\beta F} \,. \end{aligned} \tag{2.15}$$

First we replaced the integration over microstates by an integration over all the different Energy hypersurfaces. Then we exploited the definition of the entropy and thereafter the

fact that E and S are extensive. The final integral was then carried out by the saddle point method, which amounts to the fact that the integrand is dominated by its largest value in the thermodynamic limit.

The **free energy**

$$F(\beta) = E - \frac{1}{\beta} S(E) = -\frac{\ln Z(\beta)}{\beta}\,, \tag{2.16}$$

is completely analogue to (2.9) and a physical system at temperature T will take an energy E and entropy S, such that *the free energy $E - S(E)/\beta$ is minimized*. In equilibrium the physical system the thermodynamically optimal states are those that minimize the free energy. Although $F(\beta)$ is the right **physical potential** to look at, in the sense that it has a nice interpretation in thermodynamic terms (it is the accessible energy that can be extracted from the system in as work), it is often beneficial to work with the dimensionless quantity $-\ln Z(\beta)$ instead. With some abuse of notation we will also refer to $-\ln Z(\beta)$ as a free energy. Once $\ln Z(\beta)$ is known, many other observables can be derived from it. As an example, the average energy, E, is obtained from the first derivative

$$E(\beta) = \frac{\int \mathrm{d}\boldsymbol{x} \mathcal{H}(\boldsymbol{x}) \mathrm{e}^{-\beta \mathcal{H}(\boldsymbol{x})}}{\int \mathrm{d}\boldsymbol{x} \mathrm{e}^{-\beta \mathcal{H}(\boldsymbol{x})}} = -\frac{\mathrm{d}}{\mathrm{d}\beta} \ln Z(\beta)\,. \tag{2.17}$$

More generally, $\ln Z(\beta)$ is the moment generating function of $\mathcal{H}(\boldsymbol{x})$.

In general, once $\ln Z(\beta)$ is known, it is possible to access any other thermodynamic potentials from it. To that end, let us consider the **free entropy**, which just differs from the free energy by a minus sign and is a unitless quantity:

$$\Phi(\beta) \coloneqq \ln Z(\beta) = S(E) - \beta E\,. \tag{2.18}$$

The energy, as well as the entropy are **extensive** quantities, i.e., they scale with the number of particles/constituents in the system, N. We have already used the fact in the derivation of the free energy in (2.15). Let us therefore work with the per particle quantity,

$$\phi(\beta) = \lim_{N\to\infty} \frac{\ln Z(\beta)}{N}\,. \tag{2.19}$$

The physical value of $\phi(\beta)$ for a given β corresponds to the maximum (we switched sign) of the per particle quantities of the r.h.s. of (2.18)

$$\phi(\beta) = \max_{\epsilon} \left[s(\epsilon) - \beta\epsilon \right] \tag{2.20}$$

and thus

$$\frac{\mathrm{d}\left[s(\epsilon) - \beta\epsilon \right]}{\mathrm{d}\epsilon} = 0\,, \quad \Leftrightarrow \quad \left.\frac{\mathrm{d}s}{\mathrm{d}\epsilon}\right|_{\epsilon=\epsilon^*} = \beta \tag{2.21}$$

Apparently the above is equivalent to

$$\phi(\beta) = s(\epsilon) - \beta\epsilon \quad \text{such that} \quad \frac{\mathrm{d}s}{\mathrm{d}\epsilon} = \beta \tag{2.22}$$

In the above equation, there are two dependent variables, ϵ and β. For a fixed temperature one should select the energy ϵ, such that the r.h.s. of the above equation is maximized. However, such a procedure might be impractical in certain situations. In particular, we will encounter this issue when investigating the problem of circular coloring in chapter 4. Typically it is much simpler to control the temperature (or β). This presents us with the problem of inverting the above equations. In practice this can be achieved by realizing that the l.h.s. of (2.20) is the **Legendre transformation** of $s(\epsilon)$. It can be inverted, which leads to an optimization problem over β:

$$s(\epsilon) = \min_{\beta} \left[\phi(\beta) + \beta\epsilon\right] . \tag{2.23}$$

This is not surprising if one recalls (2.9). Again, we may write the above in the form

$$s(\epsilon) = \phi(\beta) + \beta\epsilon \quad \text{such that} \quad \epsilon = \frac{\mathrm{d}\phi}{\mathrm{d}\beta} . \tag{2.24}$$

In turn this now allows us to access the energy by tuning the temperature such that the above equations are minimized.

2.2 Disordered Systems

In disordered systems the Hamiltonian is subject to randomness. It becomes a random object. In a system with **quenched disorder** the Hamiltonian explicitly depends on the realization of a RV, say $\boldsymbol{J}$. We write

$$\mathcal{H}\left(\boldsymbol{x} \mid \boldsymbol{J}\right) . \tag{2.25}$$

We have already encountered an example of such a system with explicit disorder in Sec. 1.3.4. In the dense DS model for *crowdsourcing*, the explicit disorder are the labels $\boldsymbol{L}$. In our model of crowdsourcing there is another source of randomness due to the random assignment of tasks to workers. In the dense DS model this randomness is, however, washed out by the CLT (each node only sees an effective mean field when the assignment graph is densely connected). A better example of **self-generated disorder** is *circular coloring*. The associated Hamiltonian (1.57) contains no explicit disorder. However, the underlying interaction network is disordered, if one studies the problem on random graphs. For problems on random graphs, $\mathcal{G}$ plays the role of the disorder $\boldsymbol{J}$.

An undesired consequence of the present disorder in the Hamiltonian is that the physical quantities become functions of the disorder. Fortunately, in the thermodynamic limit, physical properties do not depend on the realization of the disorder. They are **self-averaging**. If the macroscopic observable $A(\boldsymbol{\vartheta} \mid \boldsymbol{J})$

$$\lim_{N\to\infty} A(\boldsymbol{\vartheta} \mid \boldsymbol{J}) = A(\boldsymbol{\vartheta}) . \tag{2.26}$$

Although we might expect that the above equality holds, it is in practice very difficult to derive the r.h.s. of the above equation. Given an explicit form of $A(\boldsymbol{\vartheta} \mid \boldsymbol{J})$ it is usually not

easy to *see* the dependence of $\boldsymbol{J}$ dropping out, once the $N \to \infty$ limit is taken. The usual approach to circumvent these problems is to simply re-define the meaningful macroscopic physical quantities as those that are obtained after averaging over the disorder. The average in (2.12) is adapted as follows

$$A(\boldsymbol{\vartheta}) = \int \mathrm{d}\boldsymbol{J} P(\boldsymbol{J}) A(\boldsymbol{\vartheta} \mid \boldsymbol{J}) \,. \tag{2.27}$$

We have established the free energy (or free entropy) as important thermodynamic potential that contains all macroscopic information about the system. And we have just introduced a useful trick that facilitates its computation.

Statistical Physics, Inference and Constraint Satisfaction Problems

The link between statistical physics and inference comes through the interpretation of the posterior probability as the GB measure of a disordered system. In inference problems the observed data is the quenched disorder (cf. Box 2.3). In CSPs the graph $\mathcal{G}$ represents the disorder, however, it does not appear explicitly in the Hamiltonian here. Nevertheless, one can set up a teacher-student scenario – known as **planting**. Let us consider the problem of circular coloring as an example. One first produces a set of N nodes with colors $x_i \in \{1, \dots, q\},\ i \in N$, and degree according to some prior distribution $P(\boldsymbol{x})$. Next one produces the graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$ according to $P(\mathcal{G} \mid \boldsymbol{x}) = \mathrm{e}^{-\beta \mathcal{H}(\boldsymbol{x})}$, cf. (1.57). For the problem of circular coloring this can be achieved by sorting the N nodes x_i into one of q bins, according to their color. Each node i carries d_i half-edges. Subsequently a random matching between the nodes in half-edges in bin b and bins $b \pm 1 \mod q$ is drawn (if it exists). However, CSPs are typically *not studied in the planted setting*. Instead a graph ensemble is considered, from which the graph is drawn and subsequently one studies the zero temperature limit $\lim_{\beta \to \infty} \mathrm{e}^{-\beta \mathcal{H}(\boldsymbol{x})}$. In that sense CSPs are somewhat similar to maximum likelihood problems.

Now that the link between inference problems, CSPs and disordered systems is established, we can discuss the *common* difficulties: the necessity to evaluate high dimensional (infinite dimensional) integral/sum. That is the partition function

$$Z(\boldsymbol{\vartheta} \mid \boldsymbol{J}) = \int \mathrm{d}\boldsymbol{x} \mathrm{e}^{-\beta \mathcal{H}(\boldsymbol{x} \mid \boldsymbol{J})} \,. \tag{2.28}$$

The explicit computation of the partition function can be avoided by some clever methods. One path to circumvent the problem is a **Monte Carlo** (MC) approach, presented in Box 2.1. A major issue with MC is that it can be impractically inefficient. In particular for glassy systems, that break the ergodicity. An alternative approach is the replica method that often renders an analytic computation of the logarithm of the above integral (the free entropy) possible. It is presented in the upcoming section. Anticipating the outcome of the replica method, it will typically yield an expression of the free entropy that is of the form

$$\phi(\boldsymbol{\vartheta}) = \operatorname*{extr}_{\boldsymbol{Q}} A_{\boldsymbol{\vartheta}}[\boldsymbol{Q}] \,. \tag{2.29}$$

The minimizers of this equation are deeply related to the Bayes risk in Sec. 1, cf. (1.73). We have already outlined that the Bayes optimal setting is special. In particular it will lead to a very simple re-parametrization of $\boldsymbol{Q}$ in terms of (effectively) a single free parameter, q, that is directly related to the fixed points of the state evolution equations from Sec. 1.3.6. Note also that, according to the above, a physical system minimized the free energy, as opposed to the Energy (Hamiltonian). This is reminiscent and in analogy to the difference between the minimum *mean* square error and the maximum a-posteriori estimator.

Box 2.1: Monte Carlo Methods

The aim is to compute physical observables, $A(\boldsymbol{\vartheta} \mid \boldsymbol{J})$, such as the average energy

$$E(\boldsymbol{\vartheta} \mid \boldsymbol{J}) = \int \mathrm{d}\boldsymbol{x}\, \mathcal{H}(\boldsymbol{x} \mid \boldsymbol{J}) \frac{\mathrm{e}^{-\beta \mathcal{H}(\boldsymbol{x}|\boldsymbol{J})}}{Z(\boldsymbol{\vartheta} \mid \boldsymbol{J})} \,. \tag{2.30}$$

The problem that we are facing is that $Z(\boldsymbol{\vartheta} \mid \boldsymbol{J})$ is unknown. Monte Carlo methods circumvent this problem by direct *simulation* of N particles, that can take state variables x_i, $i = 1, \ldots, N$, and that are subject to the interactions $\mathcal{H}(\boldsymbol{x} \mid \boldsymbol{J})$.

The royal road to such problems in high dimensions is the **Metropolis–Hastings algorithm**. It uses a Markov chain (cf. box 1.1), which asymptotically reaches the equilibrium distribution $\mathrm{e}^{-\beta \mathcal{H}(\boldsymbol{x}|\boldsymbol{J})}/Z(\boldsymbol{\vartheta} \mid \boldsymbol{J})$. Denoting the state variables at simulation time t by $\boldsymbol{x}(t)$, the probability that the system is found in state $\boldsymbol{s}_j$ at time t reads

$$P(\boldsymbol{x}(t) = \boldsymbol{s}_j) = \sum_k P(\boldsymbol{x}(t) = \boldsymbol{s}_j, \boldsymbol{x}(t-1) = \boldsymbol{s}_k) \tag{2.31}$$

$$= \sum_k P(\boldsymbol{x}(t) = \boldsymbol{s}_j \mid \boldsymbol{x}(t-1) = \boldsymbol{s}_k) P(\boldsymbol{x}(t-1) = \boldsymbol{s}_k) \,. \tag{2.32}$$

Denoting the *transition probabilities* by $T(\boldsymbol{s}_k \to \boldsymbol{s}_j) \coloneqq P(\boldsymbol{x}(t) = \boldsymbol{s}_j \mid \boldsymbol{x}(t-1) = \boldsymbol{s}_k)$ and $P_t(\boldsymbol{s}) \coloneqq P(\boldsymbol{x}(t) = \boldsymbol{s})$, this can be rewritten as

$$P_t(\boldsymbol{s}) = P_{t-1}(\boldsymbol{s}) + \sum_{\boldsymbol{s}'} \left[T(\boldsymbol{s}' \to \boldsymbol{s}) P_{t-1}(\boldsymbol{s}') - T(\boldsymbol{s} \to \boldsymbol{s}') P_{t-1}(\boldsymbol{s}) \right] . \tag{2.33}$$

For the probabilities $P_\infty(\boldsymbol{s})$ to reach the equilibrium distribution after long simulation times, $t \to \infty$, the following *detailed balance* condition must hold

$$\frac{T(\boldsymbol{s} \to \boldsymbol{s}')}{T(\boldsymbol{s}' \to \boldsymbol{s})} = \frac{P_\infty(\boldsymbol{s}')}{P_\infty(\boldsymbol{s})} = \mathrm{e}^{-\beta\left(\mathcal{H}(\boldsymbol{s}'|\boldsymbol{J}) - \mathcal{H}(\boldsymbol{s}|\boldsymbol{J})\right)} =: \mathrm{e}^{-\beta \Delta\mathcal{H}(\boldsymbol{s}', \boldsymbol{s}|\boldsymbol{J})} \,. \tag{2.34}$$

Therefore, the transition probabilities in the simulation should be set accordingly. The way this is achieved in the Metropolis-Hastings algorithm is by choosing the following acceptance rule

$$T(\boldsymbol{s} \to \boldsymbol{s}') = \begin{cases} 1 & \text{if } \Delta\mathcal{H}(\boldsymbol{s}', \boldsymbol{s} \mid \boldsymbol{J}) \leq 0 \\ \mathrm{e}^{-\beta \Delta\mathcal{H}(\boldsymbol{s}', \boldsymbol{s}|\boldsymbol{J})} & \text{if } \Delta\mathcal{H}(\boldsymbol{s}', \boldsymbol{s} \mid \boldsymbol{J}) > 0 \end{cases} . \tag{2.35}$$

The strategy is as follows: one creates a set of sampling points $\{\boldsymbol{s}_1, \boldsymbol{s}_2, \dots\}$ by successively proposing a new state $\boldsymbol{s}'$, according to some symmetric transition rule (that would cancel in the above equations for detailed balance) and then accepting/rejecting it according to the above rule. One can then simply compute the physical quantities by sampling, e.g.

$$E(\boldsymbol{\vartheta} \mid \boldsymbol{J}) = \frac{1}{P} \sum_{i=1}^{P} \mathcal{H}(\boldsymbol{s}_{k_i} \mid \boldsymbol{J}) \tag{2.36}$$

where $\boldsymbol{s}_{k_i}$ is the state of the system during the k_ith step of the simulation. Note that successive samples are typically correlated and one should thus choose the sample points in the above sum with care.
There are several other caveats. The system must be ergodic, i.e., it must be possible to reach any state from any state in finite time. Once the temperature is low, i.e. β large, and the energy landscape rugged, Monte Carlo Methods will get stuck in metastable states. In practice the computation of $\Delta\mathcal{H}(\boldsymbol{s}', \boldsymbol{s} \mid \boldsymbol{J})$ must be "feasible": if we change the state of one of the particles it is desired, that not all the other particles are affected by this change, otherwise each simulation step would take $\Theta(N)$ steps.

Replica Symmetric Free Energies and Inference

Let us introduce an important tool that will permit the analytic computation of the disorder average of the free energy

$$F(\tilde{\boldsymbol{\vartheta}}) = \int \mathrm{d}\boldsymbol{J} P(\boldsymbol{J}) \ln Z(\boldsymbol{\vartheta} \mid \boldsymbol{J}) \,. \tag{2.37}$$

Except from some rare cases, this integration is difficult to perform. The difficulty enters through the logarithm. A smart way to circumvent this problem is the **replica trick**. It amounts to combining the following useful identity that expresses the logarithm as a power function

$$\ln Z = \lim_{n \to 0} \frac{\partial}{\partial n} Z^n \tag{2.38}$$

We denote the limit $n \to 0$ the **replica limit**. We write

$$\lim_{N\to\infty} \mathbb{E}_{\boldsymbol{J}} \left[\frac{\ln Z(\boldsymbol{\vartheta} \mid \boldsymbol{J})}{N} \right] = \lim_{N\to\infty} \mathbb{E}_{\boldsymbol{J}} \left[\frac{1}{N} \lim_{n\to 0} \frac{\partial}{\partial n} Z(\boldsymbol{\vartheta} \mid \boldsymbol{J})^n \right] \tag{2.39}$$

Next the average of $Z(\boldsymbol{\vartheta} \mid \boldsymbol{J})^n$ w.r.t. $\boldsymbol{J}$ is computed as if n was an *integer power* to obtain an explicit formula that does no longer contain the disorder $\boldsymbol{J}$. Usually, it requires the *exchange of the two limits* in the above calculation to continue the calculation as well as some further **black magic**. After having done all, pretending that n is an integer, an analytic continuation of n is carried out in order to take the replica limit, $n \to 0$. What exactly happens is easiest understood by consideration of the concrete example in Box 2.3.

Although the replica trick seems just like a dirty trick to circumvent averaging the logarithm of the partition function, lots of physics is emerging from it. We have first introduced n

replicas of the system, at equal temperature $T = 1$.

$$Z(\boldsymbol{\vartheta} \mid \boldsymbol{J})^n = \sum_{\boldsymbol{x}^1,\dots,\boldsymbol{x}^n} \mathrm{e}^{\sum_{a=1}^n \mathcal{H}(\boldsymbol{x}^a|\boldsymbol{J})} \,. \tag{2.40}$$

These replicas are not directly coupled (the Hamiltonian is additive w.r.t. the replicas), however, they do all depend on the same disorder. An additional replica enters when the average over the disorder is carried out. This replica is special, it is *the one to bind them all in disorder*. Let's call it the *disorder replica*

$$\mathbb{E}_{\boldsymbol{J}}\left[Z(\boldsymbol{\vartheta} \mid \boldsymbol{J})^n\right] = \sum_{\boldsymbol{J}} \sum_{\boldsymbol{x}^0,\boldsymbol{x}^1,\dots,\boldsymbol{x}^n} \mathrm{e}^{\beta\mathcal{H}(\boldsymbol{x}^0|\boldsymbol{J})+\sum_{a=1}^n \mathcal{H}(\boldsymbol{x}^a|\boldsymbol{J})} \,. \tag{2.41}$$

In the crowdsourcing analysis we have introduced the disorder replica at the Nishimori line, i.e., at the same temperature. In the above equations we have kept the setting slightly more generic and introduced the disorder replica at a general temperature β. Otherwise the disorder replica is no different from the other replicas (in particular it has the same Hamiltonian). The temperature $\beta = 1$ is the physical equivalent to the Bayes optimal setting, while for $\beta \neq 1$ we are in the mismatched setting.

Once the disorder average is carried out in the above equation, one typically finds that it assumes the following (or a similar) shape for the replicated partition that couples the same sites i from different replicas

$$\mathcal{Z}_n(\boldsymbol{\vartheta}) \coloneqq \mathbb{E}_{\boldsymbol{J}}\left[Z(\boldsymbol{\vartheta} \mid \boldsymbol{J})^n\right] = \int \mathrm{d}\boldsymbol{Q} \mathrm{e}^{N A_{\boldsymbol{\vartheta}}[\boldsymbol{Q}]} \,. \tag{2.42}$$

The integration over the microstates is replaced by an integration over the macroscopic order parameters

$$\boldsymbol{Q} \coloneqq (q^{ab}) = \frac{1}{N}\sum_i x_i^a x_i^b. \tag{2.43}$$

The free entropy can be obtained from the above equation by a saddle point evaluation

$$\phi(\boldsymbol{\vartheta}) = \operatorname*{extr}_{\boldsymbol{Q}} A_{\boldsymbol{\vartheta}}[\boldsymbol{Q}] \,. \tag{2.44}$$

Box 2.2: Nishimori's Condition

Consider a teacher student inference problem and denote $\boldsymbol{x}^{(0)}$ the ground truth. Denoting the GB average with $\langle . \rangle_{\boldsymbol{x}|\boldsymbol{J}}$ and the average over the disorder with $[.]_{\boldsymbol{J}}$ the following equality is readily verified, if we are in the Bayes optimal setting with $P(\boldsymbol{x} \mid \boldsymbol{J}) = P_0(\boldsymbol{x} \mid \boldsymbol{J})$,

$$[\langle A(\boldsymbol{x}, \boldsymbol{x}^{(0)})\rangle_{\boldsymbol{x}|\boldsymbol{J}}]_{\boldsymbol{x}^{(0)},\boldsymbol{J}} = [\langle A(\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)})\rangle_{\boldsymbol{x}^{(1)},\boldsymbol{x}^{(2)}|\boldsymbol{J}}]_{\boldsymbol{J}} \,. \tag{2.45}$$

If the quantity behaves self-averaging, i.e. the measure concentrates with respect to the realizations of the disorder, $[\langle A(\boldsymbol{x}, \boldsymbol{x}^{(0)})\rangle_{\boldsymbol{x}|\boldsymbol{J}}]_{\boldsymbol{x}^{(0)},\boldsymbol{J}} = \langle A(\boldsymbol{x}, \boldsymbol{x}^{(0)})\rangle_{\boldsymbol{x}|\boldsymbol{J}}$ holds almost surely. Under the assumption of self-averaging we therefore find Nishimori's identity [Nis80]

$$\langle A(\boldsymbol{x}, \boldsymbol{x}^{(0)})\rangle_{\boldsymbol{x}|\boldsymbol{J}} = \langle A(\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)})\rangle_{\boldsymbol{x}^{(1)},\boldsymbol{x}^{(2)}|\boldsymbol{J}}\,. \tag{2.46}$$

The last identity states that two samples from the posterior distribution will carry the same average properties as one sample at fixed disorder. The immense importance of the above equality becomes vivid, once one considers the overlap between two configurations for $A(\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)})$. Then Nishimori tells us that comparing two different samples among each other in terms of their overlap, results in the same as comparing them with the ground truth (disorder). Identifying the two overlaps as

$$A(\boldsymbol{x}, \boldsymbol{x}^{(0)}) \coloneqq \frac{1}{N}\sum_{i=1}^{N} x_i^{(0)} x_i \tag{2.47}$$

$$A(\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}) \coloneqq \frac{1}{N}\sum_{i=1}^{N} x_i^{(1)} x_i^{(2)} \tag{2.48}$$

then the Nishimori conditions state that

$$A(\boldsymbol{x}, \boldsymbol{x}^{(0)}) = A(\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)})\,. \tag{2.49}$$

This is important because it states that the replica symmetric Ansatz in box 2.3 is correct in the Bayes optimal setting. It also allows a reduction of the set of order parameters to *one* instead of three. Another essential consequence of the Nishimori conditions is that the Bayes optimal MMSE estimate is not 'glassy'. In contrast, the MAP estimate may result in a solution space in the condensed phase.

The **replica symmetric** (RS) solution of the above equations corresponds to the particular parametrization of $\boldsymbol{Q}$, such that all the $n+1$ replicas are symmetric:

$$q^{ab} = Q\delta_{a\,b} + (1 - \delta_{a\,b})q\,. \tag{2.50}$$

For $\beta = 1$ it is physically intuitive that the $n+1$ replicas are all equal and it seems natural to assume that the order parameters adopt this symmetry. As is shown in Box 2.2, this is always true in the teacher-student scenario. Therefore the RS solution is the thermodynamically dominating one in the Bayes optimal case. However, there is a more general class of solutions that break this replica symmetry and will be discussed in the next section.

Now that we can compute the free energy, let us turn to a quick discussion of the interesting physics that can be extracted from it. We restrict the discussion to the simple Bayes optimal case and sketch the rough phenomenology, rather than giving an exhaustive treatment. Since we assume a Bayes optimal setting, $\boldsymbol{Q}$ is parametrized by the two parameters Q and q. Furthermore, we can set the selfoverlap $Q = 1$, which just amounts to a rescaling of the variables. The phase that the system lives in for a given (inverse) temperature, β, can be classified by the behavior of the minimizer of $A_\beta(q)$, say q^*. A **second order phase**

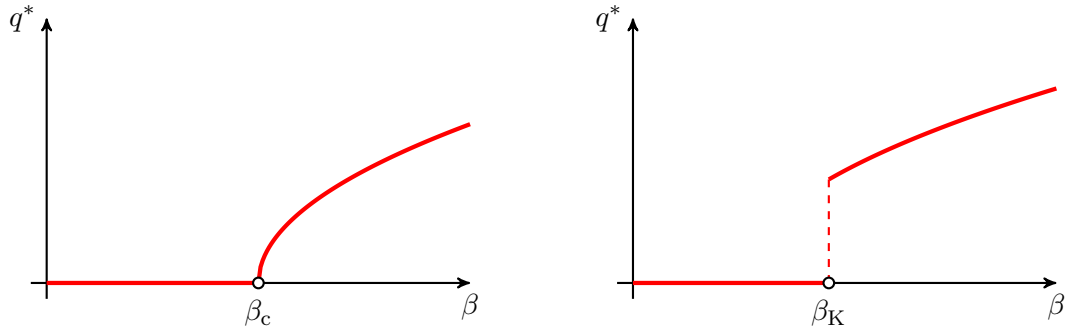

Figure 2.1: Exemplary, schematic, plot of the minimizer of the replica free energy, q^*, as a function of the inverse temperature β. Recalling the definition of q, we can identify the region with $q^* > 0$ as the ferromagnetic and $q^* = 0$ as the paramagnetic phase. Left: a second order phase transition, right: a first order phase transition. The variety of scenari that can appear in a first order transition are quite diverse and will not be discussed here in full detail.

transition is characterized by the property that the minimizer of $A_\beta(q)$ varies continuously. In contrast to this stands the **first order transition**, that is characterized by a discontinuous jump in the minimizer (more precisely by a discontinuity in the first derivative of the free energy, which manifests itself in this very jump). In Fig. 2.1 we are plotting the situation exemplarily.

Box 2.3: The Free Energy of Crowdsourcing (Bayes Optimal)

The posterior probability of the dense DS model for crowdsourcing is of the form (cf. sec. 1.3.4)

$$P\left(\boldsymbol{\theta}, \boldsymbol{v} \mid \boldsymbol{L}\right) = \frac{1}{Z(\boldsymbol{L})} \prod_{1 \leq j \leq M} P_v\left(v_j\right) \prod_{1 \leq i \leq N} P_\theta\left(\theta_i\right) \prod_{1 \leq j \leq M,\, 1 \leq i \leq N} P\left(L_{ij} \mid \theta_i, v_j\right) . \quad (2.51)$$

Consequently the Hamiltonian reads

$$\mathcal{H}\left(\boldsymbol{\theta}, \boldsymbol{v} \mid \boldsymbol{L}\right) = \sum_{j=1}^{M} \ln P_v\left(v_j\right) + \sum_{i=1}^{N} \ln P_\theta\left(\theta_i\right) + \sum_{i,j} \ln P\left(L_{ij} \mid \theta_i, v_j\right) \quad (2.52)$$

The associated free energy can now be computed with the replica trick. In the most general setting this was first done in [LKZ17]. We have

$$Z(\boldsymbol{L}) = \int \mathrm{d}\boldsymbol{\theta} \mathrm{d}\boldsymbol{v} P(\boldsymbol{\theta}) P(\boldsymbol{v}) P(\boldsymbol{L} \mid \boldsymbol{\theta}, \boldsymbol{v}) . \quad (2.53)$$

The computation becomes easer once we remove a constant from the free energy and compute $\tilde{Z}(\boldsymbol{L}) = Z(\boldsymbol{L})/P(\boldsymbol{L} \mid \boldsymbol{0})$ instead. The term $P(\boldsymbol{L} \mid \boldsymbol{0})$ only contain parameters that are related to the channel; it will not contain any contribute any "physics" in terms of the physical parameters that we will introduce further below. We adapt the notation from sec. 1.3.4 and write $P(\boldsymbol{L} \mid \boldsymbol{\theta}, \boldsymbol{v}) = \mathrm{e}^{g(\boldsymbol{L},\boldsymbol{w})}$ with $\boldsymbol{w} = \boldsymbol{\theta}\boldsymbol{v}^\intercal/\sqrt{N}$. We also introduce the shorthand $\mathrm{D}\boldsymbol{x} = \mathrm{d}\boldsymbol{x}P(\boldsymbol{x})$. For positive integer n the replicated partition function then reads

$$\tilde{Z}(\boldsymbol{L})^n = \int \prod_{a=1}^{n} \mathrm{D}\boldsymbol{\theta}^a \mathrm{D}\boldsymbol{v}^a \mathrm{e}^{\sum_{a=1}^n g(\boldsymbol{L},\boldsymbol{w}^a) - \sum_{a=1}^n g(\boldsymbol{L},\boldsymbol{0})} \,. \tag{2.54}$$

Next, this equation is averaged over the disorder. Since

$$P(\boldsymbol{L}) = \int \mathrm{d}\boldsymbol{\theta}\mathrm{d}\boldsymbol{v} P(\boldsymbol{L} \mid \boldsymbol{\theta}, \boldsymbol{v}) P(\boldsymbol{\theta}) P(\boldsymbol{v})$$

we simply obtain

$$\mathbb{E}_{\boldsymbol{L}}\left[\tilde{Z}(\boldsymbol{L})^n\right] = \int \mathrm{d}\boldsymbol{L} \prod_{a=0}^{n} \mathrm{D}\boldsymbol{\theta}^a \mathrm{D}\boldsymbol{v}^a \mathrm{e}^{\sum_{a=0}^n g(\boldsymbol{L},\boldsymbol{w}^a) - \sum_{a=1}^n g(\boldsymbol{L},\boldsymbol{0})} \,. \tag{2.55}$$

It is shown in appendix A.6 that an expansion around $\boldsymbol{w} = \boldsymbol{0}$ and integration over $\boldsymbol{L}$ yields

$$\mathbb{E}_{\boldsymbol{L}}\left[\tilde{Z}(\boldsymbol{L})^n\right] = \int \prod_{a=0}^{n} \mathrm{D}\boldsymbol{\theta}^a \mathrm{D}\boldsymbol{v}^a \mathrm{e}^{\frac{1}{2}(R+\frac{1}{\Delta})\sum_{a=0}^n \sum_{i,j} (\theta_i^a v_j^a)^2 + \frac{1}{2\Delta}\sum_{a\neq b}^n \theta_i^a v_j^a \theta_i^b v_j^b} \tag{2.56}$$

The following order parameters are emerging

$$q_\theta^{ab} = \frac{1}{N}\sum_{i=1}^{N} \theta_i^a \theta_i^b \quad \text{and} \quad q_v^{ab} = \frac{1}{M}\sum_{j=1}^{M} v_j^a v_j^b \,. \tag{2.57}$$

The r.h.s. of (2.56) can be expressed in terms of these order parameters, such that

$$\mathbb{E}_{\boldsymbol{L}}\left[\tilde{Z}(\boldsymbol{L})^n\right] = \int \mathrm{d}\boldsymbol{q}_\theta \mathrm{d}\hat{\boldsymbol{q}}_\theta \mathrm{d}\boldsymbol{q}_v \mathrm{d}\hat{\boldsymbol{q}}_v \, \mathrm{e}^{NS(\boldsymbol{q}_\theta, \hat{\boldsymbol{q}}_\theta, \boldsymbol{q}_v, \hat{\boldsymbol{q}}_v)} \,. \tag{2.58}$$

The action, S, is now an intensive quantity, and does not scale with N. In the thermodynamic limit this integral will be dominated by its saddle point, that is, the choice of $\{q_\theta^{ab}, q_v^{ab}, \hat{q}_\theta^{ab}, \hat{q}_v^{ab}\}$ that extremizes the action. The extremization allows us to eliminate the auxiliary variables However, it is difficult to find a solution for generic $n \times n$ matrices $\boldsymbol{q}_\theta, \hat{\boldsymbol{q}}_\theta, \boldsymbol{q}_v, \hat{\boldsymbol{q}}_v$. Therefore, it is common to reduce the set of solutions to be **replica symmetric**

$$q_x^{ab} = \begin{cases} Q_x & \text{if } a = b \\ q_x & \text{else} \,. \end{cases} \tag{2.59}$$

It can be shown, that in the Bayes optimal setting such a solution is always correct (cf. Nishimori condition, Box 2.2). As is shown in the appendix, the replica symmetric setting simplifies the above equations to

$$S_n(q_\theta, q_v) = -\frac{\alpha}{2}\frac{1}{\Delta}n(n+1)q_\theta q_v + \ln\left[\int \mathrm{D}W \mathrm{d}\theta P_\theta(\theta)\left[\int \mathrm{d}\theta P_\theta(\theta)\mathrm{e}^{-\frac{\alpha}{\Delta}q_v\theta^2+\sqrt{\frac{\alpha}{\Delta}q_v}W\theta}\right]^n\right] + \\ + \alpha\ln\left[\int \mathrm{D}W \mathrm{d}v P_v(v)\left[\int \mathrm{d}v P_v(v)\mathrm{e}^{-\frac{q_\theta}{\Delta}v^2+\sqrt{\frac{q_\theta}{\Delta}}Wv}\right]^n\right]. \tag{2.60}$$

Where $\mathrm{D}W = \mathrm{d}W\exp(-w^2/2)/\sqrt{2\pi}$ is the standard Gaußian measure. This action should be evaluated in the replica limit, where $n \to 0$, utilizing the identity (2.38). Finally, the replica symmetric free entropy reads

$$\phi(q_\theta, q_v) = -\frac{\alpha}{2\Delta}q_\theta q_v + \int \mathrm{D}W \mathrm{d}\theta P_\theta(\theta)\ln\left[\int \mathrm{d}\theta P_\theta(\theta)\mathrm{e}^{-\frac{\alpha}{\Delta}q_v\theta^2+\sqrt{\frac{\alpha}{\Delta}q_v}W\theta}\right] + \\ + \alpha\int \mathrm{D}W \mathrm{d}v P_v(v)\ln\left[\int \mathrm{d}v P_v(v)\mathrm{e}^{-\frac{1}{\Delta}q_\theta v^2+\sqrt{\frac{1}{\Delta}q_\theta}Wv}\right]. \tag{2.61}$$

2.2.1 Replica Symmetry Breaking

The first correction to the replica symmetric form of $\boldsymbol{Q}$ is the **one step replica symmetry breaking** (1RSB). In the 1RSB framework, the matrix $\boldsymbol{Q}$ takes a different form: the off-diagonal elements are grouped into different groups; the overlap of two replicas from within one group is different from the overlap among two replicas from different groups. In the Bayes optimal case we have

$$q^{ab} = \begin{cases} Q & \text{if } a = b \\ q_0 & \text{if } a, b \text{ in the same group} \\ q_1 & \text{if } a, b \text{ in different groups} \end{cases}. \tag{2.62}$$

This approach can be generalized and the groups can be divided into groups again etc., which leads to an *ultrametric structure* of $\boldsymbol{Q}$ [MPV87].

The breaking of the replica symmetry is deeply related to the concept of pure states [MPV87]. If the system is not on the Nishimori line, at low temperatures, and when $N \to \infty$, the *ergodicity of the system may break*. Instead of exploring the whole phase space, the system only explores a "small" part that is separated by extensive barriers from other parts of the phase space (cf. Fig. 2.2). The phase space breaks into **clusters** of solutions! Inside one such cluster ergodicity is restored and the corresponding microstates are said to be in a **pure state**. If one were to observe the system, one would not *see* the **Gibbs state**, i.e., states sampled from the GB measure, but instead one would observe a pure state, trapped in one of the clusters. The GB measure decomposes into contributions from each cluster (labeled by α), we write

$$P(\boldsymbol{x}) = \sum_\alpha w_\alpha \mu_\alpha(\boldsymbol{x}), \tag{2.63}$$

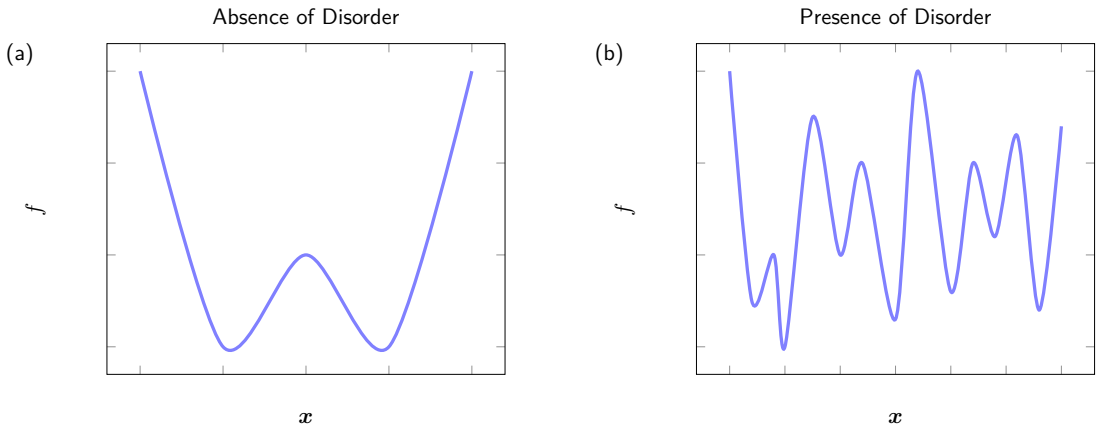

Figure 2.2: (a) Schematic sketch of a simple free energy landscape, e.g., the ferromagnetic Ising model below T_c. (b) Schematic sketch of a roughed free energy landscape with many local minima that are separated by extensive barriers.

with $\mu_\alpha(\boldsymbol{x})$ describing the measure within the pure state. Referring to Fig. 2.2 the α corresponds to different minima in the free energy landscape.

For a moment we go back to the physical picture in terms of phase space trajectories from the very beginning of this chapter. Let the system be prepared in some random initial state and subsequently let it evolve and relax into "equilibrium". If this experiment is repeated many times, and the free energy landscape contains many minima, separated by extensive barriers, the trajectories will each time get trapped into one of the different minima, α. A natural order parameter of interest is then the overlap between the magnetizations obtained in two different pure states, α and β,

$$q_{\alpha\beta} = \frac{1}{N}\sum_{i=1}^{N}\langle x_i\rangle_\alpha\langle x_i\rangle_\beta \tag{2.64}$$

Let us briefly outline how this is related to the replica calculation from the previous section. Let us assume a CSP on a graph $\mathcal{G}$, the disorder and consider the square magnetization of the Gibbs state

$$q_{\mathcal{G}}^{(1)} = \frac{1}{N}\sum_{i=1}^{N}\langle x_i\rangle^2\,, \tag{2.65}$$

Carrying out the averages w.r.t. (2.63) yields

$$q_{\mathcal{G}}^{(1)} = \sum_{\alpha,\beta} w_\alpha w_\beta \frac{1}{N}\sum_{i=1}^{N}\langle x_i\rangle_\alpha\langle x_i\rangle_\beta \tag{2.66}$$

$$= \sum_{\alpha,\beta} w_\alpha w_\beta q_{\alpha\beta}\,. \tag{2.67}$$

Let us denote the distribution of $q_{\alpha\beta}$ by

$$P_{\mathcal{G}}(q) = \sum_{\alpha,\beta} w_\alpha w_\beta \delta(q - q_{\alpha\beta})\,. \tag{2.68}$$

Where the subscript denotes that the distribution of $q_{\alpha\beta}$ does in principle depends on the disorder. The equation for $q^{(1)}_{\mathcal{G}}$ then reduce to the first moment of $P_{\mathcal{G}}(q)$. And after averaging over the disorder

$$q^{(1)} = \int \mathrm{d}q P(q)\, q\,, \tag{2.69}$$

with $P(q) = \mathbb{E}_{\mathcal{G}}[P_{\mathcal{G}}(q)]$. At the same time, the analogue replica calculation yields

$$q^{(1)} = \int \mathrm{d}\boldsymbol{Q} \mathrm{e}^{NA[\boldsymbol{Q}]}\, q^{ab} \quad \text{for a} \neq \text{b}\,. \tag{2.70}$$

This suggests, that the off-diagonal elements of the replica matrix $\boldsymbol{Q}$ are distributed according to $P(q)$. In the replica symmetric phase, the different pure states therefore all have the same overlap among each other. When the replica symmetry is broken the overlap between different pure states will be ordered into different groups. In the 1RSB framework the function $P(q)$ will have support on two different values. In the 2RSB framework it will have support on three values and so on. The **fullRSB** framework then corresponds to a $P(q)$ having finite continuous support.

Each cluster/pure state is characterized by the property that *long range correlations vanish at large distance*, i.e.,

$$\langle x_i x_j\rangle_\alpha - \langle x_i\rangle_\alpha \langle x_j\rangle_\alpha \xrightarrow{|i-j|\to\infty} 0\,. \tag{2.71}$$

This property must not hold within the GB measure:

$$\begin{aligned}\lim_{|i-j|\to\infty} \langle x_i x_j\rangle = \sum_\alpha w_\alpha \lim_{|i-j|\to\infty} \langle x_i x_j\rangle_\alpha &= \sum_\alpha w_\alpha \langle x_i\rangle_\alpha \langle x_j\rangle_\alpha \\ &\neq \langle x_i\rangle\langle x_j\rangle = \sum_\alpha w_\alpha \langle x_i\rangle_\alpha \sum_\alpha w_\alpha \langle x_j\rangle_\alpha\,.\end{aligned} \tag{2.72}$$

A sufficient condition for the failure of a RS solution is that long range correlations decay, $\lim_{|i-j|\to\infty}\langle x_i x_j\rangle \nrightarrow \langle x_i\rangle_\alpha\langle x_j\rangle_\alpha$.

In many diluted systems this condition does not hold and it is required to extend the replica symmetric version of the cavity method to compute the *correct* thermodynamic properties of the system. Formally, such an extension is necessary when the Gibbs measure loses its **extremality** [MM09; ZK07]. In such a situation, the extremality can be "restored" by decomposing the measure as in (2.63) with each cluster contributing with a weight that is exponential in its free energy [MPV87]

$$w_\alpha = \frac{\mathrm{e}^{-\beta N f_\alpha}}{\sum_\alpha \mathrm{e}^{-\beta N f_\alpha}}\,. \tag{2.73}$$

The partition function becomes

$$Z(\beta) = \sum_{\alpha} \mathrm{e}^{-\beta N f_\alpha} \tag{2.74}$$

and in analogy to the procedure in Sec. 2.1.2 we replace the summation over the pure states, by a summation over the characteristic free energy hypersurface

$$Z(\beta) \simeq \int \mathrm{d}f \Omega(f) \mathrm{e}^{-\beta N f} = \int_{f_{\min}}^{f_{\mathrm{th}}} \mathrm{e}^{N[\Sigma(f)-\beta f]} \simeq \mathrm{e}^{N[\Sigma(f^*)-\beta f^*]} =: \mathrm{e}^{-\beta N f_{\mathrm{tot}}(\beta)} \tag{2.75}$$

The **complexity**, $\Sigma(f)$ counts the log-number of pure states of free energy f. Depending on the temperature of the system, the following situations are typically encountered [Zam10]

$T > T_d$**:** The high temperature phase. It exists a single (typically paramagnetic) state with free energy f_{para} that is smaller than $f - T\Sigma(f)$ for all $f \in [f_{\min}, f_{\mathrm{th}}]$. The GB measure is therefore dominated by this state. However, other states exist, but they are not providing any thermodynamically relevant contribution to the measure.

$T_{\mathrm{K}} \leq T \leq T_{\mathrm{d}}$**:** The dynamic glass region. The temperature T_{d} denotes the **dynamic temperature** at which an **ergodicity breaking transition** takes place (also known as dynamical transition or clustering transition). The phase space decomposes into exponentially many clusters. $\Sigma(f_{\min}) > 0$. Yet, the total energy is given by the analytic continuation of the paramagnetic free energy $f_{\mathrm{tot}} = f^* - T\Sigma(f^*) = f_{\mathrm{para}}$. Thus, the dynamical transition is not a true thermodynamic transition. However, the dynamical behavior of the system (2.1) changes and no equilibrium distribution will be reached. This is the point were, e.g., Monte Carlo methods get trapped in one of the pure states that is separated by extensive barriers from other pure states. The total entropy is the sum of the internal entropy and the complexity, $s_{\mathrm{tot}}(T) = s_{\mathrm{int}} + \Sigma(T)$. Note that, although the high temperature free energy is still correct, there is no more single pure state dominating in this region. Instead, the fact that the free energy analytically continues is an average property that results from the contribution of the exponentially many pure states.

$T < T_{\mathrm{K}}$**:** The static glass region or **condensed phase**. For temperatures below the **Kautzmann temperature**, T_{K}, the dominating contribution comes from clusters with zero complexity and minimum free energy, $f_{\mathrm{tot}} = f^* - T\Sigma(f^*) = f_{\min}$. The total entropy is thus given by the internal contribution, $s_{\mathrm{tot}}(T) = s_{\mathrm{int}}(T)$. The GB measure condenses over the states with smallest free energy.

We have anticipated these results without any explicit calculations. This is the picture, established in the study of glassy systems [Zam10; CKP+17] (and it finds its analogies in constrained satisfaction problems, as will be outlined in Chap. 4). The above transitions can be extracted from the consideration of the complexity, $\Sigma(f)$. Therefore let us now address the question of how to compute $\Sigma(f)$. To that end one can introduce a reweighting

parameter m in the partition function that allows to control the average free energy within clusters, in analogy to (2.8),

$$\mathcal{Z}(m,\beta) = \sum_{\alpha} \mathrm{e}^{-\beta m N f_\alpha} \,. \tag{2.76}$$

Such that

$$\mathcal{Z}(m,\beta) = \int \mathrm{d}f \mathrm{e}^{N(\Sigma(f)-\beta m f)} =: \mathrm{e}^{\Phi(m,\beta)} \,. \tag{2.77}$$

Although the original system corresponds to $m = 1$, it is instructive to study $\Phi(m,\beta)$ for generic $0 \leq m \leq 1$: variation of m allows one to reweight solutions with different free energy. From this it is possible to extract $\Sigma(f)$, as is outlined in the following. From the saddle point evaluation of the above integral one obtains $\Phi(m,\beta)$ as the Legendre transformation of $\Sigma(f)$:

$$\Phi(m,\beta) = \max_f \left[-\beta m f + \Sigma(f)\right] = -\beta m f^* + \Sigma(f^*) \,. \tag{2.78}$$

The free variable is either f or βm. The free energy is then related to $\Phi(m,\beta)$ through

$$f^*(m,\beta) = -\frac{1}{\beta}\frac{\partial \Phi(m,\beta)}{\partial m} \,. \tag{2.79}$$

And the inversion of the Legendre transformation leads to

$$\Sigma(m,\beta) = \beta m f^*(m,\beta) + \Phi(m,\beta) \,. \tag{2.80}$$

Therefore, if $\Phi(m,\beta)$ can be computed it provides access to $f^*(m,\beta)$ and $\Sigma(m,\beta)$. From the knowledge of these two $\Sigma(f)$ can be constructed from a parametric plot of $f(m,\beta)$ and $\Sigma(m,\beta)$. In the next section we turn to the question of how to actually compute $\Phi(m,\beta)$ for pairwise MRFs.

2.2.2 Free Energies in Sparse Pairwise Markovian Random Fields

Let us give a quick note on *the difference between inference and constraint satisfaction problems*. In the natural setting of CSPs the disorder is not generated in the teacher-student scenario and $P(\boldsymbol{J}) \neq \sum_{\boldsymbol{x}} \exp\left(\beta \mathcal{H}(\boldsymbol{x} \mid \mathcal{G})\right)$. The exception are planted CSPs [KZ09; ZK11], that can be understood as inference problems in the teacher-student scenario at *zero temperature*. In the *sparse* graphical models that we are interested in when studying CSPs it is usually not possible to derive an explicit form of the free energy from the replica method. There are no apparent order parameters on which the equations can be closed. However, the physical phenomenology still applies. The equations that one obtains from the evaluation of the saddle point equations on sparse graphical models are the fixed point equations of the density evolution equations (1.70) [FLM+13; Mon98].

In order to compute the replicated free energy, $\Phi(m,\beta)$, for a CSP that is defined by the Hamiltonian $\mathcal{H}(\boldsymbol{x} \mid \mathcal{G})$ it is first necessary to compute the free energy within a pure state. This can be done in the **Bethe approximation**. For a pairwise MRF (for which $\mathcal{H}(\boldsymbol{x} \mid \mathcal{G}) =$

$\sum_{(ij)\in\mathcal{E}} \mathcal{H}(x_i, x_j \mid \mathcal{G})$) the Bethe approximation decomposes the partition function into its node and edge contributions, Z^i and Z^{ij} respectively,

$$\Phi_{\text{Bethe}} = -\beta F_{\text{Bethe}} = \sum_{i\in\mathcal{V}} \ln Z^i - \sum_{(ij)\in\mathcal{E}} \ln Z^{ij} \,. \tag{2.81}$$

Recalling the message passing equations for a pairwise MRF (1.69), the two contributions are

$$Z^i = \sum_{x_i} \prod_{k\in\partial i} \sum_{x_k} \psi_{ik}\left(x_i, x_k\right) m_{k\to i}\left(x_k\right) \tag{2.82}$$

$$Z^{ij} = \sum_{x_i, x_j} m_{i\to j}(x_i)\psi_{ij}(x_i, x_j) m_{j\to i}(x_j) \,. \tag{2.83}$$

The first term adds up the contributions from each node and the second is a correction term that is due to the fact that, when summing over Z^i, one counts the contributions from each edge twice. The Bethe approximation is exact on tree-like graphs [MM09].

The **1RSB cavity equations**, were developed in [MP01; MP03] and deal with weighted averages over pure states

$$\mathcal{P}_{i\to j}(m_{i\to j}) = \frac{1}{\mathcal{Z}^{i\to j}} \int \left[\prod_{k\in\partial i\setminus j} \mathrm{d}m_{k\to i} \mathcal{P}_{k\to i}(m_{k\to i})\right] \left(Z^{i\to j}\right)^m \delta\left(m_{i\to j} - \mathcal{F}\left(\{m_{k\to i}\}\right)\right) \,, \tag{2.84}$$

where $m_{j\to i}\left(x_j\right) = \mathcal{F}\left(\{m_{k\to j}\left(x_k\right)\}\right)$ are BP fixed point equations. In the 1RSB formalism, the $\mathcal{P}_{k\to i}(m_{k\to i})$ are distributions over the distributions $m_{k\to i}(x_k)$. The "functional fixed point", say $\{\mathcal{P}^*(m)\}$, provides the distribution (averaged over the disorder) of messages that live on the edges of a graph for a given graph ensemble. Note, that the distributions over the messages carry a subscript to stress that the depend on the nature of the particular edge. On random regular graphs, all the edges are similar (in the sense that the neighborhoods of each edge, up to a distance $\Theta(\log N)$, are identical).

In practice they can hardly be solved analytically and instead resorts to solve them by sampling (cf. Sec. 1.3.2). The **1RSB Bethe approximation** of the replicated free energy (2.77) decomposes as before into node and edge contributions, $\Phi_{\text{Bethe}} = \sum_i \Delta\Phi^i - \sum_{ij} \Delta\Phi^{ij}$, with

$$\Delta\Phi^{i,ij} = -\frac{1}{\beta m} \int \prod_k \mathrm{d}m_k \mathcal{P}_k^*(m_k) \left(Z^{i,ij}\right)^m \,. \tag{2.85}$$

Where Z^i and Z^{ij} are computed according to (2.83).

We are now able to compute the log partition function for the circular coloring, as well as the crowdsourcing problem. The other thermodynamical potentials, such as energy and entropy that are of physical interest, can be accessed through a Legendre transformation. We will demonstrate this in Chap. 4 for the problem of circular coloring.

— 3 —

Elements from Random Matrix Theory

This chapter establishes some tools from random matrix theory that will be useful to analyze disordered systems with spins of extensive size. We focus on the large N limit (the planar limit) and the Coulomb gas approach. We show how the free energy of the rotationally invariant ensembles can be extracted by a saddle point evaluation. The saddle point equations lead to a singular integral integration. The solution to these equations can be obtained in two complementary ways that are outlined and explained.

At first sight random matrices are nothing but random variables that carry two indices. For such random matrices one is then interested in the statistics of the eigenvalues and eigenvectors. The first prominent contribution is due to Wishart [Wis28], who computed the joined distribution of eigenvalues (and vectors) for matrices of the form $\boldsymbol{X}\boldsymbol{X}^\intercal$, with $x_{\mu k} \sim \mathcal{N}(0,1)$. In physics, random matrices were introduced in nuclear physics to study the properties of large nuclei that are otherwise too complex to handle. The idea was to treat such systems not in terms of the properties of a single nucleus, but instead study the *typical* properties of an ensemble of nuclei [Wig51], very much in the spirit of statistical physics. Instead of investigating the energy levels and eigenfunctions of the Hamilton operator of one specific nucleus,

$$\hat{\boldsymbol{H}}\boldsymbol{\Psi}_i = E_i \boldsymbol{\Psi}_i$$

one studies the ensemble of all nucleus that belong to a specific ensemble that is determined by the underlying symmetry class, such as time-reversal invariance. The Hamilton operator $\hat{\boldsymbol{H}}$ then becomes a random object that is drawn from this ensemble and so do the energy levels (the eigenvalues of $\hat{\boldsymbol{H}}$). This idea was first introduced by Wigner and later elaborated on by many others [Dys62]. It now finds applications in a vast amount of different fields. We refer the reader to the great lecture notes [EKR15] for a summary.

A **random matrix model** for an $N \times N$ matrix $\boldsymbol{M} \in \Omega_\beta$ is characterized through the measure $P_\beta(\boldsymbol{M})$ that acts on the ensemble Ω_β. Here we adapted the RMT notation and indicated with β the ensemble of random matrices; $\beta \in \{1, 2, 4\}$ respectively for the orthogonal, unitary and symplectic ensembles (we will see below why it makes sense to introduce

this notation). We are interested in the class of matrix ensembles that are induced by the potential, $V(\boldsymbol{M})$, such that

$$P_\beta(\boldsymbol{M}) = \mathrm{e}^{-\beta \mathrm{Tr} V(\boldsymbol{M})}\,, \tag{3.1}$$

with the partition function

$$Z_\beta = \int \mathrm{d}\boldsymbol{M} \mathrm{e}^{-\beta \mathrm{Tr} V(\boldsymbol{M})}\,. \tag{3.2}$$

The three most prominent examples of this class are the Gaußian (or Wigner) ensembles with $V(\boldsymbol{M}) = \boldsymbol{M}\boldsymbol{M}^\dagger/\sigma$ (where the $\dagger$ indicates the transpose, Hermitian conjugation and the quaternion self-dual respectively for $\beta = 1,\, 2,\, 4$):

- The *Gaußian Orthogonal Ensemble (GOE)* ($\beta = 1$). The matrices $\boldsymbol{M}$ are real and symmetric. This ensemble applies to systems with even spin, invariant under time reversal and with a rotational symmetry.
- The *Gaußian Unitary Ensemble (GUE)* ($\beta = 2$), for which the matrices $\boldsymbol{M}$ are complex and Hermitian. These systems are not time reversal invariant.
- The *Gaußian Symplectic Ensemble (GSE)* ($\beta = 4$), for which the matrices $\boldsymbol{M}$ are again Hermitian, but they carry quaternions as entries. This ensemble applies to systems with odd-spin, invariant under time reversal, but with no rotational symmetry.

All these ensembles share the property that there are no correlations among the unique elements of $\boldsymbol{M}$ and that they can be decomposed as

$$\boldsymbol{M} = \boldsymbol{\Omega}\boldsymbol{L}\boldsymbol{\Omega}^\dagger\,, \tag{3.3}$$

where $\boldsymbol{\Omega}$ is an orthogonal, unitary or symplectic matrix respectively and $\boldsymbol{L}$ is the matrix of eigenvalues of $\boldsymbol{M}$.

This decomposition imposes a measure on the eigenvalues and eigenvectors of $\boldsymbol{M}$

$$P_\beta(\boldsymbol{M})\,\mathrm{d}\boldsymbol{M} = P_\beta(\boldsymbol{\Omega}\boldsymbol{L}\boldsymbol{\Omega}^\dagger)\,\mathrm{D}\boldsymbol{\Omega} \prod_{i<j} |l_i - l_j|^\beta\,(\mathrm{d}\boldsymbol{L})\,. \tag{3.4}$$

Here $D\boldsymbol{\Omega}$ indicates the flat (Haar) measure that is invariant with respect to a translation with a group element of $G(N) \in \{\mathrm{O}(N), \mathrm{U}(N), \mathrm{Sp}(N)\}$ [EKR15]. The term $\prod_{i<j} |l_i - l_j|$ is known as **Vandermonde determinant**. We see that the only difference among the ensembles enters through the power of the Vandermonde determinant. If the distribution P_β is invariant with respect to the group action, i.e., $P_\beta(\boldsymbol{\Omega}\boldsymbol{L}\boldsymbol{\Omega}^\dagger) = P_\beta(\boldsymbol{L})$, then the partition function reduced to an integration over the eigenvalues

$$Z_\beta = \int \mathrm{d}\boldsymbol{L} \prod_{i<j} |l_i - l_j|^\beta\ \mathrm{e}^{-\beta \mathrm{Tr} V(\boldsymbol{L})}\,. \tag{3.5}$$

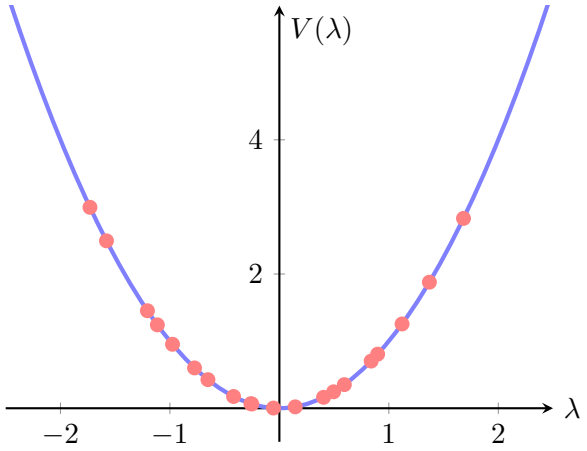

Figure 3.1: An exemplary realization of the eigenvalue gas for the GOE resulting from a diagonalization of a 20×20 Wigner matrix.

3.1 The Planar Limit

In what follows we will only be concerned with real symmetric matrices (i.e. we set $\beta = 1$) and the limit where the size of the matrix is infinitely large: the thermodynamic (or planar) limit $N \to \infty$. In random matrix theory this limit is known as planar limit, because the leading order Feynman diagrams that provide the dominant contributions to the connected Green's functions are planar diagrams. In this limit, it is natural to expect that (3.5) may be accessible by a saddle point evaluation.

First of all, we bring the above in a more amenable form for a physicist: we write

$$Z = \int \prod_{i=1}^{N} \mathrm{d}l_i \, \mathrm{e}^{-N^2 E[\{l_i\}]}, \quad \text{with} \quad E[\{l_i\}] = \frac{1}{N} \sum_{i=1}^{N} V(l_i) - \frac{1}{N^2} \frac{1}{2} \sum_{i \neq j} \ln |l_i - l_j| \,. \tag{3.6}$$

Where we have rescaled the potential in such a way that the energy of the system is an extensive quantity and the eigenvalues are $O(1)$ and assumed that $V(\boldsymbol{L}) = \sum_{i=1}^{N} V(l_i)$. Consider for example the GOE, in which this rescaling corresponds to elements that have variance $O(1/N)$. Brought into this form it is possible to extract some physics from the above equations. Our original system, in which the (unique) matrix elements were all independently experiencing the potential V, is apparently equivalent to a system of N interacting particles, each of them moving in the same potential, but this time they are all interacting through a logarithmic repulsion (cf. Fig. 3.1). This is often referred to as **Coulomb gas analogy** because the repulsive term is the same as for a two dimensional Coulomb gas of charged particles. However, note that we actually have a one dimensional gas and the analogy is slightly misleading.

For large N and $E[\{l_i\}] = O(1)$ the Gas will settle in the most probable configuration that minimizes the energy and the integral in (3.6) can be evaluated by the saddle point

method

$$\frac{\partial E}{\partial l_i} \stackrel{!}{=} 0 \quad \Rightarrow \quad V'(l_i) = \frac{1}{N} \sum_{\substack{j=1 \\ j \neq i}}^{N} \frac{1}{l_i - l_j} \quad \forall i = \{1, \ldots, N\} \,. \tag{3.7}$$

The saddle point condition leads to a equilibrium between the forces acting from the external potential (l.h.s.) and the interactions (r.h.s.). In order to solve these equations it is useful to transition to the continuous limit of the above equations. In the thermodynamic limit the eigenvalues will eventually approach a continuous limit and it is natural to introduce the **eigenvalue (or spectral) density**

$$\rho_{\boldsymbol{M}}(\lambda) \coloneqq \lim_{N \to \infty} \mathbb{E}_{\boldsymbol{M}} \left[\frac{1}{N} \sum_{i=1}^{N} \delta \left(\lambda - l_i \right) \right] . \tag{3.8}$$

Typically we drop the subscript of $\rho_{\boldsymbol{M}}$, if it is clear from the context.

Before it is possible to move on to solve the saddle point equations, it is necessary to introduce some useful transforms from random matrix theory. For an arbitrary $z \in \mathbb{C}$ the normalized trace of the resolvent is defined as

$$\mathfrak{g}_{\boldsymbol{M}}^{N}(z) \coloneqq \frac{1}{N} \mathrm{Tr} (z \mathbf{1} - \boldsymbol{M})^{-1} = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{z - l_i} \,. \tag{3.9}$$

We simply refer to it as **resolvent** and may drop the subscript that indicates the matrix of interest, if there is no danger of confusion. We see that the resolvent is a nice function everywhere in $\mathbb{C}$, except at the points where z coincides with one of the eigenvalues of $\boldsymbol{M}$, where it has poles. Indeed, it contains all the information about the eigenvalues of $\boldsymbol{M}$ in its poles. For physicists the resolvent is also known as two-point Green's function and the $N \to \infty$ limit is commonly known as **Stieltjes or Cauchy transform**:

$$\mathfrak{g}_{\boldsymbol{M}}(z) \coloneqq \lim_{N \to \infty} \mathfrak{g}_{\boldsymbol{M}}^{N}(z) = \int \mathrm{d}\lambda \, \frac{\rho_{\boldsymbol{M}}(\lambda)}{z - \lambda} \,. \tag{3.10}$$

The resolvent enjoys the following properties [BIPZ78]:

1) It is analytic in the whole complex plane except from the cuts along $\mathrm{supp}(\rho_{\boldsymbol{M}})$.

2) It behaves as z^{-1} for $|z| \to \infty$.

3) It is real for $z \in \mathbb{R} \backslash \mathrm{supp}(\rho_{\boldsymbol{M}})$.

4) It has a jump when approaching the support of $\rho_{\boldsymbol{M}}$ from above/below:

$$\lim_{\eta \to 0^+} \mathfrak{g}_{\boldsymbol{M}}(x \pm \mathrm{i}\, \eta) = \mathfrak{h}_{\boldsymbol{M}}(x) \mp \mathrm{i}\, \pi \rho_{\boldsymbol{M}}(x) \quad \text{for} \quad x \in \mathrm{supp}(\rho_{\boldsymbol{M}}) \,, \tag{3.11}$$

with $\mathfrak{h}_{\boldsymbol{M}}(x)$ denoting the Hilbert transform of $\rho_{\boldsymbol{M}}(x)$.

The first point is a consequence of the definition. The second follows from the **normalization of the eigenvalue density** and expanding the integral for large z. Point three is also obvious. Finally, the last, and most important point is not immediately obvious: it tells us that we can extract the eigenvalue density of $\boldsymbol{M}$ from the imaginary part of the $g_{\boldsymbol{M}}(z)$ on approaching the support of $\rho_{\boldsymbol{M}}(\lambda)$, which is known as **Sokhotski formula**:

$$\mathfrak{g}(x - \mathrm{i}\,0) - \mathfrak{g}(x + \mathrm{i}\,0) = 2\pi\,\mathrm{i}\,\rho(x)\,. \tag{3.12}$$

Having defined the resolvent and its analytic properties, we are ready to outline the approach of [BIPZ78] to solve the saddle point equations.

3.2 Solving the Saddle Point Equations

Let us now outline how to solve the saddle point equations. In this section we outline two different methods that lead to a solution of the continuous saddle point equations

$$V'(\lambda) = \fint \mathrm{d}\lambda' \frac{\rho(\lambda')}{\lambda - \lambda'}\,. \tag{3.13}$$

The integral is a principal value integral, because the $O(1)$ contribution of the r.h.s. in (3.7) is contributed by the few eigenvalues that are close by. We are aware of two different approach to the inversion of the above **singular integral equation**. In an approach due to Brézin, Itzykson, Parisi and Zuber [BIPZ78] (**BIPZ**) one solves the problem by passing into the complex plane and exploiting general properties of analytic functions. Another approach is to invert the equations directly (once one has taken the continuous limit) by means of **Tricomi's theorem** [Tri85; Gak90; MR08].

3.2.1 The BIPZ Approach

In the BIPZ approach one starts from (3.7) and introduces the resolvent by multiplication of both sides of (3.7) with $N^{-1}(z - l_i)^{-1}$ and summing over all i:

$$\frac{1}{N}\sum_{i=1}^{N} \frac{V'(l_i)}{z - l_i} = \frac{1}{N^2}\sum_{i \neq j} \frac{1}{(z - l_i)(l_i - l_j)}\,. \tag{3.14}$$

These equations can be rewritten, which leads to a quadratic equation for $\mathfrak{g}(z)$:

$$\mathfrak{g}^2(z) - 2V'(z)\mathfrak{g}(z) + \frac{2}{N}\sum_{i=1}^{N} \frac{V'(z) - V'(l_i)}{z - l_i} = 0\,. \tag{3.15}$$

In principle one can now solve the above equations for $\mathfrak{g}(z)$. Let us introduce the term

$$P(z) \coloneqq \frac{1}{N}\sum_{i=1}^{N} \frac{V'(z) - V'(l_i)}{z - l_i} \tag{3.16}$$

and solve the equations (3.15) for $\mathfrak{g}(z)$, which leads to

$$\mathfrak{g}(z) = V'(z) \pm \sqrt{V'(z)^2 - 2P(z)}\,. \tag{3.17}$$

However, a difficulty remains: it appears necessary to evaluate the term $P(z)$ in the limit $N \to \infty$ in order to get a useful explicit expression. However, one can circumvent this issue by exploiting some generic properties of the functions involved: let us assume that $V'(z)$ can be expressed as a Laurent polynomial of degree d and order l, and that $\rho(x)$ support on $[a, b]$, then

$$\mathfrak{g}(z) = V'(z) - Q(z)\sqrt{(z-a)(z-b)}\,. \tag{3.18}$$

with $Q(z)$ a Laurent polynomial of degree $d-1$ and order l:

$$Q(z) = \sum_{k=-l}^{d} c_k z^k\,. \tag{3.19}$$

It seems like it is necessary to know the *explicit* form of the term (3.16) in order to extract the eigenvalue density from the (3.18) via the Sokhotski formula. However, this is not the case, because the coefficients of $Q(z)$, as well as a and b, can be extracted from the the requirements that $\mathfrak{g}(z)$ must obey, namely that $\mathfrak{g}(z) \to z^{-1}$ for large z.[1]

3.2.2 The Boundary Value Approach

In the boundary value approach, one starts from the continuous version of the saddle point equations (3.13). This singular integral equation (3.13) for $\rho(\lambda)$ can be inverted directly [MR08; Gak90; Tri85]. The solution basically depends on whether or not the solution is bounded at its ends. The most general inversion formula reads [Tri85]

$$\rho(\lambda) = \frac{1}{\sqrt{(\lambda-a)(b-\lambda)}}\left[\frac{1}{\pi^2}\fint_a^b \mathrm{d}\lambda'\,\sqrt{(\lambda'-a)(b-\lambda')}\,\frac{V'(\lambda')}{\lambda-\lambda'} + a_0\right]\,. \tag{3.20}$$

Depending on the result of the integral and the constant a_0, the formula allows for a solution (i) unbounded at both ends, (ii) bounded at one of the two ends and (iii) bounded at both ends. The free constants must be obtained from the constraints that are imposed on $\rho(\lambda)$ (symmetries, normalization,...). If it is known that the solution is bounded at the end a and unbounded at the end b the solution reduces to [Gak90]

$$\rho(\lambda) = \frac{1}{\pi^2}\sqrt{\frac{\lambda-a}{b-\lambda}}\fint_a^b \sqrt{\frac{b-\lambda'}{\lambda'-a}}\frac{V'(\lambda')}{\lambda-\lambda'}\,. \tag{3.21}$$

[1] This condition is not always sufficient to obtain all free parameters. In that case one must employ additional constraints that $\mathfrak{g}(z)$ must obey.

Similarly, a solution bounded at the two ends, a and b, reads

$$\rho(\lambda) = \frac{1}{\pi^2}\sqrt{(\lambda-a)(b-\lambda)}\oint_a^b \mathrm{d}\lambda' \frac{V'(\lambda')}{\sqrt{(\lambda'-a)(b-\lambda')}(\lambda-\lambda')} \tag{3.22}$$

and must obey the additional condition

$$\int_a^b \mathrm{d}\lambda' \frac{V'(\lambda')}{\sqrt{(\lambda'-a)(b-\lambda')}} \stackrel{!}{=} 0\,. \tag{3.23}$$

As an example of the methods and ideas introduced, we conclude the chapter with the Gaußian orthogonal ensemble as an example in Box 3.1.

Box 3.1: The Macroscopic Limit of the Gaußian Ensemble The GOE is the ensemble of *real symmetric* $N \times N$ matrices $\boldsymbol{M}$ with $\mathbb{E}m_{ij} = 0$ and $\mathbb{E}m_{ij}^2 = (1+\delta_{ij})\sigma^2/N$, such that

$$P(\boldsymbol{M}) \propto \mathrm{e}^{-\frac{N}{2\sigma^2}\mathrm{Tr}\boldsymbol{M}\boldsymbol{M}^\intercal}\,.$$

This is evidently orthogonally invariant as $\mathrm{Tr}\boldsymbol{M}\boldsymbol{M}^\intercal = \mathrm{Tr}\boldsymbol{O}\boldsymbol{L}\boldsymbol{O}^\intercal\boldsymbol{O}\boldsymbol{L}\boldsymbol{O}^\intercal = \mathrm{Tr}\boldsymbol{L}^2$. In the notion of (3.6), this leads to a potential of the form

$$V(x) = \frac{x^2}{4\sigma^2}\,.$$

The saddle point equations, in their continuous version, become

$$\frac{\lambda}{2\sigma^2} = \fint \mathrm{d}\lambda' \frac{\rho(\lambda')}{\lambda-\lambda'}\,.$$

From the shape of the potential (cf. 3.1) it is quite clear that we should expect the eigenvalue density to have a single support.

We first solve the saddle point equations in the BIPZ approach. The derivative of the potential is a regular polynom of degree one and thus we assume $Q(z) = c_0$. It remains to fix the free parameters, c_0, a, b, by the asymptotic constraint that $\mathfrak{g}(z) \to z^{-1}$. An asymptotic expansion of (3.18), with $Q(z) = c_0$, leads to

$$\mathfrak{g}(z) \stackrel{|z|\to\infty}{=} \left(\frac{1}{2\sigma^2} - c_0\right) z + \frac{c_0}{2}(a+b) + \frac{c_0}{8}(a-b)^2\frac{1}{z} + O(\frac{1}{z^2})\,.$$

We require $\mathfrak{g}(z) \to z^{-1}$, which implies

$$c_0 = \frac{1}{2\sigma^2}\,, \quad \frac{c_0}{2}(a+b) = 0\,, \quad \frac{c_0}{8}(a-b)^2 = 1\,.$$

The first condition fixes c_0, the second implies that $a = -b$ (which we must expect from the symmetry of the potential) and the last equation yields $b = -a = 2\sigma$. In summary, we obtain

$$\mathfrak{g}(z) = \frac{z - \sqrt{z^2 - 4\sigma^2}}{2\sigma^2}\,.$$

From which we can extract the eigenvalue density by Sokhotski's formula

$$\rho(\lambda) = \frac{\sqrt{4\sigma^2 - \lambda^2}}{2\pi\sigma^2}\,.$$

Next, we approach the same problem by direct inversion. From the shape of the potential we clearly require the solution to be bounded at both ends and employ the equations (3.22) and (3.23) and solve the integral

$$\fint_a^b \mathrm{d}\lambda' \frac{V'(\lambda')}{\sqrt{(\lambda' - a)(b - \lambda')}(\lambda - \lambda')} = \frac{1}{2\sigma^2} \fint_a^b \mathrm{d}\lambda' \frac{\lambda'}{\sqrt{(\lambda' - a)(b - \lambda')}} \frac{1}{(\lambda - \lambda')}\,.$$

This can be done by drawing the contour around the cut $[a, b]$ and deforming the contour such that it approaches the cut. This integral can be evaluated through its residuum at infinity. This contour integral is equal to two times the above integral because the square root switches sign when the contour goes from one side of the cut to the other. We obtain

$$\fint_a^b \mathrm{d}\lambda' \frac{\lambda'}{\sqrt{(\lambda' - a)(b - \lambda')}} \frac{1}{(\lambda - \lambda')} = \pi$$

and thus

$$\rho(\lambda) = \frac{1}{2\pi\sigma^2} \sqrt{(\lambda - a)(b - \lambda)}\,.$$

Next, it is necessary to invoke (3.23), which leads to what we already know from the intrinsic symmetry:

$$\frac{\pi}{2}(a + b) = 0 \qquad \Rightarrow \qquad a = -b\,.$$

Finally, we are left with one last free parameter, a, that we can fix by imposing the normalization of the eigenvalue density:

$$\int_a^{-a} \mathrm{d}\lambda \rho(\lambda) \stackrel{!}{=} 1 \qquad \Rightarrow \qquad \frac{a^2}{4\sigma^2} = 1\,.$$

We are pleased to find back

$$\rho(\lambda) = \frac{\sqrt{4\sigma^2 - \lambda^2}}{2\pi\sigma^2}\,.$$

Part II

Applications and Contributions

— 4 —

Circular Coloring

The CCOL problem was introduced in the course of chapter 1 as an example of a CSP. Our work on this problem inscribes in a line of other works that apply the tools and ideas from spin glass theory to study the typical (average case) properties of CSPs.

My main contribution comprises in the 1RSB analysis of the problem. It revealed that CCOL exhibits several interesting features, not common to other discrete CSPs. Further more it establishes a conjecture by Nešetřil about the 5-CCOLorability of sub-cubic graphs, from a statistical physics point of view.

This chapter is organized as follows. After a short introduction of the problem, the RS and 1RSB cavity method are outlined. The different zero temperature limits, necessary to study the satisfiability of the problem, are subsequently introduced and results for random regular graphs are presented. In the context of 5-CCOLorability of sub-cubic graphs, it will be necessary to consider the entropic zero temperature limit in Sec. 4.5.4. Subsequently the 1RSB stability of this particular case is considered. The results of the stability analysis will lead us to some algorithmic considerations of the problem, that will be presented in the very end of the chapter.

Although the problem was already introduced, as an example in Sec. 1.3.1, we briefly recall the definition and draw the line to the canonical coloring problem. Whereas in the canonical coloring two nodes of a graph that are connected by an edge are required to have different colors, in circular coloring the colors are ordered into a circle, and two adjacent nodes are required to have two adjacent colors (compare Fig. 1.7 and Fig. 1.7).

To define *circular coloring* (CCOL) consider a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ where each node $i \in \mathcal{V} = \{1, \dots, N\}$ can attain the discrete values (colors) $x_i \in \{1, 2, \dots, q\}$ and two nodes (i, j) are connected if $(i, j) \in \mathcal{E}$. We denote j as a neighbor of i, $j \in \partial i$, if $(i, j) \in \mathcal{E}$. Then the graph is q-circular colorable if and only if there exists an assignment of q colors to the nodes such that, if a node $i \in \mathcal{V}$ is of color x_i, then all nodes $j \in \partial i$ are of color $x_j \in \{x_i - 1, x_i + 1\}$ modulo q.

Note that for $q = 2$ and $q = 3$ there is no difference between the canonical and circular coloring. Hence all the works on canonical coloring apply. For this reason only $q > 3$ (and mostly odd q, see below) will be considered. Note further that CCOL with an even number of colors is closely related to 2-coloring. The minimum number of edges that must be violated

in $2z$-circular coloring of a given graph is the same for every $z = 1, 2, 3, \ldots$. On the one hand, consider an assignment of 2 colors that minimizes the number of violated edges for 2-coloring, then the same number of violated edges is achievable also for $2z$-circular coloring for any integer z (we simply use only 2 of the $2z$ colors). On the other hand, if a graph is $2z$-circular colorable with a given number of violated edges for some integer z then it is also 2-colorable with less or equal number of violated edges because every odd color can be replaced by the first one and every even color by the second one. Therefore the ground state of q-circular coloring is the same for every q even.

4.1 Context in Mathematics

CCOL belongs to a larger class of problems that generalizes the canonical graph coloring problem and is often explained using graph homomorphisms, objects of more general interest in mathematics. Given graphs $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ and $\mathcal{G}' = (\mathcal{V}', \mathcal{E}')$, a *homomorphism* is any mapping $f : \mathcal{V} \to \mathcal{V}'$ which satisfies $(ij) \in \mathcal{E} \Rightarrow (f(i)f(j)) \in \mathcal{E}'$. The existence of a coloring of a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with q colors is hence equivalent to an existence of a homomorphism of that graph onto a complete graph on q nodes. CCOL is equivalent to a homomorphism onto a cycle of q nodes. Clearly, all the other possibilities for the graph $\mathcal{G}'$ are of interest in mathematics.

Conjecture 4.1 (**Pentagon Problem** [Neš13]). *There exists an integer l with the following property: If $\mathcal{G}$ is a subcubic graph (i.e. every vertex has degree ≤ 3) with girth $\geq l$ then $\mathcal{G} \to \mathcal{C}_5$.*

Nešetřil's Pentagon Conjecture states that all graphs of maximum degree three without short cycles are 5-circular colorable. This conjecture is inspired by the aim to generalize classical results known for coloring, for instance that every graph with maximum degree 3 is 3-colorable unless it contains a complete graph of 4 nodes [Bro41].

A series of mathematical works established that there are sub-cubic graphs with large girth that are not q-circular colorable for $q \geq 7$ [KNS01; WW01; Hat05]. These proofs show that a random 3-regular graph is not q-circular colorable for $q \geq 7$ using variants of the first moment method. Therefore, $q = 5$ is the remaining open case for colorability of sub-cubic graphs. The use of random graphs in the proofs [WW01; Hat05] is an important motivation to study the behavior of 5-circular coloring on the same class of graphs. The existing powerful, non-rigorous techniques, from statistical physics, can easily be adapted to study CCOLs on random graphs. CCOL of particular graphs or deterministic classes of graphs is well studied, for a recent review we refer to [Zhu06].[1] Concerning CCOL on random graphs not much is known, apart from [WW01; Hat05].

[1] In the paper by Zhu, the following definition of CCOL is employed. For two integers $1 \leq r \leq q$, a (q, r)-coloring of a graph $\mathcal{G}$ is a coloring $\boldsymbol{x}$ of the vertices of $\mathcal{G}$ with colors $\{0, 1, 2, \ldots, q-1\}$ such that $(ij) \in \mathcal{E}(\mathcal{G}) \Rightarrow r \leq |x_i - x_j| \leq q - r$. The circular chromatic number is defined as $\chi_c(\mathcal{G}) = \inf\{k = d : \text{there is a } (r, q) - \text{coloring of } \mathcal{G}\}$. This is related to our definition by setting $q - 2d = 1$, such that $d \leq |x_i - x_j| \leq d + 1$. For odd q this coincides with our CCOL definition.

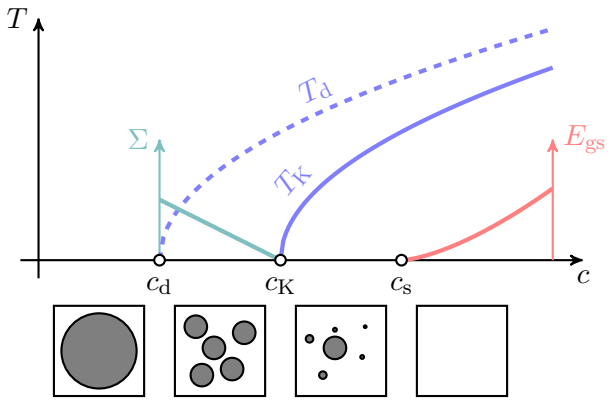

Figure 4.1: Schematic phase diagram in the temperature, T, vs. average degree, c, plane. The dashed and solid blue lines mark the dynamical and Kautzmann temperatures respectively, as discussed in 2.2.1. Their intersection point at zero temperature with the c-axis marks the respective zero transition points c_{d} and c_{K}. The SAT/UNSAT (or COL/UNCOL) transition is denoted c_{s}. It is the point where no more zero energy groundstate configurations exist. The array of pictures depicts a figurative way of the three transition points in terms of the clusters of solutions at zero temperature.

From the known mathematical results the most remarkable one is perhaps the one of [Hat05] that established the non-colorability for 7-circular coloring of random 3-regular graphs. A simple calculation of the expected number of proper circular colorings where every color is present on the same number of nodes shows that the vanilla 1st moment method is not sufficient to show that 3-regular graphs are with high probability not 7-circular colorable. This suggest that something non-trivial is happening for 7-circular coloring of random 3-regular graphs. And the upper bound established by [Hat05] is non-trivial along the lines of the upper bound of [Coj13] for coloring.

4.2 Context in Physics

The contributions presented in this chapter inscribes in a line of work where statistical physics methods, developed in the field of spin glasses [MPV87; MM09], are applied to study random instances of constraint satisfaction problems. The most well known works in this direction are those of random graph coloring and random K-satisfiability [MPZ02; MPWZ02; KMR+07; ZK07]. The success of these previous investigations and the subsequent spark of rigorous results [CZ12; BCH+16; CP16] motivated the extension of these investigations to CCOL.

In the course of chapter 2 we have outlined the typical phase space picture, found in spin glass systems: when the temperature is decreased, the phase space decomposes into clusters of pure states at the dynamical temperature T_{d}. Cooling the system further down, the GB measure condenses on the largest of these clusters with smallest free energy at the Kautzmann temperature, T_{K}. On the other hand, investigating the satisfiability of CSPs corresponds to

the minimum of (1.53), which, in turn, is the zero temperature limit of the GB measure, $\lim_{\beta\to\infty} \mathrm{e}^{-\beta\mathcal{H}(\boldsymbol{x})}$. Finding satisfiable CCOL assignments on a random instance becomes harder when (a) the average degree increases and (b) the number of colors increases. At zero temperature, the average degree of the underlying graph, c, plays the role of an inverse temperature. Bringing this together with the above spin glass picture of the phase space, one would expect a $T-c$ phase diagram as in Fig. 4.1, which is indeed the rough behavior found in K-SAT and canonical coloring [MM09; ZK07].[2] In light of the Pentagon Problem, we will be mostly interested in the satisfiability threshold. The point where the ground state energy becomes larger than zero.

4.3 Cavity Method for Circular Coloring

The cavity method, as developed in [MP01; MP03] and presented in Sec. 1.3.2 of this thesis, can be applied to the CCOL Hamiltonian

$$\mathcal{H}(\boldsymbol{x}) = \sum_{(i,j)\in\mathcal{E}} H(x_i, x_j) = \sum_{(i,j)\in\mathcal{E}} (1 - \delta_{x_i,x_j-1} - \delta_{x_i,x_j+1}) \,, \tag{4.1}$$

where the algebra in the indices is modulo q.

4.3.1 Replica symmetric solution

It was shown in Box 1.3 that for $q > 2$ the BP equations for CCOL, at generic (inverse) temperature, $\beta = 1/T$, read (after a slight lightening of the notation)

$$m_{x_i}^{i\to j} = \mathcal{F}(\{\boldsymbol{m}^{k\to i}\}) \coloneqq \frac{1}{Z^{i\to j}} \prod_{k\in\partial i\setminus j} [\mathrm{e}^{-\beta} + (1-\mathrm{e}^{-\beta})(m_{x_i+1}^{k\to i} + m_{x_i-1}^{k\to i})] \,. \tag{4.2}$$

The term multiplied by $1-e^{-\beta}$ assigns more probability when neighbors have consecutive colors. Here and in the following, bold messages $\boldsymbol{m}^{i\to j}$ indicate the whole distribution over the discrete variables $1 \le x_i \le q$. The Bethe approximation, cf. (2.81), for the free entropy for CCOL becomes

$$\begin{aligned} -\beta F &= \sum_i \log Z^i - \sum_{(ij)\in\mathcal{E}} \log Z^{ij} \\ &= \sum_i \log \left[\sum_{s=1}^{q} \prod_{k\in\partial i} [\mathrm{e}^{-\beta} + (1-\mathrm{e}^{-\beta})(m_{s+1}^{k\to i} + m_{s-1}^{k\to i})] \right] \\ &\quad - \sum_{(ij)\in\mathcal{E}} \log \left[\mathrm{e}^{-\beta} + (1-\mathrm{e}^{-\beta}) \sum_s m_s^{i\to j} (m_{s+1}^{j\to i} + m_{s-1}^{j\to i}) \right] . \end{aligned} \tag{4.3}$$

[2] The actual picture is a little more refined, but that has been extensively studied before and won't be further discussed here.

In (4.3) the free entropy is an extensive variables, denoted by capital letters. Subsequently we will be working with the corresponding densities and denote them by lower case letters, e.g., $\epsilon = E/N$.

To examine the validity of the RS solution one can investigate when it loses its stability towards small perturbations in the messages. The high temperature phase is characterized by the decay of long range correlation, i.e., the clustering property (2.71). The two-point correlation functions can be related to the leading eigenvalue of the Jacobian,

$$J^{\tau\sigma} = \frac{\partial m_{\tau}^{1\to 0}}{\partial m_{\sigma}^{2\to 1}}\Big|^{\mathsf{RS}}, \tag{4.4}$$

by virtue of the fluctuation-dissipation theorem (cf. Appendix C of [Zde09] for a summary). Here $m_{\tau}^{1\to 0}$ represents an arbitrary outgoing messages that depends on the incoming message $m_{\sigma}^{2\to 1}$. When the leading eigenvalue of (4.4) becomes larger than one, the two point correlation functions will diverge. This establishes a lower bound on the dynamic/clustering temperature (respectively an upper bound on c, cf. Fig. 4.1). Note, however, that a 1RSB solutions with positive complexity may appear before (at higher temperature) the two point correlations diverge. The clustering property is only a sufficient condition for the existence of a clustered solution space, not a necessary one. The temperature/(average) degree where this transition sets in is referred to as *linear*, *local* or *Kesten-Stigum* instability. This is equivalent to the non-convergence of BP on a single graph. For $q \geq 3$, the entries of the Jacobian in terms of the BP messages read

$$J^{\tau\sigma} = (1-\mathrm{e}^{-\beta})\, m_{\tau}^{1\to 0} \left\{ \frac{(\delta_{\sigma,\tau-1} + \delta_{\sigma,\tau+1})}{e^{-\beta} + (1-\mathrm{e}^{-\beta})(m_{\tau-1}^{2\to 1} + m_{\tau+1}^{2\to 1})} \right. \tag{4.5}$$

$$\left. - \frac{m_{\sigma+1}^{1\to 0}}{\mathrm{e}^{-\beta} + (1-\mathrm{e}^{-\beta})(m_{\sigma}^{2\to 1} + m_{\sigma+2}^{2\to 1})} - \frac{m_{\sigma-1}^{1\to 0}}{\mathrm{e}^{-\beta} + (1-\mathrm{e}^{-\beta})(m_{\sigma-2}^{2\to 1} + m_{\sigma}^{2\to 1})} \right\}; \tag{4.6}$$

The explicit eigenvalues will be given in the results section.

4.3.2 One Step Replica Symmetry Breaking

Anticipating results from the later sections of this chapter, the CCOL problem exhibits replica symmetry breaking in a broad range of the phase space. In order to take the clustered structure of the phase space into account, the 1RSB cavity equations (2.84) must be considered. In the 1RSB framework, each state corresponds to a different fixed point of the BP equations. The 1RSB equations then deal with weighted averages over all such fixed points and repeated here for convenience

$$\mathcal{P}_{i\to j}(\boldsymbol{m}^{i\to j}) = \frac{1}{\mathcal{Z}^{i\to j}} \int \left[\prod_{k\in\partial i\setminus j} \mathrm{d}\boldsymbol{m}^{k\to i} \mathcal{P}^{k\to i}(\boldsymbol{m}^{k\to i}) \right] \delta\Big(\boldsymbol{m}^{i\to j} - \mathcal{F}(\{\boldsymbol{m}^{k\to i}\})\Big) \left(Z^{i\to j}\right)^{m}, \tag{4.7}$$

Likewise the thermodynamic quantities must be re-evaluated. For that purpose the *replicated free energy* $\Phi(\beta, m)$ was introduced in Sec. 2.2.2.

$$\Phi(\beta, m) = -\beta m f + \Sigma(f)\,. \tag{4.8}$$

The replicated free entropy Φ is the Legendre transformation of Σ. The Bethe approximation is adopted to compute the 1RSB free energy as outlined in Sec. 2.2.2 and recalled here

$$\begin{aligned}
\Phi_{\text{Bethe}} &= \frac{1}{N}\left(\sum_{i\in\mathcal{V}} \Delta\Phi^i - \sum_{(ij)\in\mathcal{E}} \Delta\Phi^{ij}\right)\,, \\
\Delta\Phi^{i,ij} &= -\frac{1}{\beta m}\ln \int \prod_{a=1}^{l} \mathrm{d}\boldsymbol{m}^a \mathcal{P}_a(\boldsymbol{m}^a)\,\delta\Big(\boldsymbol{m} - \mathcal{F}(\{\boldsymbol{m}^a\})\Big)\,\Big(Z^{i,ij}\Big)^m\,.
\end{aligned} \tag{4.9}$$

Where l is the number of neighbors on which Z^i and Z^{ij} depend. The 1RSB equations can be solved efficiently using population dynamics as introduced in [MP01] and detailed in appendix A.3: $\mathcal{P}_a(\boldsymbol{m}^a)$ is approximated by a population of messages that is updated according to (4.7). After convergence, the population $\{\boldsymbol{m}^*\}$ is an approximation of the true distribution over fixed points $\mathcal{P}(\boldsymbol{m})$. Each element in $\{\boldsymbol{m}^*\}$ corresponds to a different cluster.

Following the same line of arguments as for the replica symmetric solution, the stability of the 1RSB solution towards further steps of replica symmetry breaking can be tested. Small perturbations can occur in the 1RSB solution either in $\mathcal{P}(\boldsymbol{m})$ *or* in $\boldsymbol{m}$. In this work only the second case is investigated and will already suffice to exclude the exactness of the 1RSB framework. A description of this stability analysis was for instance given in [Zde09]. Numerically, one first waits τ iterations until the 1RSB equations converged, for a given value of the re-weighting parameter m, to $\{\boldsymbol{m}^*\}$. Then the population is duplicated and a small noise is introduced $\{\boldsymbol{m}^* + \delta\boldsymbol{m}^*\}$. Subsequently t further iterations are performed. If both populations do not converge towards the same fixed point in the limit of many iterations, the 1RSB solution is said to be unstable. For a given m the convergence can be tracked by means of the evolution of the noise $Q_m(t) = \sum_{\{\boldsymbol{m}^*\}} |\boldsymbol{m}^*_{\tau+t}|$. If $\lim_{t\to\infty} Q_m(t) \not\to 0$ the solution is unstable towards further steps of RSB. The value of m at which the instability sets in will be denoted by m^{Δ}.

4.4 Cavity Method at Zero Temperature

In CCOL the ground state quantities are particularly interesting, as they yield information about the satisfiability. Therefore the zero temperature limit of the 1RSB equations should be considered. There are complementary ways to take the zero temperature limit that we shall now briefly recall.

If one is merely interested in proper assignments for which the total energy is zero one considers the so-called *entropic zero-temperature limit* [MPR05; Zde09]. If the per-particle

energy is zero, $\epsilon = 0$, then the internal free entropy reads

$$-\beta f = s \tag{4.10}$$

and can be computed via (4.3). With this, and after defining $\Phi_{\mathrm{s}}(m) \coloneqq \lim_{\beta\to\infty} \Phi(\beta, m)$, the replicated free entropy, (4.8), might be rewritten as

$$\Phi_s(m) = ms + \Sigma(s)\,. \tag{4.11}$$

The last equation states that the total entropy is composed of the internal entropy within a cluster and the log number of clusters with this specific internal entropy. The replicated entropy $\Phi_s(m)$ and hence the number of clusters of size s of proper assignments, $\Sigma(s)$, can be computed from (4.9). The remaining free parameter m can be varied in order to access $\Sigma(s)$ over the whole range of s.

On the other hand, in the *energetic zero temperature limit* we take $\beta \to \infty$ while $y \coloneqq m\beta$ remains finite. In this case the internal free energy equals the internal energy, i.e.,

$$f = \epsilon \tag{4.12}$$

and from (4.8) one obtains

$$\Phi(y) = -y\,\epsilon + \Sigma(\epsilon)\,. \tag{4.13}$$

It's called *energetic zero temperature limit* because it allows to compute the actual ground state energy of the physical system. In this limit the structure of the messages alter considerably as developed in the work of [MP03] and explained below.

4.4.1 Warning Propagation

To realize the energetic zero temperature 1RSB analysis, one starts with the *warning propagation*, which is a zero temperature limit of the belief propagation (4.2). The cavity fields $h_{x_i}^{i\to j}$ are introduced as $m_{x_i}^{i\to j} \coloneqq \exp(-\beta h_{x_i}^{i\to j})$ and the sum in the BP equations (4.2) is replaced by taking the maximal marginal (the only survivor in this limit), i.e., $\sum_{x_k} \to \max_{x_k}$. Adapting the generic BP equations (1.69) accordingly yields

$$m_{x_i}^{i\to j} = \exp(-\beta h_{x_i}^{i\to j}) \cong \exp(-\beta \sum_{k\in\partial i\setminus j} \min_{x_k}[\mathcal{H}(x_i, x_k) + h_{x_k}^{k\to i}])$$

and therefore, after taking the logarithm, the equations for the fields read

$$h_{x_i}^{i\to j} = \sum_{k\in\partial i\setminus j} \min_{x_k} \left[\mathcal{H}(x_i, x_k) + h_{x_k}^{k\to i}\right] - \min_{x_k} h_{x_k}^{k\to i}\,. \tag{4.14}$$

The term on the right hand side can be interpreted as a *warning* (0 or 1) from node k, incoming to node i; let's denote it by $u_{x_k}^{k\to i}$. The sum over the warnings $\sum_{k\in\partial i\setminus j} u_{x_k}^{k\to i}$ yields

the *cavity field* acting on node i, assuming that node j is absent and the above equation can be re-expressed as

$$h_\tau^{i\to j} = \sum_{k\in\partial i\setminus j} u_\tau^{k\to i}\,, \tag{4.15}$$

where the τth component of $u(h_1,\dots,h_\tau)$ reads

$$\begin{aligned}\tilde{u}_\tau(h_1,\dots,h_q) = \min\big(&h_1+1,\dots,h_{\tau-2}+1,h_{\tau-1},h_\tau+1,h_{\tau+1},h_{\tau+2}+1,\dots,h_q+1\big)\\ &-\omega(h_1,\dots,h_q)\,.\end{aligned} \tag{4.16}$$

with

$$\omega(h_1,\dots,h_q) = \min\left(h_1,h_2,\dots,h_{q-1},h_q\right)\,, \tag{4.17}$$

assuring that $u_{x_k}^{i\to j}\in\{0,1\}$. Identifying a contradiction, or *energy shift*, along the directed edge $i\to j$ as an assignment of x_i such that $\mathcal{H}(x_i,x_j)=1$. Then a contradiction is related to $u_{x_i}^{i\to j}=1$. Contrarily, if $u_{x_i}^{i\to j}=0$ no contradiction is caused by the assignment. The τth component of the field $h^{i\to j}$ is therefore related to the number of contradictions along all incoming edges caused by assigning color τ to node i if the edge j is absent. Accordingly, an *energy shift* can be assigned to a single node: this is the number of contradictions from *all* incoming edges. Subsequently q-component vectors will be denoted by bold symbols $\boldsymbol{h}\coloneqq(h_1,h_2,\dots,h_q)$ to simplify the notation.

4.4.2 Survey Propagation (SP)

The reasoning of [BMP+03] was generic enough to be applied to CCOL as well (up to the point where the list of relevant warnings is explicated). This is because all the information on the Hamiltonian is absorbed in $\omega(\boldsymbol{h})$. Therefore the 1RSB equation (4.7) reads (up to a normalization, recalling $y=m\beta$)

$$\mathcal{P}_{i\to j}(\boldsymbol{h}) \cong \int \prod_{k\in\partial i\setminus j} \mathrm{d}\boldsymbol{u}_k \mathcal{Q}_{k\to i}(\boldsymbol{u}_k)\,\delta\left(\boldsymbol{h}-\sum_{k\in\partial i\setminus j}\boldsymbol{u}_k\right)\,\exp\left[-y\,\omega\left(\sum_{k\in\partial i\setminus j}\boldsymbol{u}_k\right)\right]\,, \tag{4.18}$$

$$\mathcal{Q}_{i\to j}(\boldsymbol{u}) = \int \mathrm{d}\boldsymbol{h}\mathcal{P}_{i\to j}(\boldsymbol{h})\,\delta\left(\boldsymbol{u}-\tilde{\boldsymbol{u}}(\boldsymbol{h})\right)\,. \tag{4.19}$$

The above equations are referred to as the SP-y solution, or as the energetic zero temperature limit of the cavity solution.

At this point another limit can be taken, namely $y\to\infty$, which prohibits any kind of contradiction as only those local fields $\boldsymbol{h}=\sum_k\boldsymbol{u}_k$ contribute in (4.18) that contain at least one zero component, i.e., no contradiction. This is due to the reweighting term $\exp\left[-y\omega(\boldsymbol{h})\right]$ that, in this limit, only contributes when the energy is zero. The resulting equations are known as *survey propagation*. In this limit we can characterize the distributions over the warnings $\mathcal{Q}(\boldsymbol{u})$ by the set of parameters $\{\eta\}$ that are associated with the different possible warnings $\{\boldsymbol{u}\}$. For each distinct warning $\boldsymbol{u}^{(i)}$ we have $\eta^{(i)}=\mathcal{Q}(\boldsymbol{u}^{(i)})$. In other words, solving the

equations (4.18) and (4.19) amounts to finding the equivalence classes $[\boldsymbol{h}]_i$ that map onto the warning $\boldsymbol{u}^{(i)}$, i.e., $[\boldsymbol{h}]_i = \{\boldsymbol{h} \in \{\boldsymbol{h}\} \,|\, \tilde{\boldsymbol{u}}(\boldsymbol{h}) = \boldsymbol{u}^{(i)}\}$. This can be written in a recursion for the parameters $\eta^{(i)}$, where (i) indicates the ith equivalence class and i runs over all integers from 0 to $|\{\boldsymbol{u}\}| - 1$ and $\pi(\{i\})$ denotes the set of all possible permutations over the set of the indicators of the equivalence classes.

$$\eta_{i\to j}^{(\alpha)} \cong \sum_{\pi(\{\alpha_k\})} \mathbb{I}\left(\sum_{k\in\partial i\setminus j} \boldsymbol{u}_k^{(\alpha_k)} \in [\boldsymbol{h}]_\alpha\right) \prod_{k\in\partial i\setminus j} \eta_{k\to i}^{(\alpha_k)} \,. \tag{4.20}$$

In contrast to previously studied cases of the survey propagation equations, e.g. [BMZ05; MPWZ02], the equations (4.20) have a rather complicated structure due to the diverse cases that are to be distinguished. Unlike regular coloring, CCOL is much stronger constrained: if one fixes the color of one node to s, all nearest neighbors must take either of the two colors $s-1$ or $s+1$, second-nearest neighbors are restricted to s, $s-2$ and $s+2$ and so forth. In fact up to $k(q)$-nearest neighbors are restricted. Exemplary: for $q = 5$ we have $k = 3$. Consequently the closure of (4.20) is much more involved and we could obtain no explicit formula for the generic case. This difference is further illustrated by studying the structure of the warnings $\boldsymbol{u}$ and fields $\boldsymbol{h}$ for the 5-circular coloring on 3-regular graphs in section 4.4.3. But even if the closed form is unknown, the equations can either be generated numerically in an exhaustive approach, or they can simply be solved by computing (4.18) and (4.19) in the limit $y \to \infty$ with the population dynamics algorithm (cf. Appendix A.3).

When aiming to close the equations on $\boldsymbol{u}$, it is necessary to find the mapping of incoming fields to the outgoing warnings and identify all fields that cause the same response in (4.16). Without loss of generality, we can identify $\min(\boldsymbol{h})$ with $h_{\arg\min(\boldsymbol{h})} = 0$ and all other components $i \neq \arg\min(\boldsymbol{h})$ with $h_i = 1$ because they cause the same respond in (4.16). After this identification we must, in general, still distinguish $|\{\boldsymbol{h}\}| = 2^q$ fields.

Let's assume that the rotational and reflectional symmetry of q-circular coloring is not broken below a certain point $q_{\mathrm{sym}}(d)$ – where d denotes the degree. For $q < q_{\mathrm{sym}}(d)$ we can take advantage of this symmetry and $|\{\boldsymbol{h}\}| < 2^q$ cases must be distinguished for the incoming fields. When we count the different fields and warnings it is understood that, e.g., $(h_1, 0, 0, 0, 0)$ and $(0, 0, h_3, 0, 0)$, with h_1 not necessarily equal to h_3, is counted only once. The distinguished cases $|\{\boldsymbol{h}\}|$ are then obtained by counting in how many different ways q beads can be circularly connected when each bead is of either of two colors and rotations and reflections of this circular string are regarded as equivalent. These objects are known as *bracelets* in combinatorics and they are counted by

$$|\{\boldsymbol{h}\}| = \frac{1}{2q}\sum_{t|q} \phi(t)\, 2^{\frac{q}{t}} + \begin{cases} 2^{\frac{q-1}{2}} & q \text{ odd} \\ 3\cdot 2^{\frac{q}{2}-2} & q \text{ even} \end{cases} .$$

where the sum goes over all divisors of q and $\phi(t)$ is Euler's Totient function that counts how many integers are smaller or equal to t and share no common positive divisors with t, except 1. Since (4.16) introduces a non-injective surjection on $\{\boldsymbol{h}\}$, it is more involved to

q	$\lvert\{\boldsymbol{h}\}\rvert$	$\lvert\{\boldsymbol{u}\}\rvert$
3	3	2
4	5	2
5	7	4
6	12	6
7	17	8
8	29	13
9	45	17
10	77	26
11	125	36
12	223	56
13	379	82
14	686	128

Table 4.1: The cardinality of the set of all possible fields and warnings in the paramagnetic phase where vectors that are equivalent under the group actions of rotation and reflection are only counted once.

count the cardinality of $\{\boldsymbol{u}\}$ under consideration of the symmetry. However, it is simple to provide the first few values of the sequence from numerical evaluation, done in table 4.1.

Note that this is intrinsically different from e.g. regular coloring where, after taking into account the permutation symmetry, only two generic cases must be distinguished: the case in which $\boldsymbol{h}$ possesses a unique minimum, causing $\boldsymbol{u} = \hat{\mathbf{e}}_\tau$, where $\hat{\mathbf{e}}_\tau$ is a unit vector in direction τ, and the case in which the minimum is degenerate, causing $\boldsymbol{u} = \mathbf{0}$. Such a simplification is not possible for CCOL and, per se, $\lvert\{\boldsymbol{u}\}\rvert$ generic cases must be distinguished (following this section we will give a concrete example to illustrate the difference). Within the framework of CCOL it is therefore sensible to slightly modify the notation of what is known as *frozen* variables in other constraint satisfaction problems (CSPs), such as generic coloring or K-SAT. In previously considered CSPs the variables could either be trivial or point into one of the q possible directions $\hat{\mathbf{e}}_\tau$. As we have just seen in this section, this is no longer the case in CCOL. Instead of frozen variables the term *confined variables* is therefore introduced in order to emphasis this different nature.

4.4.3 Example of Survey Propagation for Five-Circular Coloring

In this section we consider the survey propagation equations of the example of 5-circular coloring of 3-regular graphs, in this case we can obtain explicit survey propagation equations. For this example we shall confirm in Sec. 4.5.3 that the symmetry under rotation and

$\boldsymbol{h}$		$\boldsymbol{u}$
$(0,0,0,0,0)$ $(h_1,0,0,0,0)$ $(h_1,h_2,0,0,0)$	$\rightarrow$	$\boldsymbol{u}^{(0)}=(0,0,0,0,0)$
$(h_1,h_2,h_3,0,0)$ $(h_1,0,h_3,0,0)$	$\rightarrow$	$\boldsymbol{u}^{(1)}=(0,1,0,0,0)$
$(h_1,h_2,h_3,h_4,0)$	$\rightarrow$	$\boldsymbol{u}^{(2)}=(0,1,1,0,1)$
$(h_1,h_2,0,h_4,0)$	$\rightarrow$	$\boldsymbol{u}^{(3)}=(0,0,1,0,1)$

Table 4.2: All possible non-contradictory incoming fields $\boldsymbol{h}$, i.e. all those fields that contain at least one zero component, and their corresponding outgoing warnings $\boldsymbol{u}$ (up to relevant symmetry).

reflection is not broken.

First we derive the set of warnings. Assume that no neighbor imposes any constraint on node i, i.e., $\boldsymbol{h}=(0,0,0,0,0)$, then the outgoing warning yields $\boldsymbol{u}^{(0)}=(0,0,0,0,0)$. Similarly, if the neighbors of i constrain one color only, say $x_i=1$ then $\boldsymbol{h}=(h_1,0,0,0,0)$ (with $h_1>0$) yielding $\boldsymbol{u}^{(0)}$ again. The last case that yields this "zero" warning would be $\boldsymbol{h}=(h_1,h_2,0,0,0)$. Proceeding in the same fashion, one might check all possible fields $\{\boldsymbol{h}\}$. The according mapping from $\boldsymbol{h}$ to $\boldsymbol{u}$ can, again up to the intrinsic symmetry, be found in table 4.2.

Conversely the warnings map to the fields in a more tedious manner. Assuming $d=3$ and labeling the two incoming edges as "left" and "right". One might obtain $\boldsymbol{u}^{(0)}$ by combining the two incoming warnings $\boldsymbol{u}^{(0)}$ and $\boldsymbol{u}^{(0)}$. Note that there is only one such combination because of the symmetry under exchanging left $\rightleftharpoons$ right. Another way to obtain $\boldsymbol{u}^{(0)}$ is by combining $\boldsymbol{u}^{(0)}$ from the left with $\boldsymbol{u}^{(1)}$ from the right and vice versa (no symmetry w.r.t. left $\rightleftharpoons$ right). The last possible combination would then be $\boldsymbol{u}^{(1)}$, incoming from one of the two edges and $\boldsymbol{u}^{(1)}$ from the remaining edge, such that either $(h_1,0,0,0,0)$ or $(h_1,h_2,0,0,0)$ is obtained, i.e., there are 3 such combinations. For the last three cases we must also take into account the five possible rotations. We summarize this in the following table.

left	right	left $\rightleftharpoons$ right	
$\boldsymbol{u}^{(0)} = (0,0,0,0,0)$	$\boldsymbol{u}^{(0)} = (0,0,0,0,0)$	yes	$\eta^{(0)}\eta^{(0)}$
$\boldsymbol{u}^{(0)} = (0,0,0,0,0)$	$\boldsymbol{u}^{(1)} = (1,0,0,0,0)$ + rot	no	$5 \cdot 2\eta^{(0)}\eta^{(1)}$
$\boldsymbol{u}^{(1)} = (1,0,0,0,0)$	$\boldsymbol{u}^{(1)} = (1,0,0,0,0)$ + rot	yes	$5 \cdot \eta^{(1)}\eta^{(1)}$
$\boldsymbol{u}^{(1)} = (1,0,0,0,0)$	$\boldsymbol{u}^{(1)} = (0,1,0,0,0)$ + rot	no	$5 \cdot 2\eta^{(1)}\eta^{(1)}$

Providing another example, namely the one on $\boldsymbol{u}^{(2)}$, one has the following possible combinations

left	right	left $\rightleftharpoons$ right	
$\boldsymbol{u}^{(2)} = (1,1,0,1,0)$	$\boldsymbol{u}^{(1)} = (0,0,1,0,0)$	no	$2\eta^{(1)}\eta^{(2)}$
$\boldsymbol{u}^{(2)} = (1,0,1,1,0)$	$\boldsymbol{u}^{(1)} = (0,1,0,0,0)$	no	$2\eta^{(1)}\eta^{(2)}$
$\boldsymbol{u}^{(2)} = (1,0,1,1,0)$	$\boldsymbol{u}^{(2)} = (1,1,0,1,0)$	no	$2\eta^{(2)}\eta^{(2)}$
$\boldsymbol{u}^{(2)} = (1,0,1,1,0)$	$\boldsymbol{u}^{(3)} = (0,1,0,1,0)$	no	$2\eta^{(2)}\eta^{(3)}$
$\boldsymbol{u}^{(2)} = (1,1,0,1,0)$	$\boldsymbol{u}^{(3)} = (1,0,1,0,0)$	no	$2\eta^{(2)}\eta^{(3)}$
$\boldsymbol{u}^{(3)} = (1,0,1,0,0)$	$\boldsymbol{u}^{(3)} = (0,1,0,1,0)$	no	$2\eta^{(3)}\eta^{(3)}$

Proceeding in the same fashion for $\boldsymbol{u}^{(1)}$ and $\boldsymbol{u}^{(3)}$ we end up with an explicit form of the survey propagation equations (4.20) for 5-circular coloring of random 3-regular graphs

$$\begin{aligned}
\eta^{(0)} &\cong \eta^{(0)}\eta^{(0)} + 5 \cdot 2 \cdot \eta^{(0)}\eta^{(1)} + 5 \cdot 3 \cdot \eta^{(1)}\eta^{(1)} \\
\eta^{(1)} &\cong 2 \cdot \eta^{(0)}\eta^{(3)} + 2 \cdot \eta^{(1)}\eta^{(1)} + 6 \cdot \eta^{(1)}\eta^{(3)} + 2 \cdot \eta^{(3)}\eta^{(3)} \\
\eta^{(2)} &\cong 4 \cdot \eta^{(1)}\eta^{(2)} + 2 \cdot \eta^{(2)}\eta^{(2)} + 4 \cdot \eta^{(2)}\eta^{(3)} + 2 \cdot \eta^{(3)}\eta^{(3)} \\
\eta^{(3)} &\cong 2 \cdot \eta^{(0)}\eta^{(2)} + 6 \cdot \eta^{(1)}\eta^{(2)} + 4 \cdot \eta^{(1)}\eta^{(3)} + \eta^{(2)}\eta^{(2)} + 4 \cdot \eta^{(2)}\eta^{(3)} + 2 \cdot \eta^{(3)}\eta^{(3)} ,
\end{aligned}$$

where the right- and left hand sides indicate iteration k and $k+1$ respectively. Again, assuming the symmetry, the normalization is given as $\eta^{(0)} + q \cdot \sum_i \eta^{(i)}$.

4.5 Results for Random Regular Graphs

We present the results for regular random graphs. Since every node has the same degree. The neighborhood of almost every node in regular random graphs looks the same up to a distance $\Theta(\log N)$, we therefore assume that the cavity message is the same for every node/edge. This simplifies considerably the numerical analysis of the corresponding fixed point equations.

4.5.1 Replica Symmetric Phase Diagram

In this section we investigate the solution of the replica symmetric BP equations on random d-*regular* graphs. In this case the message on every edge is equal and a physical solution

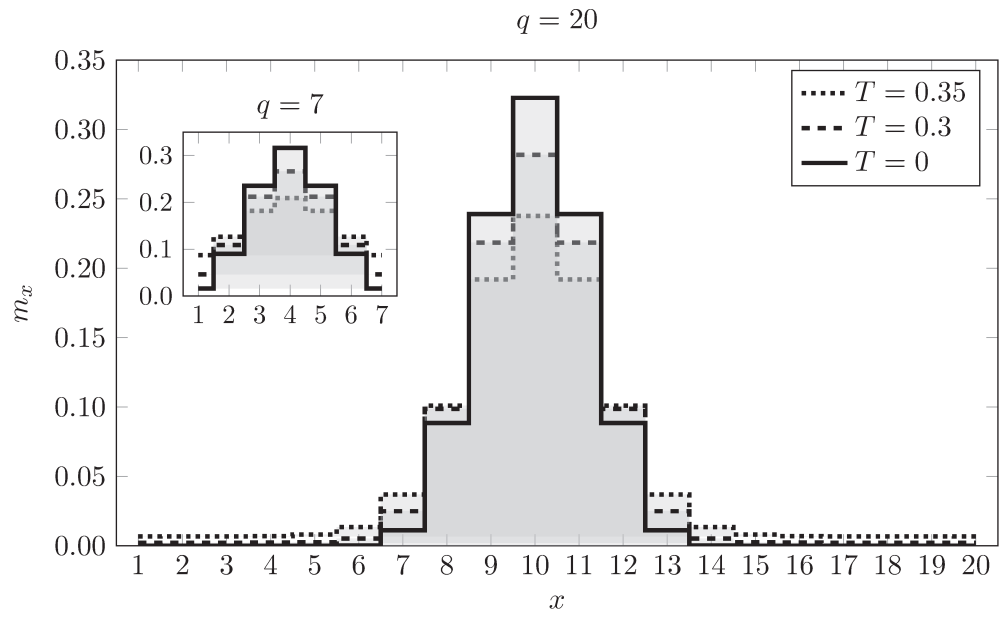

Figure 4.2: Exemplary ferromagnetic replica symmetric solutions for $q = 20$ ($q = 7$ in the inset) CCOL on random 3-regular graphs. We observed that roughly for $q \geq 11$ the RS zero temperature solutions are quasi indistinguishable.

of (4.2) must be uniform over the edges. Accordingly, we can simplify (4.2) as

$$m_x = \mathcal{F}(\{m\}) = \frac{1}{Z}[\mathrm{e}^{-\beta} + (1 - \mathrm{e}^{-\beta})(m_{x+1} + m_{x-1})]^{d-1} \,. \tag{4.21}$$

We associate the Bethe free energy (4.3) to each fixed point of (4.21) and the equilibrium solution is then obtained by selecting the fixed point that has minimal free energy.

One solution is the paramagnetic one, with a corresponding fixed point

$$m_{x_i}^{i \to j} = \frac{1}{q} \quad \forall \;\; x_i, \; (ij) \in \mathcal{E} \,. \tag{4.22}$$

However, for sufficiently large values of q we find another stable fixed point that has ferromagnetic character. In this ferromagnetic fixed point the probability to have a given color concentrates in few adjacent colors. Examples of the form on this ferromagnetic fixed point for random 3-regular graphs are depicted in Fig. 4.2. We verified by careful numerical search over initializations that no other fixed points appear for the cases we studied.

In order to draw the replica symmetric phase diagram, we evaluate the linear stability of the fixed points of (4.21) via the spectrum of the linearization matrix,

$$\mathcal{T}^{\tau\sigma} \coloneqq \left.\frac{\partial m_\tau}{\partial m_\sigma}\right|^{\mathsf{RS}} = (d-1)(1-\mathrm{e}^{-\beta}) \left\{ \chi_\tau \left(\delta_{\tau-1,\sigma} + \delta_{\tau+1,\sigma} \right) - \left(\chi_{\sigma+1} + \chi_{\sigma-1} \right) m_\tau \right\} \,, \tag{4.23}$$

where $\chi_\tau = \frac{m_\tau}{\mathrm{e}^{-\beta} + (1-\mathrm{e}^{-\beta})(m_{\tau-1} + m_{\tau+1})}$. Concerning the instability of the paramagnetic solution towards the ferromagnet we look for the smallest temperature T_{S} at which the largest

positive eigenvalue of the matrix $\mathcal{T}$ is smaller than one. The negative eigenvalues of $\mathcal{T}$ indicate an anti-ferromagnetic instability which appears on trees and other bipartite graphs, but that is unphysical on random graphs, therefore the negative eigenvalues do not lead to a physical instability.

For the paramagnetic fixed point, with $m_\tau = 1/q$, the matrix $\mathcal{T}$ is a circulant matrix and the non-zero eigenvalues ν_j read

$$\nu_j = (d-1)\frac{2}{q}\frac{1-\mathrm{e}^{-\beta}}{\mathrm{e}^{-\beta}+(1-\mathrm{e}^{-\beta})\frac{2}{q}}\cos\frac{2\pi j}{q}\ ,\ j=1,\ldots,q-1\,. \tag{4.24}$$

The largest positive eigenvalue is ν_1, the paramagnetic phase is stable if and only if $\nu_1 < 1$. At zero temperature $\beta \to \infty$ the degree of the graph needs to be larger than $d > d_F$ where

$$d_{\mathrm{F}}^{\mathrm{RS}}(q) = 1 + 1/\cos(2\pi/q) \tag{4.25}$$

for the ferromagnetic instability to appear. Analogously we will call $q_{\mathrm{F}}^{\mathrm{RS}}(d)$ the largest value of q for which the replica symmetric calculations predicts existence of the ferromagnetic phase. When this condition is satisfied the paramagnetic solution becomes unstable towards the ferromagnet for temperatures lower than the so-called paramagnetic spinodal temperature

$$T_{\mathrm{SP}} = 1/\ln\left[1+\frac{q}{2\left((d-1)\cos\frac{2\pi}{q}-1\right)}\right]\,. \tag{4.26}$$

For random 3-regular graphs we computed the detailed replica symmetric phase diagram as a function of the number of colors, q, and temperature, T, depicted in Fig. 4.3. In terms of the number of colors, q, there are three different regimes in the replica symmetric results. For $q < 7$ the only solution for all temperatures is the paramagnetic one. In the intermediate regime for $7 \leq q \leq 18$ a second order phase transition at $T_{\mathrm{c}} = T_{\mathrm{SP}}$ separates the ferromagnetic phase (where the ferromagnetic solution is the only stable one) from the paramagnetic phase (where the paramagnetic solution is the only stable one). Eventually for $q > 18$ the transition becomes first order: the ferro- and paramagnetic phases coexist for an interval of temperatures $T_{\mathrm{SP}} < T < T_{\mathrm{SF}}$. In this regime of phase coexistence the thermodynamically correct solution is the one with smallest free energy. For $T < T_{\mathrm{c}}$ the ferromagnetic free energy is smaller than the paramagnetic one and hence favorable. At $T = T_{\mathrm{c}}$ the two free energies are crossing and for $T_{\mathrm{SF}} > T > T_{\mathrm{c}}$ both states still co-exist, but now the free energy of the paramagnetic state is smaller and therefore favorable.

The existence of the ferromagnetic phase can be understood intuitively. For a large number of colors it is very unlikely that two random colors will be adjacent, hence the problem is much more constrained than with fewer colors. The energetic gain from using only few colors that are close to each other on the cycle is overwhelming the entropic gain from using many colors for low enough temperature. In Sec. 4 it was shown that CCOL with even q can be reduced to $q = 2$. Interestingly, the replica symmetric investigation does not reproduce this fact.

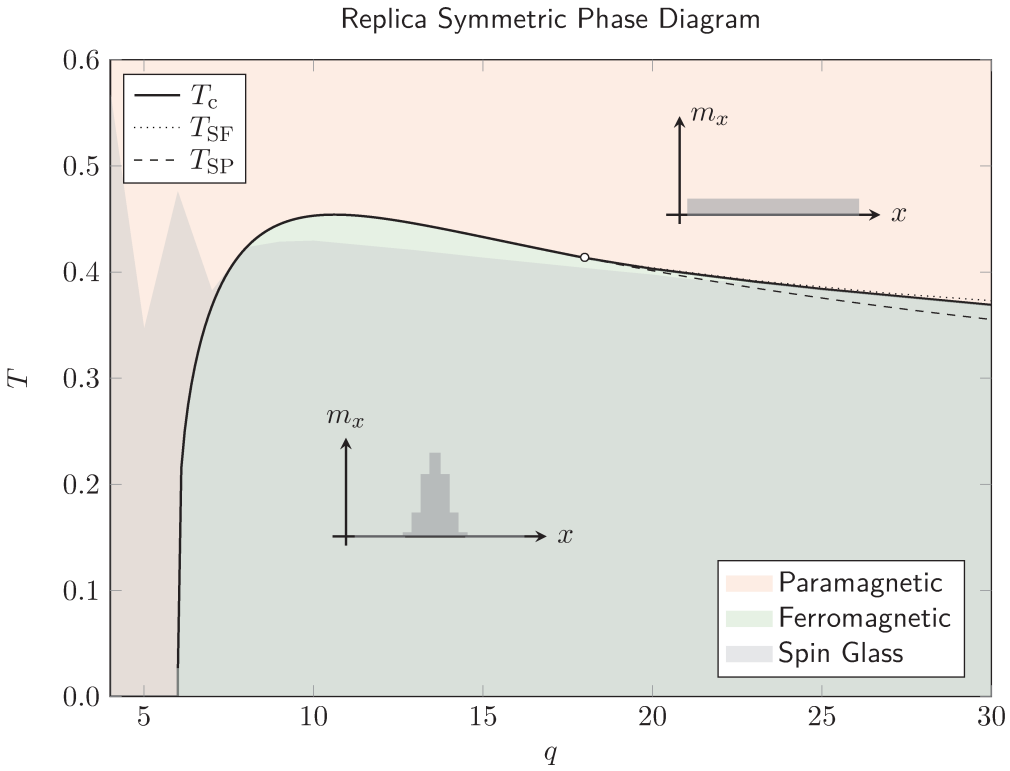

Figure 4.3: The replica symmetric phase diagram for circular coloring of 3-regular random graphs as a function of the number of colors q and the temperature T. We plotted the critical temperature T_c, the spinodal lines T_{SP} and T_{SF} as the limits of stability of the para- and ferromagnetic solution respectively, and the spin glass stability which is the temperature below which the replica symmetric Ansatz loses its validity. The critical temperature is plotted as thick (thin) solid line for $q < 18$ ($q \geq 18$). For $q < 18$ the transition is second order and the spinodal line coincides with the critical temperature. For $q \geq 18$ the transition becomes first order and the two spinodal lines separate from the critical line and a regime of phase coexistence that grows with q appears.

4.5.2 Spin-Glass Instability

In order to investigate the spin-glass stability the Jacobian (4.6) must be considered and its leading eigenvalues must be computed. For the paramagnetic fixed-point the matrix is circulant and the computation can be done analytically. There are only two kind of entries:

$$J^{\tau\sigma} = \begin{cases} (\frac{1}{q} - \frac{2}{q^2}) \frac{1-\mathrm{e}^{-\beta}}{\mathrm{e}^{-\beta}+(1-\mathrm{e}^{-\beta})\frac{2}{q}} & \text{if } \sigma = \tau \pm 1 \mod q \\ -\frac{2}{q^2} \frac{1-\mathrm{e}^{-\beta}}{\mathrm{e}^{-\beta}+(1-\mathrm{e}^{-\beta})\frac{2}{q}} & \text{else} \end{cases} \tag{4.27}$$

and we obtain the following eigenvalues

$$\lambda_j = \begin{cases} 0 & j = 0 \\ \frac{\cos\left(\frac{2\pi}{q} j\right)}{1+\frac{\mathrm{e}^{-\beta}}{(1-\mathrm{e}^{-\beta})}\frac{q}{2}} & j = 1, 2, \ldots, q-1\,, \end{cases} \tag{4.28}$$

from which the leading eigenvalue is obtained by selecting

$$j_{\mathsf{max}} = \arg\max |\lambda_j| = \begin{cases} q/2 & q \text{ even and } q \geq 4 \\ (q \pm 1)/2 & q \text{ odd and } q \geq 4\,. \end{cases} \tag{4.29}$$

With the stability condition [Tho86]

$$\kappa \lambda_{\max}^2 < 1\,,$$

where κ is the average excess degree, $\kappa = d - 1$ for d-regular graphs, and $\kappa = c$ for Erdős-Rényi graphs, we obtain the spin glass temperature

$$T_{\mathrm{SG}} = 1/\ln\left(1 + \frac{q}{2\left(\sqrt{d-1}|\cos\frac{2\pi}{q} j_{\mathsf{max}}| - 1\right)}\right)\,. \tag{4.30}$$

For the ferromagnetic solution we rely on solving (4.6) numerically and then extracting T_{SG} as the point where $(d-1)\lambda_{\max}^2 = 1$, with $\lambda_{\max}$ being the eigenvalue of the Jacobian that has largest absolute value. Note that the spin-glass transition can happen before, after or simultaneously with the ferromagnet transition.

For $2 \leq q \leq 3$ CCOL is equivalent to regular coloring and the spin-glass temperature can be computed [ZK07] from $T_{\mathrm{SG}}^{\mathrm{col}} = -1/\ln\left(1 - q/(1+\sqrt{d-1})\right)$. For 3-regular random graphs we get $T_{\mathrm{SG}} = 0.567$ when $q = 2$ and an always RS stable solution for $q = 3$. When $4 \geq q \geq 6$ no ferromagnetic solution exists and the spin-glass transition temperature is given by (4.30). We observe that for $q = 7$ the transition from RS to RSB happens in the paramagnetic phase before the transition to the ferromagnetic state happens, i.e., $T_{\mathrm{SG}} > T_{\mathrm{c}}$. At $q = 8$ the two transitions coincide (compare equations (4.30) and (4.26)) and we have a tri-critical point: the system becomes ferromagnetic at the same temperature where it loses its RS stability. For $q \geq 8$ the ferromagnetic transition happens before the RSB transition. Eventually, in the large q limit the two transitions coincide again and the ferromagnetic solution is never RS stable and $T_{\mathrm{c}} = T_{\mathrm{SG}}$. For a summary see Fig. 4.3, and Fig. 4.4 for examples of the instabilities.

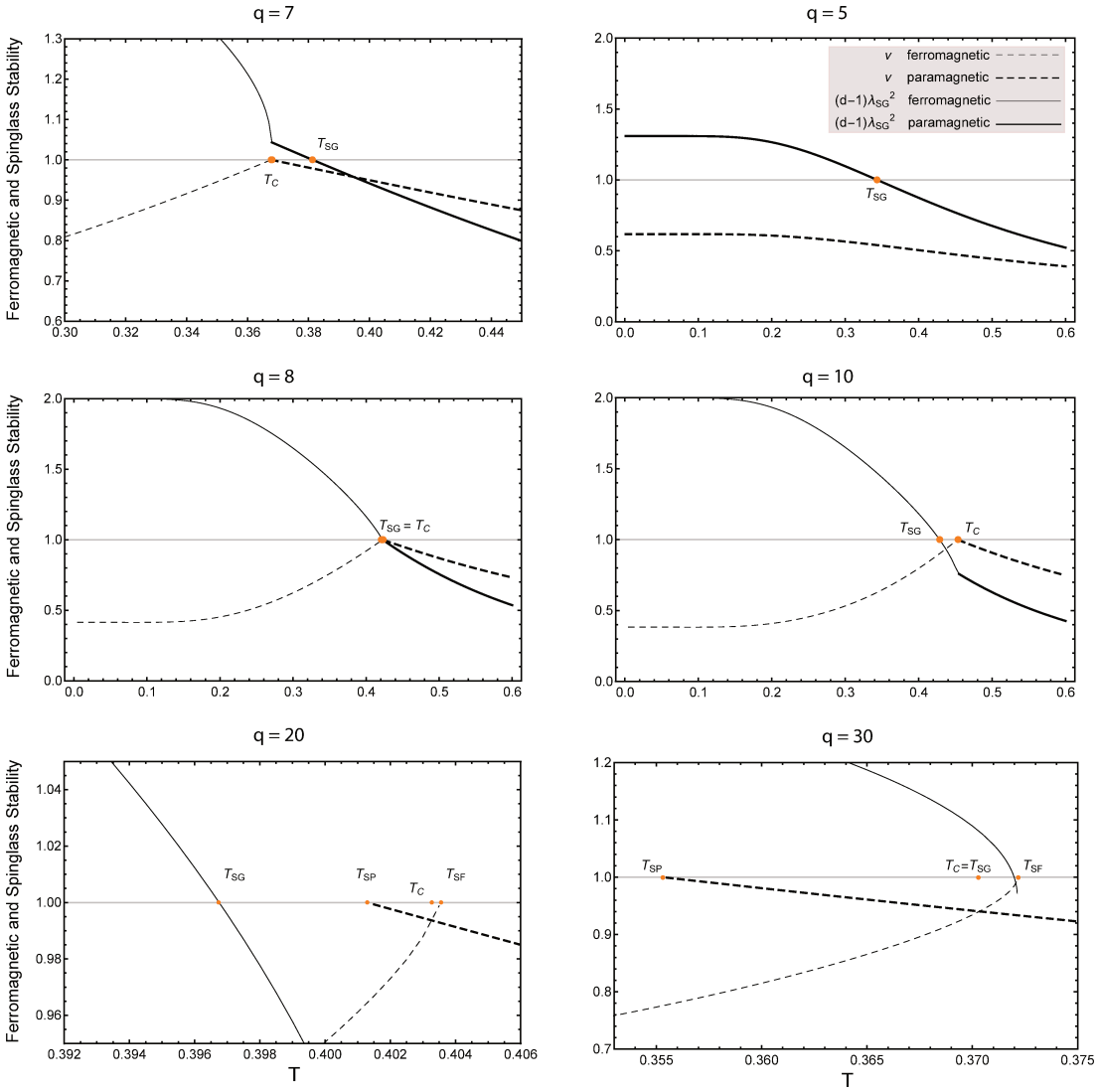

Figure 4.4: Here we show the spin-glass stability condition as a function of the temperature for representative values of q as discussed in the text. The full line represents the spin glass stability parameter of the thermodynamically stable replica symmetric solution (thick for paramagnet, thin for ferromagnet). The point where the lines cross one is the spin-glass temperature $T_{\rm SG}$. The dashed line represents the thermodynamic stability regions for the paramagnet (thick), and the ferromagnet (thin). In the last two cases ($q = 20$ and $q = 30$) the paramagnetic spinglass stability is smaller than the lower limit of the y axis and does not appear in the figures.

4.5.3 1RSB in the Energetic Zero Temperature Limit

In section 4.4 we described the 1RSB energetic zero temperature limit. Using population dynamics for equations (4.18) and (4.19) we obtain the replicated free entropy $\Phi(y)$ and the energy $\epsilon(y)$, both as a function of y, from (4.9) and (4.3) respectively. The complexity as a function of the energy is then obtained from the Legendre transformation $\Phi(y) = -y\epsilon + \Sigma(\epsilon)$. The value of ϵ for which the concave branch of $\Sigma(\epsilon)$ intersects the ϵ-axis yields the 1RSB estimate for the ground state energy.

The result of the energetic zero temperature cavity limit are very different depending on whether the number of colors is even or odd. For even number of colors, the results are exactly the same as for $q = 2$, as should be expected. For an odd number of colors, the results can roughly be deduced from the replica symmetric solution with a delay of the onset of the ferromagnetic transition.

Our results are obtained using the population dynamics, with an initialization such that each initial field $\boldsymbol{h}$ in the population is pointing into the direction of one color only, i.e. $\boldsymbol{h} = \hat{\mathbf{e}}_\tau$. And τ is drawn with probability equal to the replica symmetric estimate $P_{\mathrm{RS}}(\tau)$.

Even Number of Colors

For even number of colors the ground state energy of the circular coloring is the same as the ground state of the 2-coloring. Assume $q = 2z$ with $z \geq 2$ being an integer. Identifying $s = 2z + 1 \to 1$ and $s = 2z \to 2$ for $s \in \{1, \ldots, q\}$, the number of contradictory edges in this 2-coloring must be equal or smaller than it was before. On the other hand, any 2-coloring configuration is a valid configuration of $2z$-circular coloring.

The ground state of 2-coloring is in turn asymptotically equal (as argued e.g. in [ZB10]) to the ground state of the Viana-Bray model [VB85], for which the energetic 1RSB results were studied in [MP03]. This equality of ground state energies, together with the above mentioned work on the Viana-Bray model, makes it unnecessary to further investigate the case $q = 2z,\ z \in \mathbb{N}^+$ here. However, it is important to point out briefly that this independence of z also comes up in the energetic zero temperature limit of the cavity equations, as opposed to the RS approach.

The 1RSB cavity method predicts

$$E_{\mathrm{gs}}(q = 2z, d) = E_{\mathrm{gs}}(q = 2, c)\,, \quad \text{with} \quad \Sigma\left[E(q = 2z, d)\right] = \Sigma\left[E(q = 2, d)\right]\,.$$

The SP-y equations converge towards the same fixed point for all even q when initialized as described above. In this fixed point the population consists of only two types of messages. The alternating warnings $\boldsymbol{u} = (1, 0, 1, 0, \ldots)$ and $\boldsymbol{u} = (0, 1, 0, 1, \ldots)$ and the trivial warning $\boldsymbol{u} = (0, 0, \ldots)$ with probabilities

$$P\left[\boldsymbol{u} = (1, 0, 1, 0, \ldots)\right] = P\left[\boldsymbol{u} = (0, 1, 0, 1, \ldots)\right] = \eta/2$$

and the zero warning

$$P\left[\boldsymbol{u} = (0, 0, \ldots)\right] = 1 - \eta\,.$$

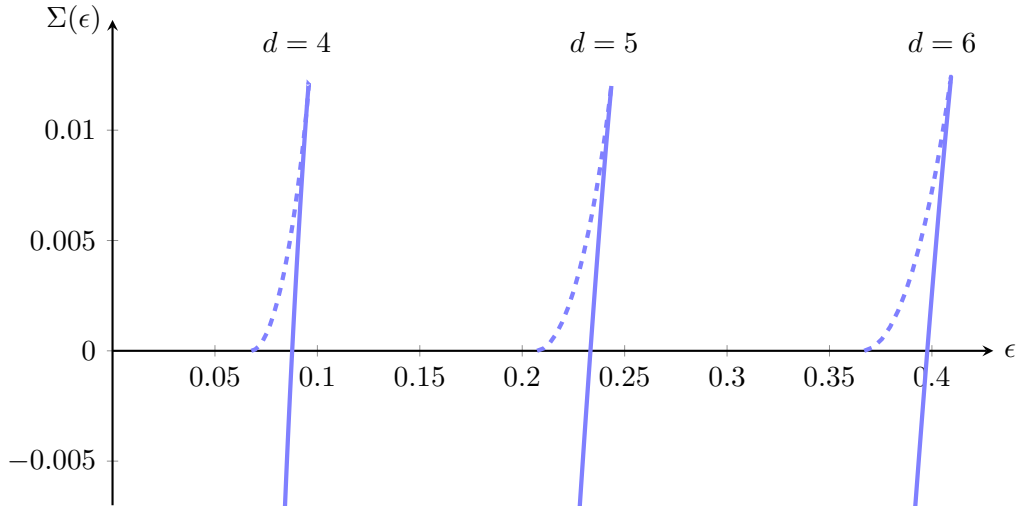

Figure 4.5: Some examples of the complexity curve as a function of the energy for $q = 5$. The physical (concave) branches are depicted as full lines and the non-physical branches are dashed. The ground state energy can be extracted from the intersection point of the physical branch with the energy axis where $\Sigma(\epsilon) = 0$.

The corresponding ground state energies and zero energy complexities were evaluated in details in previous works on the Viana-Bray model [MP03]. For instance, for random d-regular graphs, the 1RSB ground state energy is $\epsilon_{\rm gs}(d = 3) = 0.1138$, $\epsilon_{\rm gs}(d = 4) = 0.2635$, $\epsilon_{\rm gs}(d = 5) = 0.4124$ [ZB10]. It is known that in the Viana-Bray model the 1RSB solution for the ground state is not stable towards further levels of RSB. Since the equality of ground states for every $q = 2z$ is a rigorous statement also in further levels of RSB this would have to come up.

Odd Number Colors

Contrary to the case of even colors which could be reduced to $q = 2$ the population does not reduce to the simple subset of messages. In fact, the alternating warning *cannot* exist if q is an odd number. Figure 4.5 depicts some exemplary results, Tab. 4.3 summarizes the ground state energies obtained for different values of degree d and number of colors q.

In particular we find that $E_{\rm gs} > 0$ for all considered instances, except for $q = 5,\ d = 3$ the particular case for which we only obtain the trivial solution $P[\boldsymbol{u} = (0, 0, \dots)]$ when employing the energetic zero temperature limit. These results agree with what was found in previous rigorous investigations, cf. [KNS01; WW01; Hat05]. It can be expected from physical intuition that, as the number of colors, q, grows, the ground state energy for q odd converges to the one for q even. Our numerical results are consistent with this.

The RS investigation indicated the presence of a purely paramagnetic phase for all temperatures when $q < q_{\rm F}^{\rm RS}$ and the presence of a phase transition towards a ferromagnetic phase when $q \geq q_{\rm F}^{\rm RS}$. In the framework of 1RSB we monitor the magnetization by tracking an

d	q	RS phase	RSB phase	$\epsilon_{\rm gs}$
3	5	para	para	0
3	7	ferro	para	0.030
3	9	ferro	para	0.064
3	11	ferro	ferro	0.052
3	13	ferro	ferro	0.071
3	15	ferro	ferro	0.076
3	19	ferro	ferro	0.084
3	29	ferro	ferro	0.094
3	51	ferro	ferro	0.100
4	5	para	para	0.088
4	7	ferro	para	0.189
4	9	ferro	ferro	0.155
4	11	ferro	ferro	0.171
5	5	ferro	para	0.233
5	7	ferro	ferro	0.325
5	9	ferro	ferro	0.328
5	11	ferro	ferro	0.337

Table 4.3: Results of the energetic 1RSB analysis for the circular coloring. The ground state energy $\epsilon_{\rm gs}$ is obtained from the intersection point of the concave branch of $\Sigma(\epsilon)$ with the energy axis (cf. figure 4.5). We also provide which phase the zero temperature solution is found in the RS and 1RSB framework. The 1RSB zero temperature solution becomes ferromagnetic at $q = 7$ for $d \geq 5$.

order parameter $\overline{m} = \int \mathrm{d}m P(m)\, m$. As compared to the RS solution, we observe a shift of the appearance of the ferromagnet to to lower temperature. In particular, according to the 1RSB solution, the ferromagnet appears only at $q = 11$ for $d = 3$, $q = 9$ for $d = 4$ and $q = 7$ for $d \geq 5$. Recalling the RS phase diagram for $d = 3$ (Fig. 4.3), one might wonder about the finite temperature behavior. The 1RSB solution for $q = 7$ and $q = 11$ are respectively of purely paramagnetic and ferromagnetic character in the whole range $T < T_{\rm SG}$. However, anticipating the finite temperature investigation, for $q = 9$ we find a re-entrance of the 1RSB solution into a paramagnetic phase for $T < 0.125$. The zero temperature results are summarized in Tab. 4.3.

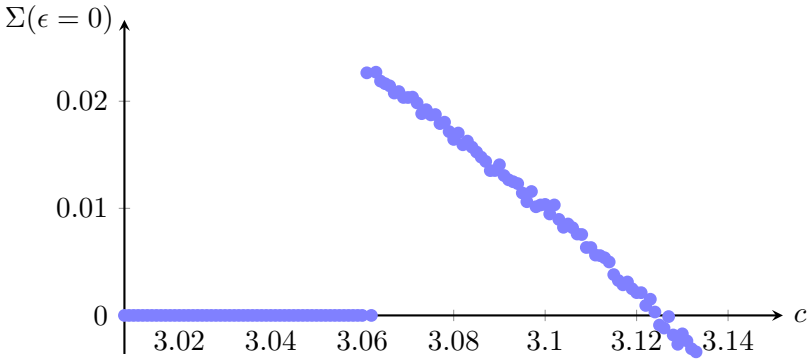

Figure 4.6: Zero energy complexity against the fraction r of degree $d = 4$ nodes in a graph with only degrees 3 and 5 for 5-circular coloring. A noncontinuous jump happens around $c = 3.06$ when the complexity suddenly becomes positive. Subsequently it is continuously decreasing and becomes negative around $c = 3.125$.

The case of 5-circular coloring

The very fact that the predictions obtained in the energetic zero temperature limit agree with previous rigorous results and the fact that even colors can always be reduced to $q = 2$ is not trivial and provides some solid ground for expanding our investigations further to the particular case of $q = 5,\ d = 3$ for which the energetic zero temperature limit does not provide a useful non-trivial estimate.

The absence of non-trivial solution suggests that no constrained fields are present, and thus $\Sigma(\epsilon = 0) = 0$. Consequently an interesting follow up question is to ask what happens when the degree is varied from $d = 3$ to $d = 4$. In the latter case constrained fields are present and $\Sigma(\epsilon = 0) \leq 0$. In order to investigate the intermediate regime where the average degree $3 < c < 4$ we work on the ensemble of graphs with a fraction $1 - r$ of degree $d = 3$ nodes and fraction r of $d = 4$. Tuning r facilitate the investigation of the zero energy complexity in the whole region $3 \leq c \leq 4$, where $c = r \cdot 4 + (1 - r) \cdot 3$. The survey propagation can be carried out for each r and the resulting $\Sigma(\epsilon = 0, r)$ is depicted in Fig. 4.6. No non-trivial solution is found and the complexity remains zero until a discontinuous jump happens around $r = 0.06$ where the complexity suddenly becomes positive and constrained messages appear in the population. Subsequently the curve is continuously decreasing and becomes negative around $r = 0.125$ where no more zero energy solution can be found. This suggests that the local constraint density is sufficiently small for $r < 0.125$ and coloring is possible. However, no confined variables are present for $r < 0.06$ and to enable a quantitative statement for the case $q = 5,\ d = 3$ we must consult the entropic zero temperature limit.

We also considered Erdős-Rényi graphs. Somewhat strangely we found that the zero energy complexity never becomes positive. Instead a solution with negative complexity appears at about $c \approx 2.27$. Recall that (4.28) yields an average degree of $c_{\mathrm{d}} = 1.53$ for the spin glass stability transition point. We solved the corresponding survey propagation equations using

population dynamics of populations with population sizes 10000. It cannot be excluded that what we observed was a finite population size effect, but it may also be that the usual 1RSB solution with confined variables, that leads to the satisfiability threshold in the widely studied K-SAT and coloring problems, is not sufficient to get a sensible estimate of the threshold for the 5-circular coloring of the Erdős-Rényi random graphs.

4.5.4 1RSB in the Entropic Zero Temperature Limit

We briefly summarize the procedure that enables access to the thermodynamic quantities for the case $q = 5,\ d = 3$ before presenting the results. An advantage of working on regular graphs is the reduction of the required computational effort necessary in the population dynamics. This is because the probability distribution over the messages in (4.7) is identical for every edge $\mathcal{P}_{i\to j}(\boldsymbol{m}) = \mathcal{P}(\boldsymbol{m})$. For a given reweighting parameter m the 1RSB equation (4.7) can be solved with the population dynamics technique. Solving it for every m yields the thermodynamic quantities as function of the reweighting parameter m. These are $\Phi_s(m)$, which follows from (4.9), $\Sigma(m)$, which is obtained from the Legendre identities and $s(m)$ which can be computed with (4.3). All in the limit $\beta \to \infty$ where $-\beta f = s$. In order to assign the right weights to each pure state in the equilibrium configuration (1RSB estimate), the reweighting parameter m must be chosen accordingly.

- If $\Sigma(m = 1) > 0$ we are in the dynamic 1RSB phase and the 1RSB and RS solutions agree.
- Else, if a non-trivial solution exists for $m = 1$ and $\Sigma(m = 1) < 0$ we must choose the thermodynamic value of the reweighting parameter m^* as the smallest positive non-zero value for which the complexity vanishes, i.e. $\Sigma(m^*) = 0$.

Figure 4.7 presents the results obtained within the entropic zero temperature limit. In this limit the energy vanishes $\epsilon = 0$ and $\Sigma(s)$ counts the total number of clusters of size s. The point where the physical (concave) part of the complexity curve intersects with the entropy axis, i.e., where $\Sigma(s) = 0$, provides the 1RSB estimate for the entropy as the entropy of the entropically dominating clusters. Following the positive part of the physical branch of the complexity curve in Fig. 4.7b more and more (increasing complexity) subdominant clusters of smaller entropy appear. Eventually, at the cusp (that is the point where the two branches of $\Sigma(s)$ meet in Fig. 4.7b where $s(m)$ is minimal in Fig. 4.7a) the 1RSB solution loses its validity (it becomes unstable towards more levels of RSB before the cusp is reached).

The thermodynamic value of the reweighting parameter was found to be $m^* = 0.299$ with an 1RSB estimate of the entropy $s_{\mathrm{1RSB}} = 0.223$ (to be compared with the replica symmetric estimate of the entropy $s_{\mathrm{RS}} = 0.235$). Finding $m^*(T = 0) \neq 1$ (or more precisely the fact that $\Sigma(m = 1) < 0$) implicates that the solution space of the problem is beyond the condensation transition [KMR+07]. This 1RSB result is a very strong indication towards 5-circular colorability of random 3-regular graphs. In the next section we will show that the 1RSB approach is actually not stable, and further steps of replica symmetry breaking are

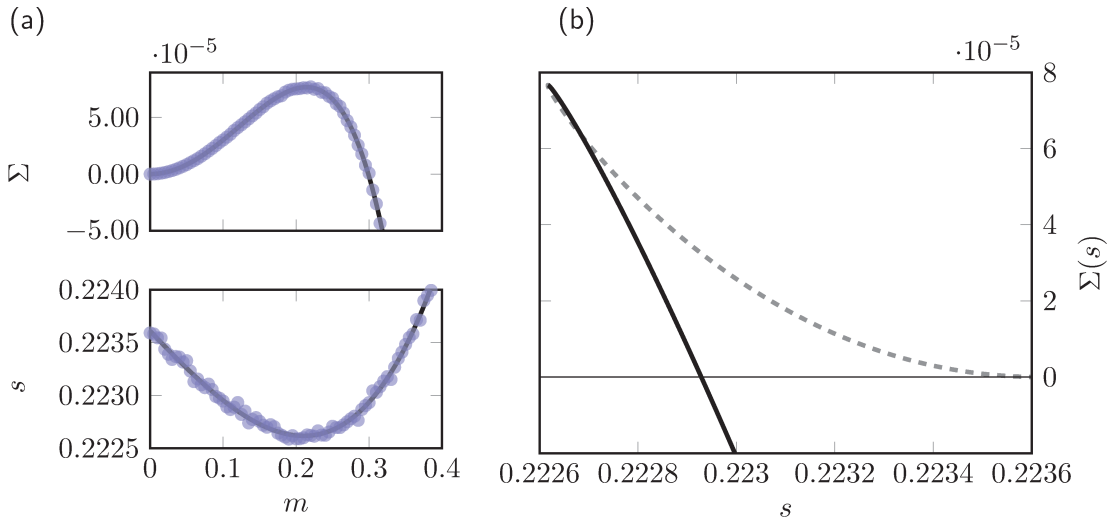

Figure 4.7: Results obtained within the entropic zero temperature limit for 5-circular coloring of 3-regular random graphs. The two plots in (a) depict the complexity and entropy as a function of the reweighting parameter – $\Sigma(m)$ and $s(m)$ respectively. Figure (b) presents the resulting complexity as a function of the entropy $\Sigma(s)$. The results were first fitted with a function of type $a + b\mathrm{e}^{-x} + c\mathrm{e}^{-2x} + d\mathrm{e}^{-3x} + \ldots$ and subsequently the change of variables from $\Sigma(m)$ to $\Sigma(s)$ was performed with the fit function. The resulting complexity function contains a physical branch (solid line) as well as an unphysical branch (dashed line). The 1RSB estimate for the entropy of clusters of zero energy are obtained from the intersection of the physical branch of the complexity function with the entropy axis.

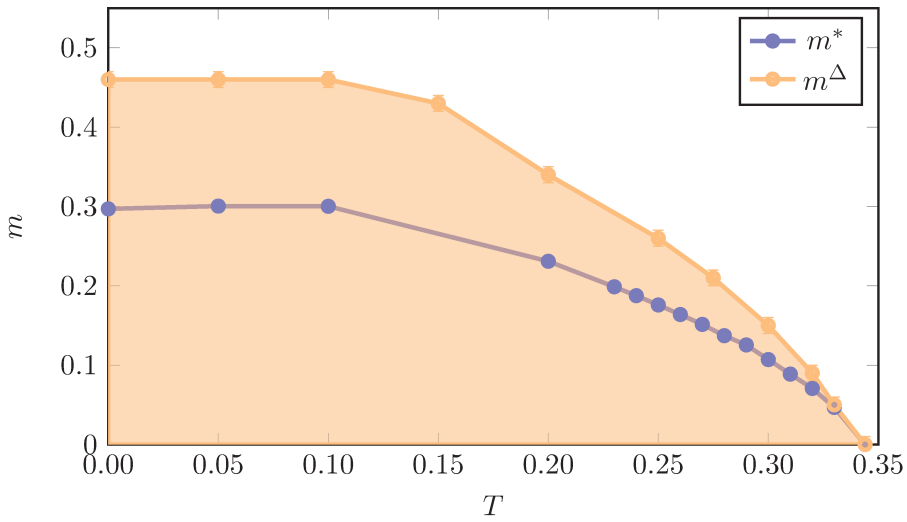

Figure 4.8: The m, T phase diagram for $T < T_{\mathrm{SG}} = 0.344$ for the 5-circular coloring of 3-regular random graphs. The blue squares mark the values m^Δ at which the 1RSB solution becomes unstable. The whole shaded area below the m^Δ-line is unstable towards further steps of RSB. The orange diamonds denote the thermodynamic values of the reweighting parameter m^*, extracted as $\Sigma(m^*) = 0$. For all relevant temperatures $m^* < m^\Delta$ holds and hence the 1RSB solution is unstable and further steps of replica symmetry breaking are necessary.

necessary. However, these effects usually change the value of the 1RSB ground state entropy very little, so we conjecture that the colorability remains.

4.5.5 Finite Temperature and 1RSB Stability

In this section the m, T phase diagram for the 5-circular coloring of random 3-regular graphs is investigated in order to study the stability of the 1RSB solution towards further steps of replica symmetry breaking within the whole regime $T < T_{\mathrm{SG}}$, where T_{SG} indicates the temperature at which the RS estimate becomes unstable. With $T_{\mathrm{SG}} = 0.344$ for the 5-circular coloring of random 3-regular graphs.

The stability of 1RSB towards more levels of RSB is related to the evolution of an infinitesimal perturbation in the messages, $Q(t)$, as we described in Sec. 4.3.2. The solution is considered stable towards further steps of RSB if $\lim_{t\to\infty} Q(t) \to 0$. The results are presented in figure 4.8, where the thermodynamic value of the reweighting parameter, m^*, as well as the value for which the instability sets in, m^Δ, are plotted as functions of the temperature. For a given temperature, T, the thermodynamic value of the reweighting parameter, $m^*(T)$, was extracted as the point where $\Sigma(m^*(T)) = 0$, and $m^\Delta(T)$ as the point where $Q(t) \geq Q(0)$ for large enough t. In the numerical tests $t = 1000$ was used.

As depicted in Fig. 4.8, the 1RSB solution lays in the unstable (shaded) region, i.e., $m^* < m^\Delta$, for all $T < T_{\mathrm{SG}}$ and thus the 1RSB solution is unstable for all relevant temper-

atures. Consequently, the 1RSB solution is not sufficient in this case and further steps of replica symmetry breaking are necessary. In the cases where the 1RSB framework fails, it is widely believed that the FRSB framework is necessary for a correct treatment. Interestingly, the $q = 5,\ d = 3$ case thus appears to be the first case of a CSP with exactly known degenerated zero ground state energy and FRSB structure of that we are aware. On the one hand, the FRSB structure might cause serious difficulties for a rigorous treatment via the cavity method. On the other hand, it renders $q = 5, d = 3$-CCOL a very interesting instance to be studied in terms of algorithmic consequences. It has previously been argued that a FRSB structure of the solution space is related to marginal stability and absence of basins that would trap physical dynamics for exponential time [CK94]. As a consequence simple search algorithms, such as simulated annealing, are expected to converge towards an optimal solution which we shall confirm, for CCOL, in the next section.

4.5.6 Algorithmic Consequences

In order to investigate the performance of simulated annealing (SA) for 5-circular coloring of random 3-regular graphs, we created random 3-regular graphs by randomly linking N nodes such that no self-loops, double edges or triangles are present (we simply dismiss the graphs containing triangles). We then assign colors to the N nodes uniformly at random and SA is performed. For each instance different system sizes, N, were considered and SA was run with different annealing rates δT for each of them. Different annealing schedules were tested. Among them, an exponential variation, with $T \to \frac{T}{1+\delta T}$, was found to be very efficient. The results of the implementation are depicted in Fig. 4.9: we plot the number of sweeps, necessary to find a proper coloring, as a function of the systems size, N, for a fixed annealing rate, δT. Each sweep takes N steps and the necessary number of sweeps is in good agreement with a logarithmic fit (or upper bounded by it). The total number of iterations thus appears to be $O(N \log N)$; which makes SA, equipped with the exponential schedule, a very efficient choice to find proper colorings for 5-circular coloring of random 3-regular graphs.

It is worth noting, that we also investigated belief propagation initialized in the solutions obtained from SA, and after sufficient number of iterations BP behaves exactly in the same way it does when initialized randomly.

In the previous section we saw that the problem has zero ground state energy, that no confined variables are present *and* that it likely features a FRSB structure. The very fact that SA, as a simple search algorithm, is able to find optimal solutions, supports the picture of a saddle point dominated landscape of solutions that may account for the algorithmic easiness of 5-CCOL on random 3-regular graphs.

This being said, the simulated annealing solves the 5-circular coloring problem on random 3-regular graphs with such an ease that we found it surprising. The circumstance that the dynamics on sparse graphs can be simulated in linear time, together with the fact that the ground state energy is known, makes the 5-circular coloring problem on random 3-regular graphs a unique playground to study numerically the nature of the FRSB phase in diluted

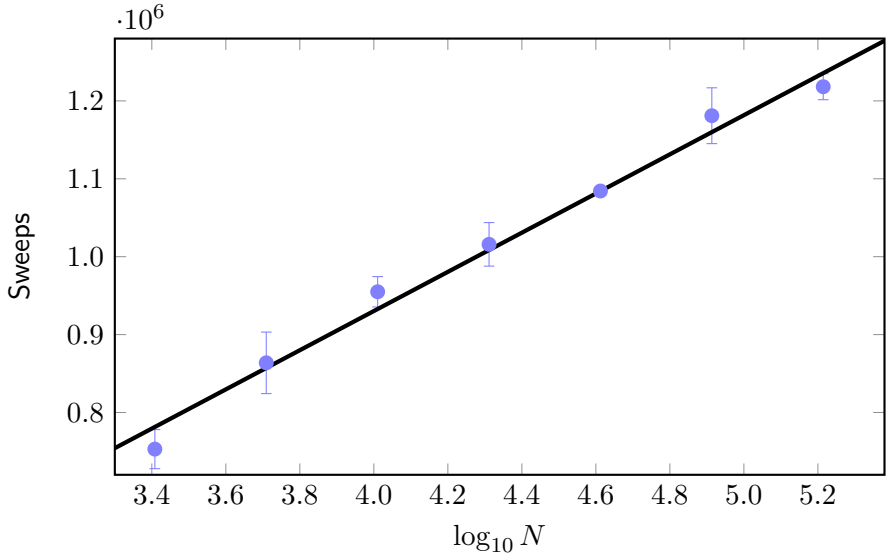

Figure 4.9: We plot the number of sweeps necessary to find a proper 5-circular coloring or a random 3-regular graph as a function of the size of the system N for exponential annealing rate $\delta T = 10^{-6}$.

systems.

The success of SA for $q = 5$ and $d = 3$ inspired us to approach the problem with an even simpler greedy strategy, that did not work. We find it, however, useful reporting this result. The greedy strategy successively assigns constraints to the nodes of the graph and hence attempts to obtain a coloring. Given an unconstrained graph, we start by picking a random node and constraining it to color s. Doing so, we constrain all direct neighbors to $\{s-1, s+1\}$, all second neighbors to $\{s, s-2, s+2\}$ and all third neighbors to $\{s-1, s-2, s+1, s+2\}$ (all $\bmod 5$). After this step we pick one of the maximally constrained nodes at random, we assign it one of the permitted colors at random and subsequently update the constraints of its neighborhood. Thereafter, a maximally constrained node is chosen and the procedure is repeated. We observe, that the probability of success is very small for small graphs and decreases further with the size of the graphs.

4.6 Conclusion

Motivated by Nešetřil's conjecture, we studied the circular colorability of random regular graphs. The statistical physics analysis showed that, w.h.p., random 3-regular graphs are 5-circular colorable. Therefore providing evidence that Nešetřil's conjecture holds. Our analysis is based on the zero temperature limit of the Cavity method. In order to resolve the satisfiability, it was necessary to employ the 1RSB energetic zero temperature limit to study the space of zero energy configurations, which lead to the above conclusion.

By virtue of the particular constraints, the problem was found exhibits several interesting

physical features. For a sufficiently large number of colors, a phase transition separates a paramagnetic high temperature phase, in which the distribution over colors is uniform, from a ferromagnetic low temperature phase, in which the distribution of colors concentrates. The energetic gain from using only few colors that are close to each other on the cycle is overwhelming the entropic gain from using many colors for low enough temperature. This is an interesting feature, caused by the strong constraints in CCOL, that is not observed in regular coloring. Within the 1RSB energetic zero temperature limit, the structure of the warning propagation equations is significantly more involved due to the difficult constraints. As compared to previously studied problems, this makes the analytical treatment rather cumbersome.

For further analysis of the stability of the 1RSB solution for 5-circular colorable on random 3-regular graphs, the finite temperature phase space had to be studied. The 1RSB solution was found to be unstable towards 2RSB, implying that the problem, likely, requires the FRSB framework. 5-circular coloring of random 3-regular graphs is hence the first instance of a *satisfiable* combinatorial problem with degenerated and *precisely known ground state* that requires FRSB. Note that such a behavior is not unique to the CCOL problem: the authors of [GDSZ17] recently found that a similar phenomenon also appears in the bicoloring of hypergraphs that can also be located in a satisfiable FRSB phase. Thereby making it a very interesting problem to be studied in the context of diluted models on its own. The fact that the ground state energy is known does also give rise to the hope for simplifications in rigorous investigations along the same lines.

Finally, the algorithmic consequences were examined. In the statistical physics community it has long been conjectured that a FRSB structure of the solution space is related to marginal stability and absence of basins that trap physical dynamics for exponential time. We could confirm this for CCOL by applying SA to typical instances. We found that an exponential annealing schedule is very efficiently finding proper coloring assignments. The numerical investigations suggest that SA, equipped with the exponential schedule, requires $O(N \log N)$ steps to succeed. Remarkably, although the statistical description of the space of proper colorings seems challengingly hard for the case of 5-circular coloring of random 3-regular graphs, the problem is algorithmically easy.

— 5 —

Crowdsourcing

Crowdsourcing is a strategy to categorize data through the contribution of many individuals. A wide range of theoretical and algorithmic contributions are based on the, previously introduced, model of Dawid and Skene.

My main contribution to this subject is the Bayes-optimal analysis of the model of Dawid and Skene in the dense regime. Additionally, numerical results are obtained on how these results translate into the sparse regime and experiments on a real world dataset are performed.

The first sections review and establish the link between the dDS model and the low-rank matrix estimation problems of [LKZ15; LKZ17]. Subsequently, some phase diagrams are drawn and discussed. Finally, the relevance of these results (obtained in the dense regime) in the sparse regime are analyzed and real world experiments are carried out.

In Sec. 1.3.4 the dense Dawid-Skene (dDS) model for crowdsourcing was introduced and the AMP equations were derived (Sec. 1.3.5) together with their corresponding state evolution (Sec. 1.3.6). The replica symmetric free energy was computed in Box 2.3. It was further shown, in section 1.3.5, that the dDS model belongs to a larger class of low-rank matrix factorization problems, as studied in [DM14; MT13; LKZ15; LKZ17; Mio17]. We briefly recall the main elements of the model, before analyzing the phase diagram.

Consider N workers and a total of M tasks. We denote with L_{ij} the label assigned to question j by worker i and assume $L_{ij} \in \{0, \pm 1\}$. If $L_{ij} = 0$, then question j was left out by worker i. We consider a system in which each worker is posed $(1-\rho)M$ questions. Letting the probability that worker i provides a correct answer be $1/2 + \sqrt{\nu/N}\theta_i$, the likelihood reads

$$\begin{aligned} P\left(L_{ij} = \pm 1 \mid \theta_i,\ v_j\right) &= (1-\rho) \cdot \frac{1}{2} \cdot \left(1 \pm \sqrt{\frac{\nu}{N}}\, \theta_i v_j\right) \\ P\left(L_{ij} = 0 \mid \theta_i,\ v_j\right) &= \rho\,, \end{aligned} \tag{5.1}$$

where the fraction of un-answered questions, ρ, was assumed to be independent of (i, j). The rest of the present section is set in the limit where $N \to \infty$ and all the other parameters $\theta_i, \nu, \alpha, \rho = \Theta(1)$. Later, in Sec. 5.2, we will discuss how to extrapolate the results into

the sparse regime where each worker is only assigned to $O(1)$ tasks. The parameter ν allows to tune the signal to noise ratio and the $1/\sqrt{N}$ scaling brings us into the interesting regime, where the problem is neither trivially hard, nor trivially easy. Denoting by $\boldsymbol{\theta} \in \mathbb{R}^N$ the vector of rescaled reliabilities for all N workers, and $\boldsymbol{v} \in \mathbb{R}^M$ the vector of labels, the rank-1 matrix

$$\boldsymbol{w} \coloneqq \frac{\boldsymbol{\theta}\,\boldsymbol{v}^T}{\sqrt{N}} \tag{5.2}$$

can be introduced and utilized to re-express (5.1) as $P\left(L_{ij} \mid w_{ij}\right) = \exp\left(g(L_{ij}, w_{ij})\right)$ with

$$g(L_{ij}, w_{ij}) = \begin{cases} \log\left(\frac{(1-\rho)}{2}\right) + \log\left(1 \pm \sqrt{\nu} w_{ij}\right) & \text{if } L_{ij} = \pm 1 \\ \log\left(\rho\right) & \text{if } L_{ij} = 0 \end{cases} . \tag{5.3}$$

The Bayes-optimal estimates, $\hat{\boldsymbol{\theta}}$, that minimize the expected mean-squared-error (MSE) on $\boldsymbol{\theta}$:

$$\mathrm{MSE}_\theta = \frac{1}{N}\sum_i (\hat{\theta}_i - \theta_i^0)^2 \,, \tag{5.4}$$

and the expected bitwise error (BE) on $\boldsymbol{v}$:

$$\mathrm{BE}_v = \frac{1}{M}\sum_j \mathbb{I}[\hat{v}_j \neq v_j^0] = \frac{1}{M}\sum_j (\frac{\hat{v}_j - v_j^0}{2})^2 = \frac{1}{2}\frac{1}{M}\sum_j (1 - \hat{v}_j v_j^0) \,, \tag{5.5}$$

were derived in Sec. 1.3.3 and read

$$\hat{\theta}_i^{\mathrm{MMSE}}(\boldsymbol{L}) = \int \mathrm{d}\theta_i \theta_i P_\theta\left(\theta_i \mid \boldsymbol{L}\right) \quad \text{and} \quad \hat{v}_j^{\mathrm{MBE}}(\boldsymbol{L}) = \mathrm{sign} \int \mathrm{d}v_j v_j P_v\left(v_j \mid \boldsymbol{L}\right) . \tag{5.6}$$

Here $P_x(x \mid \boldsymbol{L})$, with $x \in \{\{\theta_i\}_{i=1,\ldots,N}, \{v_j\}_{j=1,\ldots,M}\}$, are the marginal posterior probabilities. Inferring the reliabilities and labels in the crowdsourcing problem reduces to evaluating the marginal expectations of the posterior probability distribution. In general this is a difficult task.

In all but section 5.2.3, we assume that the distributions from which the ground truth reliabilities, θ_i^0, and labels, v_j^0, are drawn, P_{θ^0} and P_{v^0} respectively, are known. Under these assumptions we aim to achieve the following.

- Compute efficiently (if possible) the Bayes-optimal estimators of θ_i^0 and v_j^0, given the answers L_{ij}.
- Evaluate asymptotically the performance achieved by the Bayes-optimal estimators.

Our contribution, presented in this chapter, is to realize that the dDS model falls into a class of low-rank matrix estimation problems. For this class, the posterior probability distribution can be evaluated, as shown in [LKZ17], which resolves the first point. Using these results, the phase diagram can be evaluated in great detail, which treats the second point.

5.0.1 Approximate Message Passing

The approximate message passing (AMP) algorithm for low-rank matrix estimation is a simplification of belief propagation in the limit of dense graphical models and was derived in Sec. 1.3.5. For low-rank matrix factorization was first derived for special cases in [RF12; MT13] and in its general form in [LKZ15; LKZ17].

The AMP equations, stated in (1.97), can be further simplified in the Bayes-optimal setting. The coefficients of the Onsager reaction terms, can be replaced by their averages

$$\frac{1}{N}\sum_{k=1}^{M} S_{ik}^2 \sigma_{v,\,k}^t \to \alpha/\Delta\sigma_v \tag{5.7}$$

$$\frac{1}{N}\sum_{l=1}^{N} S_{lk}^2 \sigma_{\theta,\,l}^t \to \sigma_\theta/\Delta\,. \tag{5.8}$$

Further more, the terms involving R_{ij} can be set to zero by virtue of the Bayes-optimality, cf. (1.101). The resulting AMP equations are outlined in Alg. 1.

The input/denoising functions, $f_\theta(A_\theta, B_\theta)$ and $f_v(A_v, B_v)$, were introduced in (1.96). They depend on the priors, P_θ and P_v respectively and A and B are estimates for the parameters of a Gaussian distribution that are computed self-consistently in the AMP equations. The estimate $\hat{x}$ (with $x \in \{\{\theta_i\}_{i=1,\dots,N}, \{v_j\}_{j=1,\dots,M}\}$) are then computed as the mean of the prior weighted with this effective Gaussian. The estimates for their variance are obtained from the derivative w.r.t. B,

$$\begin{aligned} \hat{x} &= f_x(A_x, B_x) = \frac{1}{Z_x(A_x, B_x)} \int \mathrm{d}x\, x\, P_x(x)\, \mathrm{e}^{-\frac{1}{2}A_x x^2 + B_x x}\,, \\ \sigma_x &= \partial_{B_x} f_x(A_x, B_x)\,. \end{aligned} \tag{5.9}$$

We further recall the definition of the Fisher score matrix

$$S_{ij} \coloneqq \left.\frac{\partial g(L_{ij}, w_{ij})}{\partial w_{ij}}\right|_{w_{ij}=0} = L_{ij}\cdot\sqrt{\nu}\,, \tag{5.10}$$

where $g(L_{ij}, w_{ij})$ is defined in (5.3) and lastly, the definition of the Fisher information (inverse effective noise) of the noisy observation channel

$$\Delta^{-1} = \mathbb{E}_{P(L_{ij}|w_{ij}=0)}\left[\left(\left.\frac{\partial g(L_{ij}, w_{ij})}{\partial w_{ij}}\right|_{w_{ij}=0}\right)^2\right] = (1-\rho)\nu\,. \tag{5.11}$$

Algorithm 1: Approximate message passing for crowd sourcing. The numerical implementation might profit from an adequate damping scheme in order to enhance convergence even on small instances or when the model assumptions are not satisfied.

Data: $\boldsymbol{S}$, Δ, δ ; `// S and Δ according to (5.10) and (5.11) respectively.`

Result: MMSE estimates $\hat{\boldsymbol{v}}$ and $\hat{\boldsymbol{\theta}}$

Initialize: $\hat{\boldsymbol{v}} \leftarrow \hat{\boldsymbol{v}}^{\text{init}} \sim P_v(\boldsymbol{v})$, $\hat{\boldsymbol{\theta}} \leftarrow \hat{\boldsymbol{\theta}}^{\text{init}} \sim P_\theta(\boldsymbol{\theta})$; $\sigma_v \leftarrow 1$, $\sigma_\theta \leftarrow 1$; $\hat{\boldsymbol{v}}^{\text{old}} \leftarrow \mathbf{0}$, $\hat{\boldsymbol{\theta}}^{\text{old}} \leftarrow \mathbf{0}$;

while $\|\hat{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}^{old}\|_2^2 + \|\hat{\boldsymbol{v}} - \hat{\boldsymbol{v}}^{old}\|_2^2 > \delta$ **do**

$\boldsymbol{B}_\theta \leftarrow \frac{1}{\sqrt{N}} \boldsymbol{S}\hat{\boldsymbol{v}} - \frac{1}{\Delta} \hat{\boldsymbol{\theta}}^{\text{old}} \sigma_v$;

$\boldsymbol{A}_\theta \leftarrow \frac{1}{N\Delta} \hat{\boldsymbol{v}}^T \hat{\boldsymbol{v}}$;

$\boldsymbol{B}_v \leftarrow \frac{1}{\sqrt{N}} \boldsymbol{S}^T \hat{\boldsymbol{\theta}} - \frac{\alpha}{\Delta} \hat{\boldsymbol{v}}^{\text{old}} \sigma_\theta$;

$\boldsymbol{A}_v \leftarrow \frac{1}{N\Delta} \hat{\boldsymbol{\theta}}^T \hat{\boldsymbol{\theta}}$;

$\hat{\boldsymbol{\theta}}^{\text{old}} \leftarrow \hat{\boldsymbol{\theta}}$, $\hat{\boldsymbol{v}}^{\text{old}} \leftarrow \hat{\boldsymbol{v}}$;

$\hat{\boldsymbol{\theta}} \leftarrow f_\theta(\boldsymbol{A}_\theta, \boldsymbol{B}_\theta)$, $\sigma_\theta \leftarrow \frac{1}{N} \sum_{1 \le i \le N} \partial_{B_{\theta_i}} f_\theta(A_{\theta_i}, B_{\theta_i})$;

$\hat{\boldsymbol{v}} \leftarrow f_v(\boldsymbol{A}_v, \boldsymbol{B}_v)$, $\sigma_v \leftarrow \frac{1}{M} \sum_{1 \le j \le M} \partial_{B_{v_j}} f_v(A_{v_j}, B_{v_j})$;

end

5.0.2 State Evolution

The AMP algorithm depends on the realization of the disorder $\boldsymbol{L}$ and, consequently, so do the AMP estimates $\hat{\boldsymbol{\theta}}$, $\hat{\boldsymbol{v}}$ for the reliabilities and task labels. In the large size limit $N \to \infty$, the performance of the algorithm can be tracked with high probability by the state evolution (SE) equations, The SE for the dDS model was previously derived in Sec. 1.3.6. There it was shown that, in the Bayes-optimal setting, the overlap of the AMP estimates with the true solution can be quantified in terms of the two order parameters

$$\begin{aligned} M_\theta^t &= \frac{1}{N} \sum_{1 \le i \le N} \hat{\theta}_i^t \, \theta_i^0 \,, \\ M_v^t &= \frac{1}{M} \sum_{1 \le j \le M} \hat{v}_j^t \, v_j^0 \,. \end{aligned} \tag{5.12}$$

Where x^0 indicates the true value of x, and t the iteration step of the AMP equations (Alg. 1).

The SE equations imply that these order parameters evolve with high probability as

$$\begin{aligned} M_v^{t+1} &= \mathbb{E}_{v^0,W} \left[f_v \left(\frac{M_\theta^t}{\Delta} \,, \, \frac{M_\theta^t}{\Delta} \, v^0 + \sqrt{\frac{M_\theta^t}{\Delta}} \, W \right) \, v^0 \right] \,, \\ M_\theta^t &= \mathbb{E}_{\theta^0,W} \left[f_\theta \left(\frac{\alpha \, M_v^t}{\Delta} \,, \, \frac{\alpha \, M_v^t}{\Delta} \, \theta^0 + \sqrt{\frac{\alpha \, M_v^t}{\Delta}} \, W \right) \, \theta^0 \right] \,. \end{aligned} \tag{5.13}$$

Here W is an effective Gaussian random variable of zero mean and unit variance, $v^0 \sim P_v$, $\theta^0 \sim P_\theta$, the functions f_v and f_θ are defined in (5.9), $\alpha = M/N$ and Δ is the effective noise (5.11).

Let us denote with $M_\theta^{\rm SE}$ and $M_v^{\rm SE}$ the fixed points of the SE equations (5.13). These fixed points are then associated to the MSE and BE through

$$\mathrm{MSE}_\theta^{\rm AMP} = \mathbb{E}_\theta(\theta^2) - M_\theta^{\rm SE}\,, \tag{5.14}$$

$$\mathrm{BE}_v^{\rm AMP} = (1 - O_v^{\rm SE})/2\,, \tag{5.15}$$

where we introduced the order parameter

$$O_v^t = 1/M \sum_i \mathrm{sign}(\hat{v}_i^t)\, v_i^0\,,$$

such that

$$O_v^{\rm SE} = \mathbb{E}_{v^0,W}\left\{ \mathrm{sign}\left[f_v\left(\frac{M_\theta^{\rm SE}}{\Delta}\,,\, \frac{M_\theta^{\rm SE}}{\Delta}\, v^0 + \sqrt{\frac{M_\theta^{\rm SE}}{\Delta}}\, W \right)\right]\, v^0 \right\}\,. \tag{5.16}$$

5.0.3 Bayes-Optimal Error and Sub-Optimality of Message Passing Algorithms

As conjectured in [LKZ17] and proven rigorously in [Mio17] the performance of the Bayes-optimal estimator (5.6) can be evaluated in the large size limit $N \to \infty$ with $\alpha = \Theta(1)$ from the global minimizer of the replica symmetric Bethe free energy, which reads (cf. Box 2.3)

$$\begin{aligned}\phi_{\rm Bethe}(M_\theta, M_v) = \alpha \frac{M_\theta M_v}{2\Delta} - \alpha \mathbb{E}_{v^0,W}\left[\log Z_v\left(\frac{M_\theta}{\Delta}\,,\, \frac{M_\theta}{\Delta}\, v^0 + \sqrt{\frac{M_\theta}{\Delta}}\, W \right)\right] \\ - \mathbb{E}_{\theta^0,W}\left[\log Z_\theta\left(\frac{\alpha\, M_v}{\Delta}\,,\, \frac{\alpha\, M_v}{\Delta}\, \theta^0 + \sqrt{\frac{\alpha\, M_v}{\Delta}}\, W \right)\right]\,.\end{aligned} \tag{5.17}$$

where the functions Z_θ and Z_v are defined in (5.9) and the rest of the variables are defined in the same way as in the SE. Assume M_θ^* and M_v^* are the global minimizers of the above Bethe free energy. Then the minimum mean square error (MMSE) and the minimum bitwise error (MBE) are expressed as

$$\mathrm{MMSE}_\theta = \mathbb{E}_\theta(\theta^2) - M_\theta^*\,, \tag{5.18}$$

$$\mathrm{MBE}_v = \frac{1}{2}(1 - O_v^*)\,, \tag{5.19}$$

where O_v^* is obtained from M_θ^* via (5.16).

It is straightforward to observe that the SE equations are in fact stationarity conditions of the Bethe free energy. Hence the fixed points of the SE are critical points of the Bethe free

energy. Whether or not the SE reaches the global minimizer $M_\theta^*,\ M_v^*$ depends on the shape of the Bethe free energy and the initialization of the SE equations at $t = 0$. Canonically, the SE is initialized in such a way that the initial estimators are simply taken from the prior distributions.

We can now explain the key point of our contribution. The previous work of [OOSY16b; OOSY16a] proved asymptotic optimality of BP under certain assumptions on the parameters of the model. The present analysis of the dDS model is able to determine sharply in what regions of parameters AMP matches the Bayes-optimal estimator and when it does not, thus refining the previous picture in the limit, where AMP and BP are asymptotically equivalent.

Previously we reduced the high-dimensional model to the investigation of the two-variable free energy function (5.17). In particular, the phases in which AMP does not match the Bayes-optimal estimator can be characterized in terms of the critical points of the free energy and whether or not the SE (5.13) converges to the global minimum of the free energy (5.17). The way we check this in practice is that we initialize the SE in two different ways:

- Uninformative initialization, where $M_v^{t=0} = (\mathbb{E}_v(v))^2$ and $M_\theta^{t=0} = (\mathbb{E}_\theta(\theta))^2$. This corresponds to the uninformative initialization of the algorithm where the initial values of the estimators are simply taken equal to the mean of the prior distributions P_θ and P_v. The error achieved by the AMP algorithm is then given by iteration of (5.13) from this uninformative initialization.

- Informative initialization, where $M_v^{t=0} = \mathbb{E}_v(v^2)$ and $M_\theta^{t=0} = \mathbb{E}_\theta(\theta^2)$ so that the initial mean-squared-errors are zero. This is not possible within the algorithm without the knowledge of the ground truth and it is purely used for the purpose of the analysis. If the iteration of the SE equations (5.13) from this informative initialization leads to a different fixed point than from the uninformative initialization, then the free energies of the two fixed points need to be compared and the larger one surely does not correspond to the Bayes-optimal performance.

This procedure is sufficient, provided there are no other fixed points. If there are, the free energy of all of them needs to be compared.

Zero-mean priors and uninformative fixed point If both prior distributions P_θ and P_v have zero mean, the uninformative initialization $M_\theta = M_v = 0$ is a fixed point of the SE and equations (5.13) can be expanded around this fixed point. In first order we obtain

$$M_\theta^t = \frac{\alpha}{\Delta}\left(\mathbb{E}_\theta\left[\theta^2\right]\right)^2 M_v^t \tag{5.20}$$

$$M_v^{t+1} = \frac{1}{\Delta}\left(\mathbb{E}_v\left[v^2\right]\right)^2 M_\theta^t\,, \tag{5.21}$$

implying that the uninformative fixed point is numerically stable for

$$\Delta^2 > \alpha\left(\mathbb{E}_v\left[v^2\right]\right)^2\left(\mathbb{E}_\theta\left[\theta^2\right]\right)^2$$

and unstable otherwise. Therefore we define the critical effective noise, Δ_{c}, as

$$\Delta_{\mathrm{c}} = \sqrt{\alpha} \cdot \mathbb{E}_{\theta}\left[\theta^2\right] \mathbb{E}_{v}\left[v^2\right] . \tag{5.22}$$

For $\Delta < \Delta_{\mathrm{c}}$ the uninformative initialization becomes numerically unstable. The threshold Δ_{c} correspond to the 2nd order phase transition in the behavior of the AMP algorithm, meaning that the overlap reached by the algorithm is non-analytic and continuous at Δ_{c}.

In the case where both the priors, P_θ and P_v, have zero mean, we can divide the region of parameters into the following three phases:

- **Easy phase:** The free energy (5.17) has a unique minimum and this minimum is associated with a positive overlap with the ground-truth configuration. Consequently iterating the state evolution (5.13) yields an informative fixed point from both, the informative, as well as the (perturbed) uninformative initializations. AMP is Bayes-optimal.

- **Hard phase:** In this phase at least two minima of the free energy (5.17) coexist; at least one local minimum of small overlap and a global minimum of larger overlap. The outcome of iterating the state evolution equations now depends on the initialization: while the informative initialization yields a fixed point with large overlap, the uninformative initialization leads to a fixed point of low overlap. This is precisely the region of parameters where the AMP algorithms do not reach the information-theoretically optimal performance and *AMP is not Bayes-optimal*.

- **Impossible phase:** When the global minimum of (5.17) is associated to the trivial, non-informative, fixed point corresponding to zero overlap, we talk about a phase of impossible inference. Otherwise this region is indeed similar to the easy phase in the sense that AMP is Bayes-optimal.

If at least one of the priors has non-zero mean, then the distinction of an impossible phase is not meaningful and one would only have the easy and hard phases, the later is defined by asymptotic sub-optimality of the AMP algorithm.

Let us further define the following three thresholds that are associated with the existence of a hard phase. The hard phase is always linked to the presence of a first order phase transition, i.e., a discontinuity in the asymptotic value of the overlap reached by the Bayes-optimal estimator. The *algorithmic threshold* Δ_{alg} is the largest value of effective noise, Δ, below which the AMP algorithm asymptotically matches the Bayes-optimal performance. The *spinodal threshold*, Δ_{sp}, is the smallest values of effective noise above which the informative initialization converges to a different fixed point than the (perturbed) uninformative initialization. The *information theoretic transition*, $\Delta_{\mathrm{alg}} < \Delta_{\mathrm{IT}} < \Delta_{\mathrm{sp}}$, is where the value of the Bethe free energy of the fixed point reached from the uninformative initialization crosses with the free energy of the fixed point reached from the informative initialization. The discontinuity in overlap happens at Δ_{IT}. Remark that while in some models, such as the stochastic block model [LKZ15], we find $\Delta_{\mathrm{c}} = \Delta_{\mathrm{alg}}$, in general and in the present model $\Delta_{\mathrm{c}} \neq \Delta_{\mathrm{alg}}$.

5.1 Phase Diagrams for the Dense Dawid-Skene Model

A key property of the results we described so far is that the asymptotic behavior of the AMP algorithm and of the Bayes-optimal estimator depend only on the priors P_v, P_θ and the effective noise $\Delta = 1/[(1-\rho)\nu]$. In what follows concrete priors will be considered. It is assumed that the ground truth task labels are generated from

$$P_v(v_j) = (1-\beta)\delta\left(v_j - 1\right) + \beta\delta\left(v_j + 1\right) . \tag{5.23}$$

With the parameter $\beta \in [0,1]$ accounting for a bias in the dataset.

We start our discussion with worker reliabilities, θ_i, drawn from a skewed Rademacher-Bernoulli (RB) prior

$$P_\theta(\theta) = (1-\mu)\,\delta(\theta) + \mu\left[(1-\lambda)\delta(\theta-1) + \lambda\delta(\theta+1)\right] . \tag{5.24}$$

Besides its simplicity the phase diagram for this case comprises the essential features. Tuning μ from zero to one interpolates between an uninformative crowd of mere spammers and an informative crowd. The fraction of adversaries is controlled by λ. In physics terms the workers with $\theta = -1$ are spins that are coupled to the questions by an anti-ferromagnetic interaction, whereas the workers with $\theta = 1$ are ferromagnetically coupled. Consequently also the adversaries enhance our ability to recover the correct labels, if they can be identified, as they align anti-parallel to the truth.

The RB prior is the dense version of what is sometimes referred to as the "spammer-hammer" model in the literature [KOS11]: workers are either spammers that provide random answers or hammers that align very strongly with (or opposed to) the truth. Here the situation is slightly different as we assume a very weak alignment of $\Theta(1/\sqrt{N})$, cf. (5.1). Sending $\nu \to \infty$, and thus $\Delta \to 0$, approximates the hammers in the dDS model. The limit $\nu \to N$ will be considered in Sec. 5.2.

5.1.1 The case of symmetric priors

If $\lambda = 1/2$ and $\beta = 1/2$ both the priors P_v and P_θ have zero mean and the SE equations in (5.13) have a trivial fixed point at $M_v^* = M_\theta^* = 0$. Expansion around this fixed point yields

$$M_v^{t+1} = \alpha\frac{\mu^2}{\Delta^2}\cdot M_v^t - \alpha^2\frac{\mu^2}{\Delta^2}\left[\frac{\mu}{\Delta} + \frac{\mu^2}{\Delta^2}\right]\cdot\left(M_v^t\right)^2 + O((M_v^t)^3)\,. \tag{5.25}$$

The linear term gives the stability criterion of the trivial fixed point that we already derived in (5.22)

$$\Delta_{\mathrm{c}} = \sqrt{\alpha}\cdot\mu\,. \tag{5.26}$$

In Fig. 5.1 we present the phase diagram for several values of $\alpha = M/N$. We plot the stability threshold Δ_{c} as well as the three phase transitions associated with the existence of the hard phase. We mark the phases where inference is algorithmically easy, hard and

impossible. In particular, we find that a *hard phase appears* for small enough μ as depicted in the figure. Regions with small μ correspond to crowds that contain mostly spammers. For $\alpha = 1$ the hard phase appears only if the vast majority of the workers are spammers. When α grows (shrinks) the hard region grows (shrinks) as well. In the region where the hard phase is absent (5.26) provides the right criterion to locate the phase transition from the easy to the impossible phase.

5.1.2 Biased labels and worker reliabilities

If $\lambda \neq 1/2$ or $\beta \neq 1/2$ the trivial fixed point $M_v = M_\theta = 0$ does not exist anymore. We illustrate in Fig. 5.2 how this changes the phase diagram and the achievable MSE. For the case $\alpha = 1$ and $\mu = 0.02$ we plot the MSE reached by the SE from the informative and the uninformative initialization.

First (left top panel), we consider the unbiased case with $\beta = 1/2$, but $\lambda \neq 1/2$ as already plotted in Fig. 5.1. In the bottom-left panel we consider the case where λ changes. Due to the present symmetry it suffices to restrict the attention to $\lambda > 1/2$. When more hammers than adversaries are present, i.e. for $\lambda > 1/2$, the trivial fixed point at $M_v = 0$ disappears and instead another fixed point with low but positive overlap (i.e. error smaller than 1) appears. The hard phase shrinks as shown in the bottom-left panel of Fig. 5.2.

If the dataset is biased, i.e. $\beta \neq 1/2$, the change is quantitatively more dramatic, but phenomenologically very similar, cf. top-right panel in Fig. 5.2. Upon slight change in β the hard phase shrinks considerably. For a large range of values of β and λ the hard phase entirely disappears as in the bottom-right panel in Fig. 5.2.

5.1.3 The impact of α

Recall that $\alpha = M/N$ is the ratio of tasks to workers in our model. By virtue of the $\sqrt{\nu/N}$ scaling of the signal, cf. (5.1), we have two competing mechanisms when N is increased: on the one hand the signal becomes weaker, on the other hand we obtain more answers per question. Equation (5.26) tells us that it should be expected that inference becomes easier when α increases. If Δ is fixed and the performance changes with α is considered, it follows from the SE that it is necessary to increase the fraction of questions distributed to each worker, i.e. by increasing α, in order to achieve higher overlap. This improves the estimation of $\boldsymbol{\theta}$, which in turn improves the estimate of $\boldsymbol{v}$. We depict this by plotting the error rate against α for two different values in Fig. 5.3.

How does the hard phase vary with α? We answer this question in Fig. 5.1b where it is shown that the hard phase grows further in the impossible phase when α is increased, while it shrinks when α is decreased.

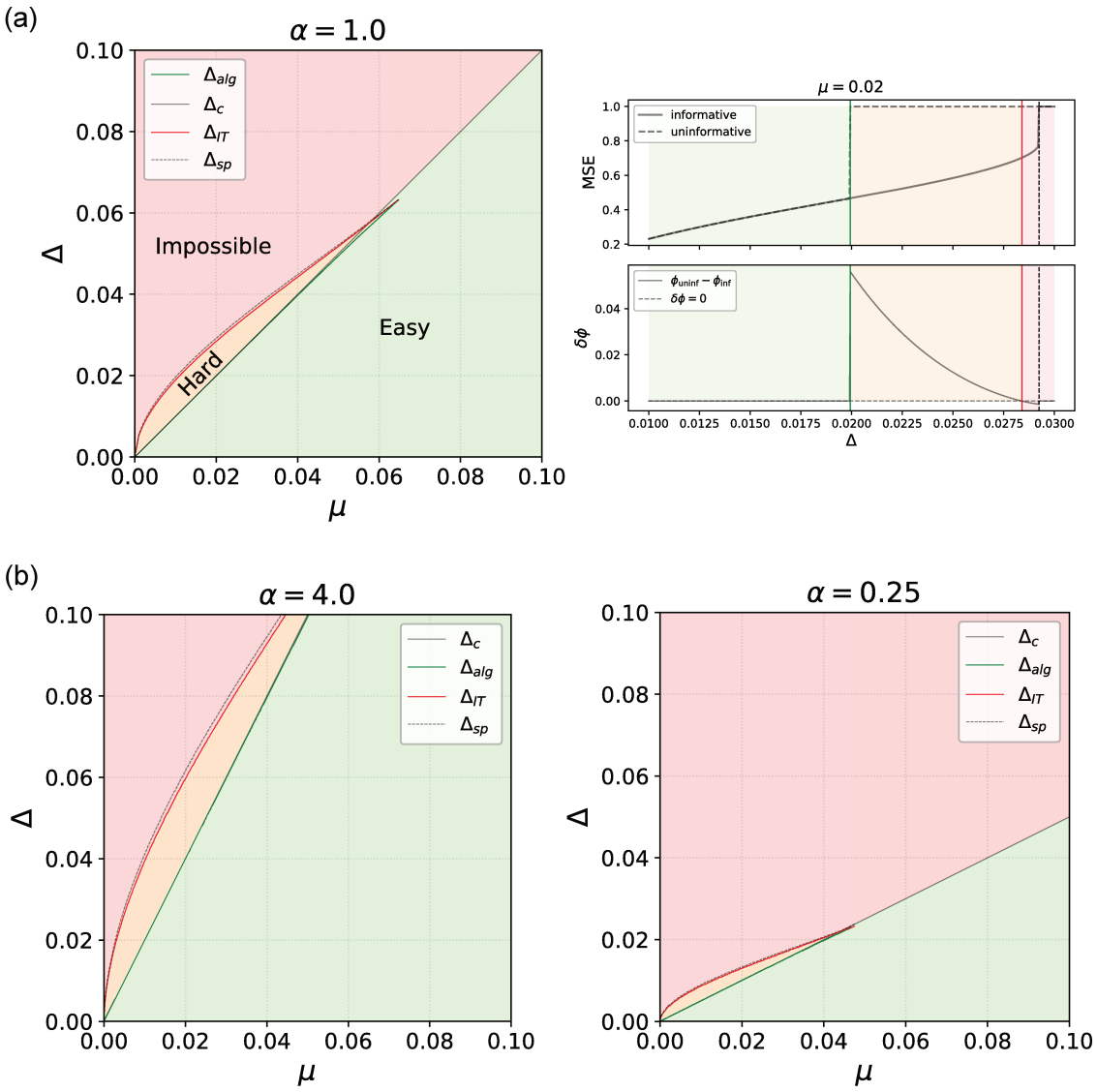

Figure 5.1: (a) left panel: The phase diagram for a Rademacher-Bernoulli prior on θ with $\lambda = 1/2$ and no bias in the distribution of the labels ($\beta = 1/2$). When the fraction of spammers is very large (small μ) a hard phase appears where the AMP algorithm is not able to reach the information-theoretically optimal performance. (a) right panel: cut of the phase diagram corresponding to $\mu = 0.02$, i.e. only 2% of non-spammers. We plot the MSE (top) and the difference in the free energy (bottom) of the two fixed points as a function of Δ. Note that in this case we still have $\Delta_{\rm c} > \Delta_{\rm alg}$ but both are indistinguishably close. In the hard region (orange) the AMP algorithm reaches $\mathrm{MSE} = 1$ while the Bayes-optimal estimator reaches the depicted MSE. (b) Phase diagrams with all parameters set to the same values, but α different. When α grows (shrinks) inference becomes easier (harder) and the hard region grows (shrinks). The tricritical point for $\alpha = 1/4$ is located around $\mu \approx 0.048$ whereas for $\alpha = 4$ it is around $\mu \approx 0.077$.

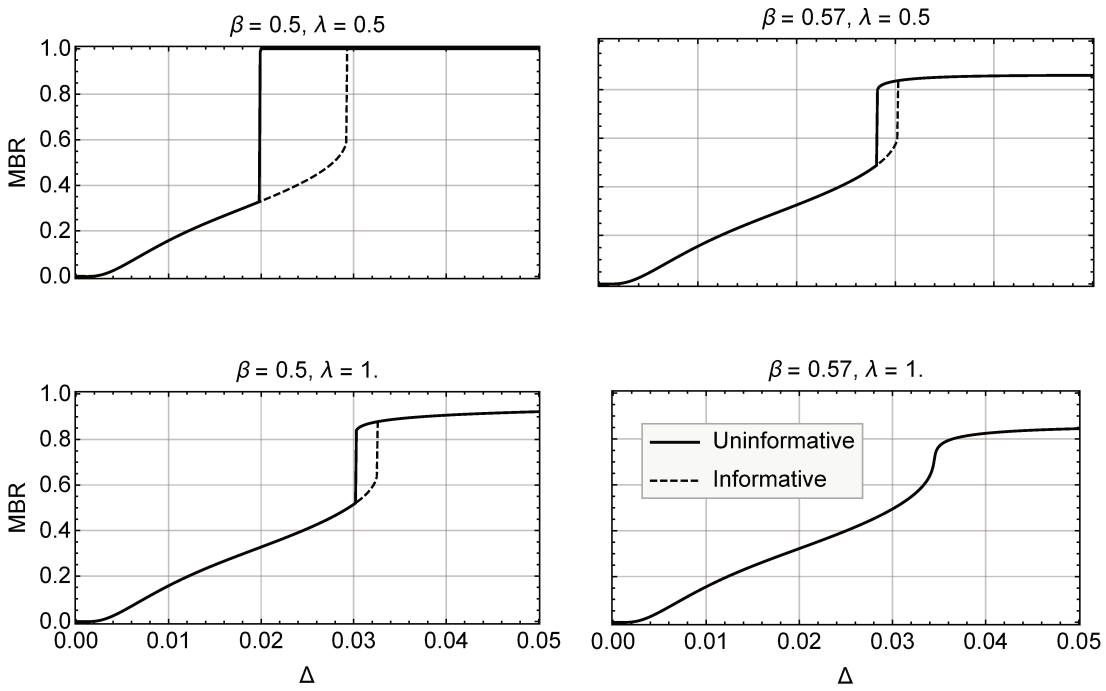

Figure 5.2: Influence of bias in the distribution of labels and worker reliabilities on the performance. Here we plot the resulting bitwise error (5.15) for $\alpha = 1$ and $\mu = 0.02$ as reached from the uninformative (bold) and informative (dashed) initialization. For bias in the labels ($\beta \neq 0.5$) or in the workers abilities ($\lambda \neq 0.5$) the trivial fixed point (error equal to one) is replaced by another fixed point with slightly lower error. The hard phase in these examples appears at larger noise and shrinks or might disappear as in the bottom right panel.

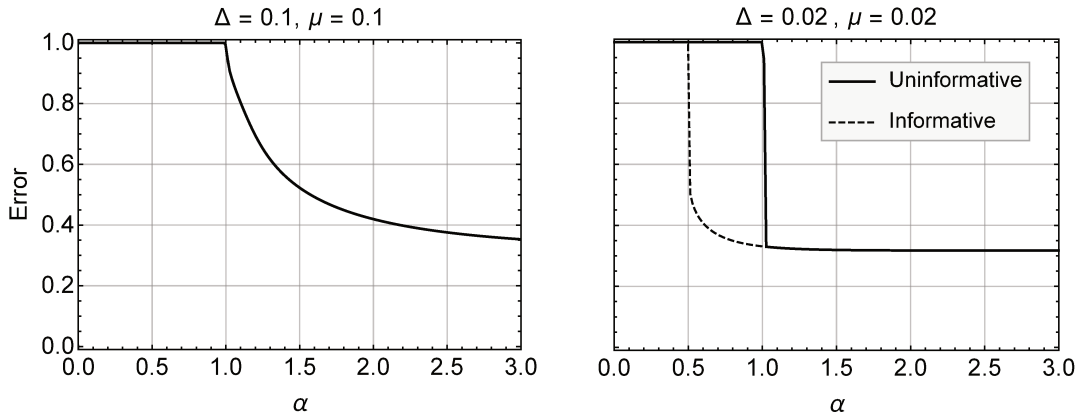

Figure 5.3: The behavior of the bitwise error versus α for the RB prior (5.24) with $\lambda = 1/2$ and bias is set to $\beta = 1/2$.

5.1.4 Dealing with other priors

The dDS model applies to any prior as long as $\theta = O(1)$. Many features persist if (5.24) is replaced by

$$P_\theta(\theta) = (1-\mu)\delta(\theta) + \mu\phi(\theta)\,.$$

with $\phi(\theta)$ some appropriate distribution (we have considered $\phi(\theta)$ being a beta distribution or a Gaussian). For instance (5.25) still holds when $\phi(\theta)$ is a standard Gaussian and as for the RB prior a first order transition is triggered by very noisy θ, i.e., only very few hammers and mostly spammers in the crowd.

One might also replace the delta distribution by some other sparsity inducing distribution. A case for which the corresponding integrals are tractable analytically is that of a mixture of two Gaußians, centered around $\bar{\theta}_L$ ($\bar{\theta}_R$) with variance σ_L^2 (σ_R^2).

$$P_\theta(\theta) = (1-\mu)\,\mathcal{N}(\theta;\bar{\theta}_L,\sigma_L^2) + \mu\,\mathcal{N}(\theta;\bar{\theta}_R,\sigma_R^2)\,.$$

Under this choice and with $\beta = 1/2$ in (5.23) the SE equations (5.13) can be expressed as

$$M_v^{t+1} = G\left(\frac{1}{\Delta}T\left(\frac{\alpha}{\Delta}M_v^t\right)\right) \tag{5.27}$$

with

$$\begin{aligned}
G(x) &= \mathbb{E}_W\left\{\tanh(x+\sqrt{x}W) - \tanh(-x+\sqrt{x}W)\right\}\\
T(q) &= \mu\cdot\mathbb{E}_W\left\{\frac{\left[F_1(W) + \frac{1-\mu}{\mu}F_2(W)\cdot\exp\left(-\frac{1}{2}Q(W)\right)\right]^2}{1+\frac{1-\mu}{\mu}\sqrt{\frac{1+q\sigma_R^2}{1+q\sigma_L^2}}\exp\left(-\frac{1}{2}Q(W)\right)}\right\}
\end{aligned}$$

where

$$\begin{aligned}
Q(W) &= \frac{1+q\sigma_R^2}{1+q\sigma_L^2}\left(W + \sqrt{\frac{q}{1+q\sigma_R^2}}(\bar{\theta}_R - \bar{\theta}_L)\right) - W^2\\
F_1(W) &= \bar{\theta}_R + \sqrt{\frac{q}{1+q\sigma_R^2}}\sigma_L^2 W\\
F_2(W) &= \left(\frac{1+q\sigma_R^2}{1+q\sigma_L^2}\right)^{\frac{3}{2}}\left(\frac{\bar{\theta}_L + q\sigma_L^2\bar{\theta}_R}{1+q\sigma_R^2} + \sqrt{\frac{q}{1+q\sigma_R^2}}W\right)\,,
\end{aligned}$$

and $\mathbb{E}_W$ indicates the average over the standard Gaussian measure on W. Varying the means ($\bar{\theta}_L$, $\bar{\theta}_R$) and variances (σ_L^2, σ_R^2) then allows to interpolate between different scenarios.

5.2 Relevance of the results in the sparse regime

Our analysis of the dense DS model is based on the ground that the underlying graphical model (the bipartite question-worker-graph) is densely connected. That is, each task-node

is connected to $\Theta(N)$ worker-nodes (and reversely each worker-node is connected to $\Theta(M)$ task-nodes). We introduced a sense of sparsity in the channel, by allowing that some of the tasks remained unanswered, cf. (5.1). However, the analysis assumed that $1-\rho = \Theta(1)$. Existing mathematical literature on low-rank matrix estimation shows that the formulas that we derived for the Bayes-optimal performance, hold true even when the degrees in the graph grow with N slower than linearly, i.e., when $(1-\rho)N$ diverges with $N \to \infty$ [DAM16; CLM17]. The regime where the above asymptotic results do not hold anymore is when $1-\rho = O(1/N)$, which we refer to as the sparse regime. In this section we investigate numerically how the behavior of the sparse DS model deviates from the predictions drawn from the dense DS model.

In the sparse regime considered here every worker is connected to d randomly chosen tasks, where $d = \Theta(1)$. Unless the quality of each answer is very high, the effective noise $\Delta = [(1-\rho)\nu]^{-1}$ is overwhelming and inference impossible, unless $\nu = \Theta(N)$. Therefore we will consider the following "mapping"

$$\rho = 1 - \frac{d}{M} \qquad \nu = n \cdot N \,, \tag{5.28}$$

with $n \in [0,1]$ being a constant. Consequently in the sparse regime we are dealing with high quality workers as compared to the dense regime. This brings us close to the setting of previous literature on the DS model [KOS11; LPI12; OOSY16b; OOSY16a].

5.2.1 Approximate message passing on sparse graphs

We studied numerically how the AMP algorithm behaves when the average degree of the nodes is small. In the following we will set $M = N$ such that the average degree of the task-nodes equals the average degree, d, of the worker-nodes.

Fig. 5.4a depicts results that were obtained by running AMP in the dense regime, where $d = \Theta(N)$, for a system with 10^4 nodes. Except from finite size effects close to the phase transition, the SE prediction agrees with the empirical results. For Fig. 5.4b we fixed different values of Δ (by adjusting n so that $\Delta = \alpha/(nd)$) and plotted the relative deviation from the SE when the degree d is varied. We also show the results obtained with the BP algorithm of [LPI12] that are obtained by matching the prior and signal to noise ratio. In the limit of large N the BP results are exact even for finite d. We find as expected that when d is increased, the AMP performance approaches the prediction of the associated dense model and so does BP. While for very small d BP slightly outperforms AMP, the difference is not very significant (up to fluctuations).

We further quantified the difference in performance of BP and AMP in the sparse regime, cf. Fig. 5.5, where ν is fixed and d (and hence Δ) varies. We compared AMP with its BP equivalent and found that BP always outperforms AMP, but again only slightly. The general trend is as expected: in the sparse regime BP is optimal and no other algorithm can outperform it. However, it is remarkable how quickly AMP becomes comparable to BP. In

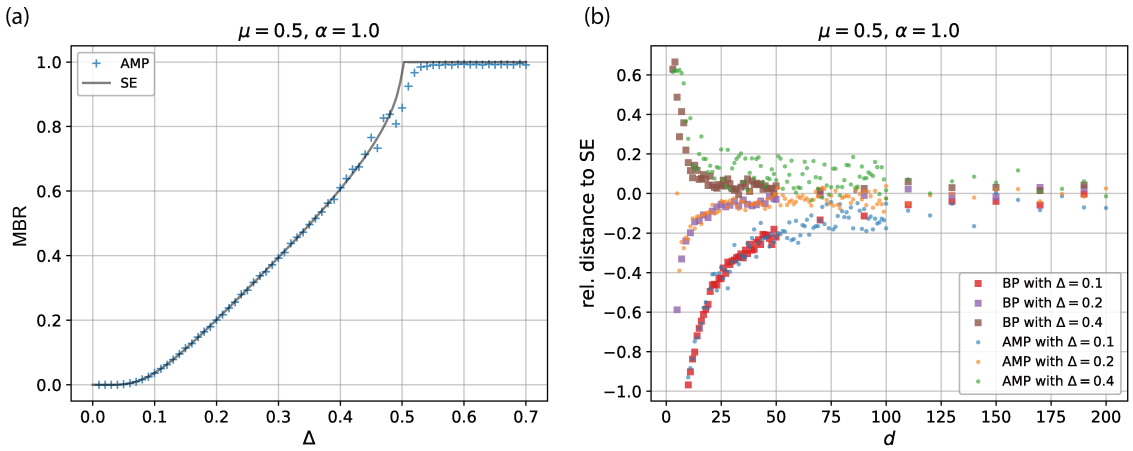

Figure 5.4: Evaluation of the AMP results in the sparse regime. (a) Comparison of the bitwise error as achieved by AMP in the dense regime for $N = M = 10^4$, averaged over 20 samples, with the SE. (b) The relative distance of the AMP results to the SE prediction of the error when the average degree, d, and signal to noise ratio, ν, are varied such that Δ remains fixed. We also compare to the BP algorithm, that is asymptotically exact in the sparse regime. The SE gives an accurate description, already for relatively small d, around $30 - 50$. Although AMP is suboptimal for low degrees and BP asymptotically optimal, AMP and BP give comparable results down to average degrees, d, around 10.

Fig. 5.5b we fix d and vary ν, such that Δ varies in the same range as in Fig. 5.5a. It is not possible to explore the full range of Δ because is is necessary to restrict $n \leq 1$.

The results clearly suggest that (for finite size systems) AMP can indeed be run even on relatively sparse instances. Compared to BP it is algorithmically less complex and more memory efficient, as fewer messages need to be stored. Further, the SE prediction seems to remain a good qualitative approximation to the algorithmic performance. It suggests that the phenomenology found in the dense limit should be rather generic and also appear in sparse systems.

5.2.2 First order phase transition in belief propagation

So far it was shown that the dense DS model can exhibit both, second and first order phase transitions. The first order transitions are more interesting algorithmically as they are associated with the presence of an algorithmically *hard* region where the corresponding message passing algorithm is suboptimal.

The authors of [OOSY16b; OOSY16a] established that BP is optimal in the sparse DS model for sufficiently large signal to noise ratio. It remains to be tested whether we can observe a first order phase transition also in the sparse version of the model. This shall be the aim of the present section.

Suboptimality of BP is associated with a region of parameters for which BP converges to different fixed points from the informative and from the uninformative initialization. We use

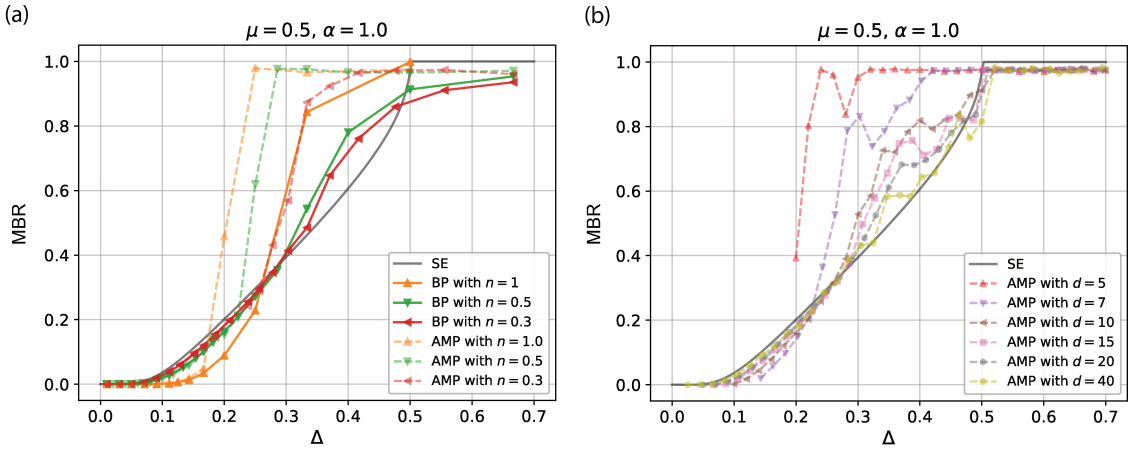

Figure 5.5: The effect of variation of either d or ν (i.e. n) on the performance of AMP. (a) Bitwise error of AMP for fixed ν. We also compared to the BP results that have the same prior and matching signal to noise ratio. (b) Bitwise error achieved by AMP for fixed d. The fact that the error found in the experiments for large Δ is slightly below the SE is due to finite size effects. Increasing the average degree pushes the results closer to the SE prediction. The experiments were carried out with $N = 10^3$ and are averaged over 100 samples.

our intuition based on the results of the dense case to show that there exist regimes where BP is sub-optimal. Fig. 5.6 depicts numerical results obtained for BP with a Bernoulli-prior on $\boldsymbol{\theta}$ ($\lambda = 0$ in (5.24)) with very sparse signals ($\mu = 0.01$). We also plot the AMP performance (in the same sparse regime) as well as the asymptotic prediction that would be expected in the dense case (5.28). Indeed, a clear first order transition appears. This establishes the suboptimality of BP by virtue of the dependency on the initialization.

5.2.3 Approximate message passing on real data

We tested the AMP algorithm (Alg. 1) on the bluebird dataset of Welinder et al. [WBPB10]. This dataset is fully connected, minimizing effects introduced by poorly designed task-worker-graphs. We used the same priors and parameters as in [LPI12] to compare AMP to other algorithms. Following [LPI12] we also implemented a "two-coin" extension of AMP that assumes that the true positive and true negative rates are different. We define $\vec{\theta}_i := (s_i, t_i)$ with s_i the sensitivity of worker i and t_i indicating its specificity. We have

$$P\left(Y_{ij} = \pm 1 \mid \vec{\theta}_i,\ v_j = +1\right) = (1-\rho) \cdot \frac{1}{2} \cdot \left(1 \pm \sqrt{\frac{\nu}{N}}\, s_i\right) \tag{5.29}$$

$$P\left(Y_{ij} = \pm 1 \mid \vec{\theta}_i,\ v_j = -1\right) = (1-\rho) \cdot \frac{1}{2} \cdot \left(1 \mp \sqrt{\frac{\nu}{N}}\, t_i\right) \tag{5.30}$$

$$P\left(Y_{ij} = 0 \mid \vec{\theta}_i,\ v_j\right) = \rho\,. \tag{5.31}$$

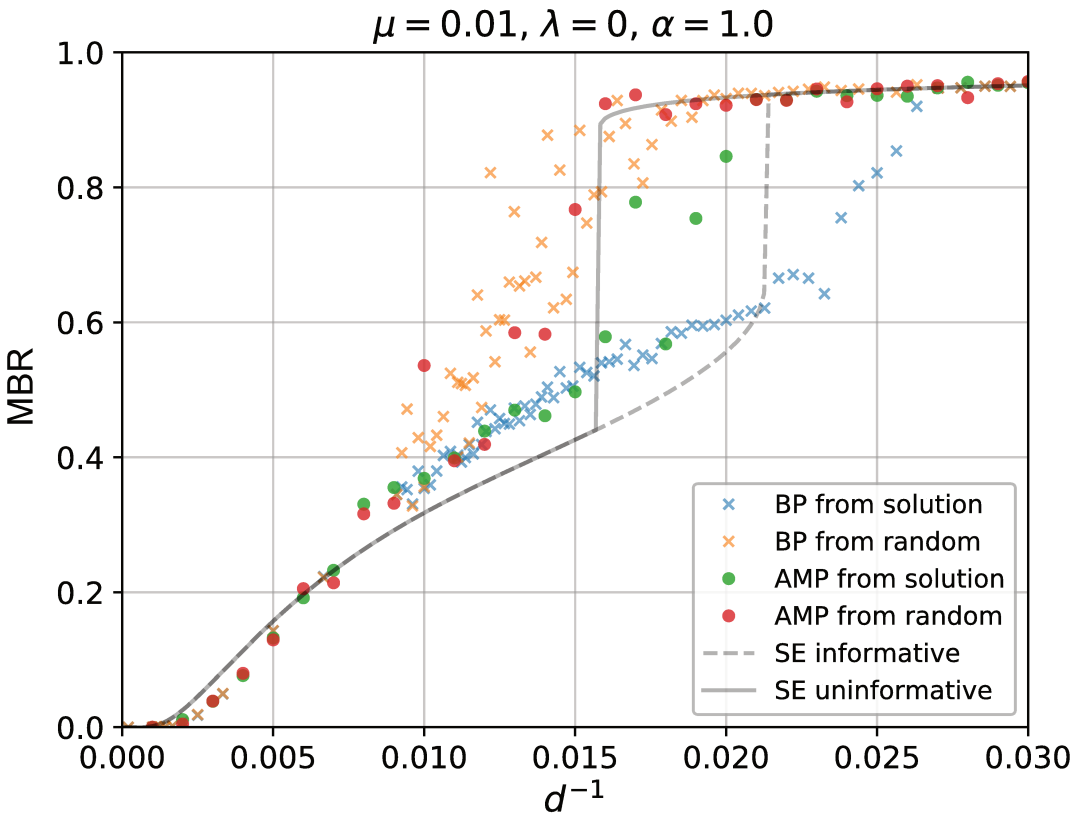

Figure 5.6: Numerical results of the achieved bitwise error obtained for the BP algorithm of references [LPI12; OOSY16b]. The experiments were carried out on graphs of size $N = M = 10^4$, and are reported as a function of the inverse average degree of the worker nodes, d. A region of coexistence associated to a first order phase transition opens up and an informative initialization leads to another fixed point than the uninformative one. This makes BP sub-optimal in the part of this region, where the free energy of the fixed point reached from the uninformative initialization is higher than the one of the fixed point reached from the informative initialization. We found in our experiments that the first order transition appears more pronounced the larger the system size, suggesting that the phenomenon persists asymptotically.

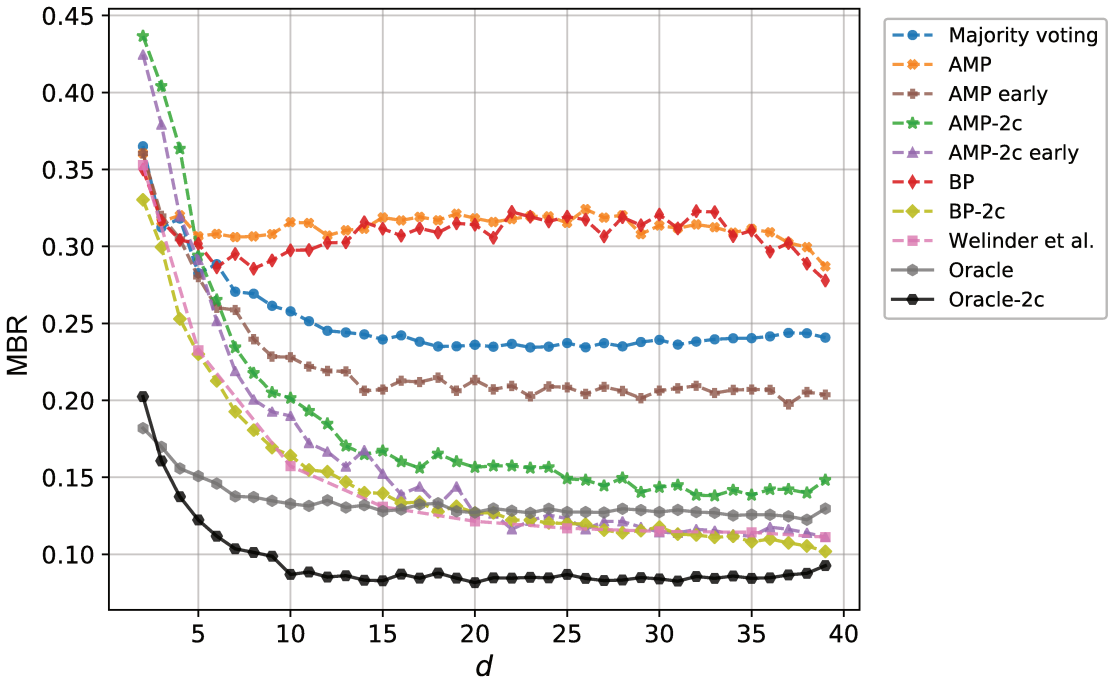

Figure 5.7: The bitwise error against the number of workers-per-task, d. Different algorithms are compared to AMP on the bluebird dataset. We compare to the results obtained with BP, majority voting and the algorithm proposed by Welinder et al. [WBPB10]. As explained in the text, we implemented two different version of AMP and BP: a symmetric one in which the sensitivity and specificity are equal and an asymmetric version (referred to as "2-c" in the legend). Finally we also plot results obtained when AMP is run with an early stopping criterion of 10 iterations. For BP and AMP the priors are set to independent $\mathrm{Beta}(2,1)$ distributions on $\boldsymbol{\theta}$. We averaged over 100 samples for each d.

The above model is cast into a rank-2 matrix factorization problem by setting

$$\vec{v_j} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \text{ if question } j \text{ is true and } \vec{v_j} = \begin{pmatrix} 0 \\ -1 \end{pmatrix} \text{ if question } j \text{ is false}.$$

The only difference is that the former rank-1 matrix $\boldsymbol{w}$, cf. (5.2), now becomes a rank-2 matrix with $\boldsymbol{\theta} \in \mathbb{R}^{M\times 2}$ and $\boldsymbol{v} \in \mathbb{R}^{N\times 2}$. The equations for a general rank were derived and given in detail in [LKZ17].

In Fig. 5.7 AMP is compared with BP, majority voting and the algorithm developed by Welinder et al. in [WBPB10]. We also compute the oracle lower bound of [KOS11] for the two versions of AMP and BP. To evaluate the oracle first estimate the true parameters $\boldsymbol{\theta}$ are estimated from the ground truth and then the resulting Bayes-optimal estimator that maximizes the posterior probability is computed. Note that the latter estimator has full information of the workers reliabilities.

We stress that analogous comparisons between existing algorithms and BP were already performed in [LPI12], where BP was found to be superior. Our main point in this section is that AMP, which is simpler than BP, gives a comparable performance to BP even on real-world data. We therefore focus on the comparison between BP and AMP. Both, BP and AMP perform badly when the original model with $s_i = t_i$ is used as can be seen from Fig. 5.7 by comparing them to majority voting as a baseline algorithm. Running the same experiments with the two-coin version improves the results significantly. Indeed BP and AMP perform essentially as well as the much more involved algorithm of [WBPB10].

The experiments were run with identical beta-priors for BP and AMP ($a = 2$, $b = 1$) for comparability with the results in [LPI12; OOSY16a]. For AMP different strategies were implemented for the prior on $\boldsymbol{v}$. Setting σ to the true value (estimated from the ground truth) or to $1/2$ led to comparable results as when it was learned. In our AMP implementation we initialize $\hat{\boldsymbol{v}}$ in the estimates obtained by majority voting.

In the symmetric case BP and AMP are very close in performance. The difference for the two-coin models tends to be slightly larger, while the general trend persists. We also observe that it can be beneficial to implement AMP with an early stopping criterion as depicted in Fig. 5.7. Early stopping can be reasonable because the assumptions made in the derivation are likely to be imprecise, especially for small system sizes.

In summary, AMP performs quite well on real world datasets. The vanilla implementation yields slightly worse results, as compared to BP. However, when AMP is stopped after few iterations (we used 10) it reaches much better performance in the rank-1 case. A significant improvement is also obtained in the rank-2 version of AMP: for small d BP outperforms AMP, but they soon become quasi indistinguishable. Our experiments on the real-world dataset also show that having a model that described data accurately is more important than the precise algorithm that is used to do do inference on the model. Besides its good performance it has the great advantage of algorithmic simplicity, better time complexity and scalability.

5.3 Conclusion

In this chapter the dDS model for crowdsourcing was considered. The mapping onto a larger class of low-rank matrix factorization problems leads to an approximate message passing algorithm for crowdsourcing and a closed-form asymptotic analysis of its performance. Due to the previous work of [BDM+16; Mio17] this analysis can be considered rigorous. While the theory only holds rigorously for the dense Dawid-Skene model, numerical experiments suggest that in the sparse regime AMP still performs well and also the asymptotic analysis provides a good qualitative prediction.

When the crowd consists mainly of spammers with only few workers that provide useful information, we found that a first order transition appears in the Bayes-optimal performance. Algorithmically this first order transition translates into the presence of a hard phase in which the AMP algorithm is sub-optimal. As a proof of concept we showed numerically that this

feature persists even in the sparse regime where the rigor of our analysis breaks down. In experiments we also found instances of first order transitions in the belief propagation algorithm of [LPI12]. This shows that there are regimes in the Dawid-Skene model where BP is not optimal. This complements recent results on [OOSY16b; OOSY16a] about regimes of optimality of BP.

We also carried out experiments on real-world data and showed that AMP performs comparable to other state-of-the-art algorithms, while being of lower time complexity.

— 6 —

Contagious Sets

Threshold models are a common approach to model collective dynamical processes on networks. A question of major importance concerns the minimal contagious sets in such dynamical processes. The work, presented in this chapter, is inspired by the generalization of the COREHD *algorithm, proposed in [ZZZ16]. This algorithm can be analyzed exactly and its analysis leads to new rigorous upper bounds on the size of the minimal contagious sets for random graphs with bounded degree, drawn from the configuration model.*

My main contribution is the derivation of the asymptotic analysis of the COREHD *algorithm. The correctness of the analysis was made rigorous in collaboration with Henry Pfister. I also propose some new heuristics that attack the problem algorithmically and outperform* COREHD*as well as other state-of-the-art algorithms.*

The rest of the chapter is divided into two parts. First, the COREHD *algorithm is analyzed. Subsequently, some new algorithms are proposed and compared to* COREHD *and other existing algorithms.*

In this chapter we study the problem of finding the smallest set of nodes in a network whose removal results in an empty k-core, the minimal contagious set. On random graphs the minimal contagious set problem undergoes a threshold phenomenon: in the limit of large graph size the fraction of nodes belonging to the minimal contagious set is w.h.p. concentrated around a critical threshold value. It is related to a widely studied model for dynamics on a network: the threshold model [Gra78], also known as bootstrapping percolation in physics [CLR79]. We refer to Sec. 1.2.3 for an introduction of the problem.

A very well performing algorithm for decycling and dismantling has been recently introduced in [ZZZ16]. In Alg. 2 the COREHD algorithm of [ZZZ16] is outlined (generalized from $k = 2$ to generic k). One of the main contributions of the present work is to analyze exactly the performance of this algorithm exactly, thus leading to upper bounds that are improving those of [BWZ02] and partly closing the gap between the best know algorithmic

upper bounds and the expected exact thresholds [GS15; BDSZ16].

Algorithm 2: Generalized COREHD algorithm.

Data: $\mathcal{G}(\mathcal{V}, \mathcal{E})$
Result: A set of nodes $\mathcal{D}$ whose removal makes the k-core vanish
Function: `core` $(k,\mathcal{G})$ returns the k-core of the graph $\mathcal{G}$

Initialize: $\mathcal{C} \leftarrow \texttt{core}(k, \mathcal{G})$, $\mathcal{D} = \{\ \}$
while $|\mathcal{C}| > 0$ **do**
    $\mathcal{M} \leftarrow \{i \in \mathcal{V}_\mathcal{C} \mid i = \arg\max [d_i]\}$; `// ` $\mathcal{V}_\mathcal{C}$ `is the set of nodes in` $\mathcal{C}$
    $r \leftarrow \text{uniform}(\mathcal{M})$;
    $\mathcal{C} \leftarrow \mathcal{C} \backslash r$;
    $\mathcal{D} \leftarrow \mathcal{D} \cup \{r\}$;
    $\mathcal{C} \leftarrow \texttt{core}(k, \mathcal{C})$;
end

6.1 The Analysis of the corehd Algorithm

The algorithm produces a *contagious set* of nodes $\mathcal{D}$, such that after their removal the resulting graph has an empty k-core. Consequently, the size of $\mathcal{D}$ provides an upper bound on the size of the minimal contagious set. The main aim is to show that the size of $\mathcal{D}$. normalized by the number of nodes, has a well defined limit, and to compute this limit.

With a proper book-keeping and dynamic updating of the k-core, the running time of COREHD on graphs with bounded degree is $O(N)$. The algorithm can be implemented such that in each iteration exactly one node is removed: if a node of degree smaller than k is present, it is removed, else a node of highest degree is removed. Thus running the algorithm reduces to keeping track of the degree of each node. If the largest degree is $O(1)$ this can be done in $O(1)$ steps by binning all the nodes of equal degree. If a node is removed only the degrees of all its $O(1)$ neighbors must be moved to the new adequate bins. An implementation of the COREHD algorithm is available through the open repository [Sch18] or the digital appendix of this thesis.

6.1.1 Reduction into a Random Process on the Degree Distribution

In the next several sections closed-form deterministic equations are derived for the macroscopic behavior of the COREHD algorithm in the limit of large random graphs that are taken from the configuration model. This is possible because, when the COREHD algorithm is applied to a random graph from the configuration model (parameterized by its degree distribution), the result (conditioned on the new degree distribution) is also distributed according to the configuration model. Thus, one can analyze the COREHD algorithm by tracking the evolution of the degree distribution.

In particular, the behavior of the COREHD procedure, averaged over the graph $\mathcal{G}$, can be described explicitly in terms of the following process involving colored balls in an urn [Pfi14].

At time step n there are N_n balls in the urn, each of which carries a color d_i, with $d_i \in \{1, \dots, d\}$ and d being the maximum degree in the graph at the corresponding time step. The number of balls of color q at iteration time step n is $v_q(n)$. The colors of the balls are initialized in such a way that, at time $n = 0$, the number of balls, N_0, is the size of the k-core of the original graph $\mathcal{G}$. The initial values of their colors are chosen from the degree distribution of the k-core.

The COREHD algorithm then consists of repeatedly applying the following two steps.

- In a first process, called **removal** (line 3-6 in Alg. 2), one ball is drawn among the $v_d(n)$ balls of maximum degree (color) d, uniformly at random. Next, d balls $\{i_1, i_2, \dots, i_d\}$ are drawn with colors following the probability that an outgoing edge from one node is incoming to another node of degree q, i.e.,

  $$K(q) \coloneqq qP(q)/c\,,$$

  with $P(q)$ the degree distribution and c the average degree. To conclude the first step, each of the d balls is replaced by a ball of color $d_{i_j} - 1$ (we assume that there are no double edges).

- In a second process, called **trimming** (line 7 in Alg. 2), we compute the k-core of the current graph. In the urn-model this is equivalent to repeatedly applying the following procedure until $v_i = 0 \; \forall \; i < k$: draw a ball of color $q \in \{1, \dots, k-1\}$ and relabel q other balls chosen according to $K(q)$.

Thus, we have a random process that depends purely on the degree distribution, $P(q)$. Note that in this process we used the fact that the graph was randomly drawn from the configurations model with degree distribution $P(q)$.

One difficulty, when analyzing the above process, is to choose the right observables. In the previous paragraph the nodes were used as observables. However, equally, one might consider the process in terms of the edges of the graph. As outlined in the previous paragraph, it is important to keep track of the distribution $K(q)$. Note that $K(q)$ is related to the fraction of edges that are connected to a node of degree q. Henceforth we will therefore be working directly with the *half-edges* to simplify the analysis.

In order to continue, some notations needs to be introduced. Let each edge (ij) consist of the two half edges that end in the nodes i and j respectively. Let h_q be the total number of half-edges that are connected to nodes of degree q at the current iteration. Furthermore, distinguish nodes of degree smaller than k from all the others. To do so one can *adapt the index notation* and identify

$$h_{k_<} \coloneqq \sum_{q<k} h_q\,. \tag{6.1}$$

The sum over the entries of a vector $\boldsymbol{h}(n) = (h_{k_<}(n), h_k(n), \dots, h_d(n))^T$ is denoted by

$$|\boldsymbol{h}(n)| \coloneqq \sum_{q=k_<}^{d} h_q(n)\,. \tag{6.2}$$

In the new index notation, the unit vectors are defined as

$$\hat{\mathbf{e}}_{k_<} = (1, 0, \ldots, 0)^\intercal,\ \hat{\mathbf{e}}_k = (0, 1, 0, \ldots, 0)^\intercal, \ldots, \hat{\mathbf{e}}_d = (0, \ldots, 0, 1)^\intercal.$$

In the urn model, each ball now represents a half-edge and its color is according to the degree of the node that this half-edge is connected to.

The two processes (trimming and removal) can be described in terms of half-edges as follows. We start with the removal process. It can be recast in the following rule

$$\begin{aligned} &(i) \quad \boldsymbol{h} \leftarrow \boldsymbol{h} - d \cdot \hat{\mathbf{e}}_d \\ &(ii) \ \text{Repeat } d \text{ times: } \boldsymbol{h} \leftarrow \boldsymbol{h} + \mathbf{A}\boldsymbol{z} \end{aligned} \tag{6.3}$$

where the vector $\boldsymbol{z} \in \mathbb{R}^{d-k+2}$ is a random vector that has zeros everywhere except in one of the $d-k+2$ directions, in which it carries a one. The probability that $\boldsymbol{z}$ is pointing in direction at iteration n is given by $h_q(n)/\left|\boldsymbol{h}(n)\right|$ for $q = k_<, k, k+1, \ldots, d$. When a node of degree d is removed from the graph, together with its half-edges, the remaining cavity leaves behind some dangling half-edges that are pruned away in step (ii) using the following **relabelling** matrix

$$\mathbf{A} = \begin{pmatrix} -1 & k-1 & & & & \\ & -k & k & & & \\ & & \ddots & \ddots & & \\ & & & -(d-1) & d-1 & \\ & & & & -d & \end{pmatrix} \in \mathbb{R}^{(d-k+2)\times(d-k+2)}\,. \tag{6.4}$$

Analogously, the trimming process can be cast in the following update rule:

$$\begin{aligned} &\text{while } h_{k_<} > 0 \text{ iterate:} \\ &(i) \quad \boldsymbol{h} \leftarrow \boldsymbol{h} - \hat{\mathbf{e}}_{k_<} \\ &(ii) \ \boldsymbol{h} \leftarrow \boldsymbol{h} + \mathbf{A}\boldsymbol{z}\,. \end{aligned} \tag{6.5}$$

Step (i) removes a *single half-edge* of degree $<k$ and subsequently step (ii) trims away the dangling cavity half-edge. The position q, in which to place the one in the random variable $\boldsymbol{z}$, is again chosen from $h_q(n)/\left|\boldsymbol{h}(n)\right|$ for $q = k_<, k, k+1, \ldots, d$. Let us refer to a single iteration ((i) and (ii)), as a **trimming step**. The trimming process consists of as many trimming steps as necessary to reach the stopping criterion $h_{k_<} = 0$.

The above equations should be compared with the procedure outlined in Sec. 1.2.3 to anticipate where the journey is going to.

The advantage of working with a representation in terms of half-edges is that we do not need to distinguish the different edges of color $k_<$. Further $\left|\boldsymbol{h}(n)\right|$ is deterministic because each column of (6.4) sums to the same constant. During the removal step, $2d$ half-edges are removed and in one trimming step, 2 half-edges (resp. d edges and one edge). However,

the number of necessary trimming steps is a random variable. We have effectively traded the randomness in $|\boldsymbol{h}(n)|$ for randomness in the running time. For now we have simply shifted the problem into controlling the randomness in the running time. In Sec. 6.1.3 it will be shown that transitioning to continuous time resolves this issue, after averaging, by determination of the running time δt as a function of $\mathbb{E}\boldsymbol{h}(t)$.

This alternating process is also related to low-complexity algorithms for solving K-SAT problems [Ach00; Ach01; CM01; DM06]. These K-SAT solution methods alternate between guessing variables, which may create new unit clauses, and trimming unit clauses via unit clause propagation. Due to this connection, the differential equation analyses for these two problems are somewhat similar.

6.1.2 Taking the Average over Randomness

As the equations stand in (6.3) and (6.5) they define a random process that behaves just as the COREHD algorithm on a random graph $\mathcal{G}$. With $\boldsymbol{z}$ and the iteration time of the trimming process implicitly containing all randomness. In terms of the urn model, the random variable $\boldsymbol{z}$ indicates the color of the second half-edge that is left behind after the first was removed. We denote the average over $\boldsymbol{z}$ as

$$\overline{\bullet} \coloneqq \mathbb{E}_{\boldsymbol{z}}\left[\bullet\right] .$$

Performing the average over the randomness, per se, only yields the average behavior of the algorithm. However, in Sec. 6.1.4 it is shown that the stochastic process concentrates around its average in the continuous limit.

Next, the combination of steps (i) *and* (ii) in eq. (6.3), for the removal, and eq. (6.5), for the trimming is considered. In order to write the average of the removal step, we recall that the probability that one half-edge is connected to a color $q \in \{k_<, k, k+1, \ldots, d\}$ is given by $h_q(n)/\left|\boldsymbol{h}(n)\right|$. In the large system limit the average drift of a full iteration of the removal step can be written as

$$\begin{aligned}
\overline{\boldsymbol{h}}(n+1) &= \left(\mathbf{1} + \frac{1}{|\overline{\boldsymbol{h}}(n)| - d - (d-1)}\mathbf{A}\right) \cdots \left(\mathbf{1} + \frac{1}{|\overline{\boldsymbol{h}}(n)| - d}\mathbf{A}\right) \left(\overline{\boldsymbol{h}}(n) - d\,\hat{\mathbf{e}}_d\right) \\
&= \left(\mathbf{1} + \sum_{j=0}^{d-1} \frac{1}{|\overline{\boldsymbol{h}}(n)| - d - j}\mathbf{A}\right) \overline{\boldsymbol{h}}(n) + \left(\sum_{k=2}^{d} c_k \mathbf{A}^k\right) \overline{\boldsymbol{h}}(n) - d\,\left(1 + \sum_{k=1}^{d} \tilde{c}_k \mathbf{A}^k\right) \hat{\mathbf{e}}_d \\
&= \left(\mathbf{1} + \sum_{j=0}^{d-1} \frac{1}{|\overline{\boldsymbol{h}}(n)| - d - j}\mathbf{A}\right) \overline{\boldsymbol{h}}(n) - d\,\hat{\mathbf{e}}_d + O(\frac{1}{|\overline{\boldsymbol{h}}(n)|}) \\
&= \left(\mathbf{1} + \frac{d}{|\overline{\boldsymbol{h}}(n)|}\mathbf{A}\right) \overline{\boldsymbol{h}}(n) - d\,\hat{\mathbf{e}}_d + O(\frac{1}{|\overline{\boldsymbol{h}}(n)|}) ,
\end{aligned} \tag{6.6}$$

where $\mathbf{1}$ represents the identity matrix. In the above estimate, d intermediate steps are used to transition from $n \to n+1$; that is the removal of a whole degree d node. We assume

that d is $O(1)$. It then follows that the coefficients c_k and $\tilde{c}_k$ are $O(|\overline{\boldsymbol{h}}(n)|^{-k})$. The last line follows from a similar estimate for the leading term in the sum.

The average removal step can now be written as

$$\overline{\boldsymbol{h}}(n+1) = \overline{\boldsymbol{h}}(n) + \mathbf{A}_d \frac{\overline{\boldsymbol{h}}(n)}{|\overline{\boldsymbol{h}}(n)|} \,. \tag{6.7}$$

with the effective average drift matrix

$$\mathbf{A}_d \coloneqq d(\mathbf{A} + \mathbf{B}_d) \,, \tag{6.8}$$

where the matrix $\mathbf{B}_d$ has all entries in the last row equal to -1 and zeros everywhere else, such that $\mathbf{B}_d \boldsymbol{v} = -\hat{\mathbf{e}}_d$ for a non-negative, normalized vector $\boldsymbol{v}$. Similarly, taking the average in one trimming step (6.5) yields the following averaged version

$$\overline{\boldsymbol{h}}(n+1) = \overline{\boldsymbol{h}}(n) + \mathbf{A}_{k_<} \frac{\overline{\boldsymbol{h}}(n)}{|\overline{\boldsymbol{h}}(n)|} \,. \tag{6.9}$$

For the trimming step the effective drift is simply

$$\mathbf{A}_{k_<} \coloneqq \mathbf{A} + \mathbf{B}_{k_<} \,. \tag{6.10}$$

where now $\mathbf{B}_{k_<}$ has all its entries in the first row equal to -1 and zeros everywhere else.

We emphasize that the two processes (6.7) and (6.9), while acting on the same vector, are separate processes and the latter, (6.9), must be repeated until the stopping condition $\overline{h}_{k_<} = 0$ is hit. Note also, that in the trimming step, one iteration $n \to n+1$ indicates the deletion of a single edge, while it indicates the deletion of a whole node in the removal process.

6.1.3 Operator and Continuous Limits

As discussed at the end of Sec. 6.1.1, a key observation, by virtue of which we can proceed, is that $|\boldsymbol{h}(n)|$ is deterministic (and hence equal to its average) during both, the removal and trimming, steps. This is due to the structure of $\mathbf{A}_{k_<}$ and $\mathbf{A}_d$ that have columns sums independent of the row index:

$$q_{k_<} \equiv \sum_i [\mathbf{A}_{k_<}]_{ij} = -2 \quad \text{and} \quad q_d \equiv \sum_i [\mathbf{A}_d]_{ij} = -2d$$

The only randomness occurs in the stopping time of the trimming process.

In this section the transition to the continuous time-variable t is performed. To that end we define the scaled process

$$\boldsymbol{\eta}(t) = \lim_{N_0 \to \infty} \frac{1}{N_0} \overline{\boldsymbol{h}}(tN_0) \tag{6.11}$$

and presume that the derivative $\boldsymbol{\eta}'(t)$ is equal to its expected change. Here N_0 stands for the initial number of vertices in the graph.

Before proceeding to the analysis of COREHD, let us first describe the solution for the two processes (removal and trimming) as if they were running separately. Let us indicate the removal process (6.7) and trimming processes (6.9) with subscripts $\alpha = d$ and $\alpha = k_<$ respectively. It then follows from (6.7) and (6.9) that the expected change is equal to

$$\boldsymbol{\eta}'_\alpha(t) = \mathbf{A}_\alpha \frac{\boldsymbol{\eta}_\alpha(t)}{|\boldsymbol{\eta}_\alpha(t)|} \,. \tag{6.12}$$

Owing to the deterministic nature of the drift terms, $\mathbf{A}_\alpha$, we have

$$|\boldsymbol{\eta}_\alpha(t)| = 1 + q_\alpha t \tag{6.13}$$

and the above differential equation can be solved explicitly as

$$\boldsymbol{\eta}_\alpha(t) = \exp\left[\frac{\mathbf{A}_\alpha}{q_\alpha} \ln\left(1 + q_\alpha t\right)\right] \boldsymbol{\eta}_\alpha(0) \,. \tag{6.14}$$

We have thus obtained an analytic description of each of the two separate processes (6.7) and (6.9).

Note that this implies that we can analytically predict the expected value of the random process in which all nodes of degree d are removed from a graph successively until none remains and then all nodes of degree smaller than k are trimmed. This already provides improved upper bounds on the size of the minimal contagious sets, that we report in Table 6.2 (cf. "two stages"). This "two stages" upper bound has the advantage that no numerical solution of differential equations is required. The goal, however, is to analyze the COREHD procedure that merges the two processes into one, as this should further improve the bounds.

Crucially, the running time of the trimming process depends on the final state of the removal process, i.e., the differential equations become nonlinear in $\boldsymbol{\eta}(t)$. As a consequence, they can no longer be brought into a simple, analytically solvable, form (at least as far as we were able to tell). To derive the differential equations that combine the removal and trimming processes and track COREHD we will be working with the operators that are obtained from the iterative steps, (6.7) and (6.9), in the continuous limit (6.12). The evolution within an infinitesimally small removal step ($\alpha = d$), respectively trimming step ($\alpha = k_<$), follows from (6.12) to

$$\boldsymbol{\eta}_\alpha(t + \delta t) = \hat{\mathbf{T}}_\alpha(\delta t, t)\boldsymbol{\eta}_\alpha(t) \,, \tag{6.15}$$

where we introduced the propagator

$$\hat{\mathbf{T}}_\alpha(\delta t, t) = \left(\mathbf{1} + \frac{\mathbf{A}_\alpha}{|\boldsymbol{\eta}_\alpha(t)|}\,\delta t\right) .$$

In what follows, we will be considering the removal and trimming processes to belong to one and the same process and therefore $\boldsymbol{\eta}(t)$ will no longer be carrying subscripts. Upon combination, a full step in the combined process in terms of the operators then reads

$$\boldsymbol{\eta}\left(t + \delta t\right) = \hat{\mathbf{T}}_{k_<}(\hat{\delta t}, t + \delta t)\, \hat{\mathbf{T}}_d(\delta t, t)\, \boldsymbol{\eta}\left(t\right) \,. \tag{6.16}$$

Note that one infinitesimal time step is the continuous equivalent of the removal of one degree d node, *together* with the resulting cascade of degree $< k$ nodes (the trimming process). It is for that reason that the final (continuous) time, after which the k-core vanishes, will be directly related to the size of the set $\mathcal{D}$ in Alg. 2. Note also that in $\hat{\mathbf{T}}_{k_<}(\hat{\delta t}, t+\delta t)$ we replaced the running time with the operator $\hat{\delta t}$. It acts on a state to its right and can be computed from the condition that *all* the nodes of degree smaller k must be trimmed after completing a full (infinitesimal) step of the combined process, so that

$$\eta_{k_<}(t+\delta t) \stackrel{!}{=} 0\,. \tag{6.17}$$

Requiring this **trimming condition** in eq. (6.16) we get, from an expansion to linear order in δt, that

$$\hat{\delta t} = \delta t \cdot \left(-\frac{[\mathbf{A}_d \boldsymbol{\eta}(t)]_{k_<}}{[\mathbf{A}_{k_<} \boldsymbol{\eta}(t)]_{k_<}} \right)\,. \tag{6.18}$$

We can now use this equation to eliminate the dependence on $\hat{\delta t}$ in the combined operator $\hat{\mathbf{T}}_{k_<}(\hat{\delta t}, t)\,\hat{\mathbf{T}}_d(\delta t, t)$. Using (6.15) and keeping only first order terms in δt in (6.16) yields

$$\boldsymbol{\eta}(t+\delta t) = \boldsymbol{\eta}(t) + \left[-\frac{[\mathbf{A}_d \boldsymbol{\eta}(t)]_{k_<}}{[\mathbf{A}_{k_<} \boldsymbol{\eta}(t)]_{k_<}}\, \mathbf{A}_{k_<} + \mathbf{A}_d \right] \frac{\boldsymbol{\eta}(t)}{|\boldsymbol{\eta}(t)|}\, \delta t, \tag{6.19}$$

which leads us to the following differential equation

$$\boldsymbol{\eta}'(t) = \left[\varphi\left(\boldsymbol{\eta}(t)\right) \mathbf{A}_{k_<} + \mathbf{A}_d \right] \frac{\boldsymbol{\eta}(t)}{|\boldsymbol{\eta}(t)|}\,. \tag{6.20}$$

The nonlinearity $\varphi(\cdot)$ is directly linked to the trimming time and defined as

$$\varphi\left(\boldsymbol{\eta}(t)\right) \equiv -\frac{[\mathbf{A}_d \boldsymbol{\eta}(t)]_{k_<}}{[\mathbf{A}_{k_<} \boldsymbol{\eta}(t)]_{k_<}} = -\frac{d\,(-\eta_{k_<}(t) + (k-1)\eta_k(t))}{-|\boldsymbol{\eta}(t)| - \eta_{k_<}(t) + (k-1)\eta_k(t)} = \frac{d\,(k-1)\eta_k(t)}{|\boldsymbol{\eta}(t)| - (k-1)\eta_k(t)}\,. \tag{6.21}$$

To obtain the last equality in (6.21) we used the trimming condition to set $\eta_{k_<}(t) = 0$. The initial conditions are such that the process starts from the k-core of the original graph. This is achieved by solving (6.14), with $\alpha = k_<$, for arbitrary initial degree distribution $\boldsymbol{\eta}(0)$ until $\eta_{k_<}(t) = 0$. Hence, the set of differential equations defined by (6.20), can be written explicitly as

$$\eta'_{k_<}(t) = 0 \tag{6.22}$$

$$\eta'_i(t) = d \cdot i \cdot \frac{-\eta_i(t) + \eta_{i+1}(t)}{|\boldsymbol{\eta}(t)| - (k-1)\eta_k(t)} \qquad \text{for } k \le i < d \tag{6.23}$$

$$\eta'_d(t) = -d + d^2 \cdot \frac{-\eta_d(t)}{|\boldsymbol{\eta}(t)| - (k-1)\eta_k(t)}\,. \tag{6.24}$$

6.1.4 Rigorous Analysis

A rigorous analysis of the k-core peeling process for Erdős-Rényi graphs is presented in [PSW96a]. This analysis is based on the Wormald approach [Wor95] but the presentation in [PSW96a] is more complicated because it derives an exact formula for the threshold and there are technical challenges as the process terminates. For random graphs drawn from the configuration model, however, the standard Wormald approach [Wor95] provides a simple and rigorous numerical method for tracking the macroscopic dynamics of the peeling algorithm when the maximum degree is bounded and the degree distribution remains positive. The primary difficulty occurs near termination when the fraction of degree $< k$ edges becomes very small.

The peeling process in COREHD alternates between deleting maximum-degree nodes (removal) and degree $< k$ edges (trimming) and this introduces a similar problem for the Wormald method. In particular, the COREHD peeling schedule typically reduces the fraction of maximum-degree nodes to zero at some point and then the maximum degree jumps downward. At this jump, the drift equation is not Lipschitz continuous and does not satisfy the necessary conditions in [Wor95]. More generally, whenever there are hard preferences between node/edge removal options (i.e., first delete largest degree, then 2nd largest degree, etc.), the same problem can occur.

For COREHD, one solution is to use weighted preferences where the removal of degree $< k$ edges is most preferred, then removal of degree-d nodes, then degree $d-1$ nodes, and so on. In this case, the drift equation remains Lipschitz continuous if the weights are finite but the model dynamics only approximate the COREHD algorithm dynamics. In theory, one can increase the weights to approximate hard preferences but, in practice, the resulting differential equations become too numerically unstable to solve efficiently. A better approach is to use the operator limit described in Sec. 6.1.3. Making this rigorous, however, requires a slightly more complicated argument.

The key argument is that the k-core peeling step (after each maximum-degree node removal) does not last too long or affect too many edges in the graph. A very similar argument (dubbed the Lazy-Server Lemma) is used in the analysis of low-complexity algorithms for solving K-SAT problems [Ach00; Ach01]. In both cases, a suitable stability (or drift) condition is required. For COREHD, we use the following lemma.

Lemma 6.1. *For some $\delta > 0$, suppose $\boldsymbol{h}(n)$ satisfies $h_k(n) \leq \frac{1-3\delta}{k-1}|\boldsymbol{h}(n)|$ and $M \triangleq |\boldsymbol{h}(n)| \geq M_0(\delta)$. Consider the COREHD process where a maximum-degree node is removed and then the trimming process continues until there are no edges with degree less than k (see (6.5)). Let the random variable T denote the total number of trimming steps, which also equals the total number edges removed by the trimming process. Then, we have*

$$\Pr\left(T > (k-1)^2\delta^{-2}\ln M\right) \leq M^{-2}.$$

Proof. See Appendix A.2.1. □

Lemma 6.2. *Let $\boldsymbol{\eta}(t)$ be the solution to the operator-limit differential equation* (6.20) *at time t starting from $\boldsymbol{\eta}(0) = \frac{1}{N}\boldsymbol{h}(0)$. Assume, for some $\delta > 0$, there is a $t_0 > 0$ such that $\eta_k(t) \leq \frac{1-4\delta}{k-1} |\boldsymbol{\eta}(t)|$, implying that the denominator of* (6.21) *is positive, and $|\boldsymbol{\eta}(t)| \geq \delta$ for all $t \in [0, t_0]$. Then, there is $C > 0$ such that*

$$\Pr\left(\sup_{t\in[0,t_0]} \left\| \boldsymbol{\eta}(t) - \frac{1}{N}\boldsymbol{h}(\lfloor Nt \rfloor) \right\| > \frac{C}{N^{1/4}} \right) = O\left(N^{-1}\right).$$

Proof. See Appendix A.2.2. □

Theorem 6.1. *The multistage* COREHD *process converges, with high probability as $N \to \infty$, to the piecewise solution of the operator-limit differential equation.*

Sketch of Proof. The first step is recalling that the standard k-core peeling algorithm results in graph distributed according to the configuration model with a degree distribution that, w.h.p. as $N \to \infty$, converges to the solution of the standard k-core differential equation [PSW96a]. If k-core is not empty, then the COREHD process is started. To satisfy the conditions of [Wor99a, Theorem 5.1] (The Theorem is outlined in Appendix A.1), the process is stopped and restarted each time the supply of maximum-degree nodes is exhausted. Since the maximum degree is finite, this process can be repeated to piece together the overall solution. Using Lemma 6.2, we can apply [Wor99a, Theorem 5.1] at each stage to show the COREHD process follows the differential equation (6.20). It is important to note that the cited theorem is more general than the typical fluid-limit approach and allows for unbounded jumps in the process as long as they occur with low enough probability. □

6.1.5 Evaluating the Results

Here we clarify how the upper bound is extracted from the equations previously derived. Note that the nonlinearity (6.21) exhibits a singularity when

$$|\boldsymbol{\eta}(t)| = (k-1)\eta_k(t)\,, \tag{6.25}$$

that is, when the gain (r.h.s.) and loss (l.h.s.) terms in the trimming process are equal. This can be either trivially true when no more nodes are left, $|\boldsymbol{\eta}(t)| = 0$, or it corresponds to an infinite trimming time. The latter is precisely the point where the size of the k-core jumps downward discontinuously, whereas the first case is linked to a continuous disappearance of the k-core. Either of these two cases define the *stopping time* t_{s} of the differential process (6.20). By construction the stopping time t_{s} provides the size of the set $\mathcal{D}$ that contains all the nodes the COREHD algorithm removed to break up the k-core. It hence also *provides an upper bound* on the size of the minimal contagious set, i.e. the smallest such set that removes the k-core.

Note that $\boldsymbol{\eta}(t_{\mathrm{s}} - \epsilon)$ (for an infinitesimally small ϵ) gives the size of the k-core, right before it disappears. For all the cases investigated in this paper we found that solving eqs. (6.22–6.24) for $k = 2$ yields a continuous disappearance of the 2-core, and for $k \geq 3$ the stopping criteria yield discontinuous disappearance of the k-core.

In order to solve the above set of ODEs numerically, we first use equation (6.14) to trim away nodes of color $k_<$, i.e. reduce the graph to it's k-core. Then we use equation (6.19) recursively, until the last component η_d is zero. Subsequently we reduce $\boldsymbol{\eta}$ by removing its last component, send $d \to d-1$, adapt the drift term (6.8) and repeat with the reduced $\boldsymbol{\eta}$ and initial condition given by the result of the previous step. All this is performed until the stopping condition (6.25) is reached. We summarize the procedure in a pseudo-code in Alg. 3. For our code that solves the differential equations we refer to the open repository [Sch18] or the digital appendix of this thesis.

Algorithm 3: Analysis of COREHD. Recall that indices are $k_<$ are referring to the first component of a vector, k to the second and so forth until the last component d.

Data: Initial degree distribution $P(q)$; k
Result: The relative size of set of removed nodes, t_s.

Initialize half-edges:
$\boldsymbol{\eta}(0) = \left(\sum_{i=1}^{k-1} i \cdot P(i), k \cdot P(k), (k+1) \cdot P(k+1), \ldots, d \cdot P(d)\right)^T$;
Compute distribution of half-edges in the k-core:
$\boldsymbol{\eta}_0 = \exp\left[-\frac{\mathbf{A}_{k<}}{2} \cdot \ln(1-2t_0)\right] \boldsymbol{\eta}(0)$ with t_0 such that $\eta_{k<}(t_0) = 0$ and $\mathbf{A}_{k<}$ defined in (6.10);
Set: $d \leftarrow$ the degree associated to the last non-zero component of $\boldsymbol{\eta}_0$; $t_\mathsf{s} \leftarrow 0$;
while $stop \neq true$ **do**
    Solve:
$$\boldsymbol{\eta}'(t) = \left[\frac{d(k-1)\eta_k(t)}{|\boldsymbol{\eta}(t)| - (k-1)\eta_k(t)} \mathbf{A}_{k<} + \mathbf{A}_d\right] \frac{\boldsymbol{\eta}(t)}{|\boldsymbol{\eta}(t)|}$$
    with initial condition $\boldsymbol{\eta}_0$. Until either $\eta_d(t^*) = 0$ or $|\boldsymbol{\eta}(t^*)| = (k-1)\eta_k(t^*)$;
    `// ` $\mathbf{A}_{k<}$ `and` $\mathbf{A}_d$ `defined by (6.10) and (6.8) with` d `set to the current largest degree`
    Update: $t_\mathsf{s} \leftarrow t_\mathsf{s} + t^*$;
    if $|\boldsymbol{\eta}(t^*)| = 0$ *or* $|\boldsymbol{\eta}(t^*)| = (k-1)\eta_k(t^*)$ **then**
        stop $\leftarrow$ true;
    else
        $\boldsymbol{\eta}_0 \leftarrow \eta_{k<:d-1}(t^*)$;
        $d \leftarrow d-1$;
    end
end

Example: two-cores on three regular random graphs. For a simple example of how to extract the upper bound consider the following case. We have $k = 2$ and $d = 3$ and we set

$k_< = 1$, then the differential equation in (6.20) becomes

$$\begin{aligned}\eta_1'(t) &= 0\\ \eta_2'(t) &= 6 \cdot \left(1 - \frac{\eta_2(t)}{\eta_3(t)}\right)\\ \eta_3'(t) &= -12\,,\end{aligned} \tag{6.26}$$

with initial condition

$$\boldsymbol{\eta}(0) = d \cdot \hat{\mathbf{e}}_d \tag{6.27}$$

because there are dN half-edges, all connected to nodes of degree d initially. The equations are readily solved

$$\begin{aligned}\eta_1(t) &= 0\\ \eta_2(t) &= 3 \cdot \sqrt{1-4t} \cdot \left(1 - \sqrt{1-4t}\right)\\ \eta_3(t) &= 3 \cdot (1-4t)\,.\end{aligned} \tag{6.28}$$

According to (6.25) the stopping time is $t_s = 1/4$, i.e. $|\mathcal{D}| = N/4$, which suggests that the decycling number (2-core) for cubic random graphs is bounded by $N/4+o(N)$. In accordance with Theorem 1.1 in [BWZ02] this bound coincides with the actual decycling number. For $d > 3$ the lower and upper bounds do not coincide an the stopping time resulting from our approach only provides an upper bound.

Finding a closed form expression for the generic case is more involved and we did not manage to do it. However, very reliable numerical resolution is possible. The simplest approach to the differential equations is to work directly with (6.19) as indicated in Alg. 3.

6.1.6 corehd Analyzed and Compared with Existing Results

In this section we evaluate the upper bound on the minimal contagious set, obtained by our analysis of COREHD. In Fig. 6.1 we compare the fraction of nodes of a given degree that are in the graph during the COREHDprocedure. We overlay the results obtained from solving the differential equations with the averaged timelines obtained from direct simulations of the algorithm.

Table 6.1 collects the comparisons between direct numerical simulations of the COREHD algorithm with the prediction obtained from the differential equations. The two are in excellent agreement. Higher precision can be obtained without much effort. When analyzing Erdős-Rényi graphs, it is necessary to restrict the largest degree to avoid an infinite set of differential equations. The resulting error is exponentially small in the maximum degree, and hence in practice insignificant.

Confident with this cross-check of our theory, we proceed to compare with other theoretical results. As previously stated, Guggiola and Semerjian [GS15] have derived the size of the minimal contagious sets for random regular graphs using the cavity method. At the same time, several rigorous upper bounds on size of the minimal contagious set exist [ABW10;

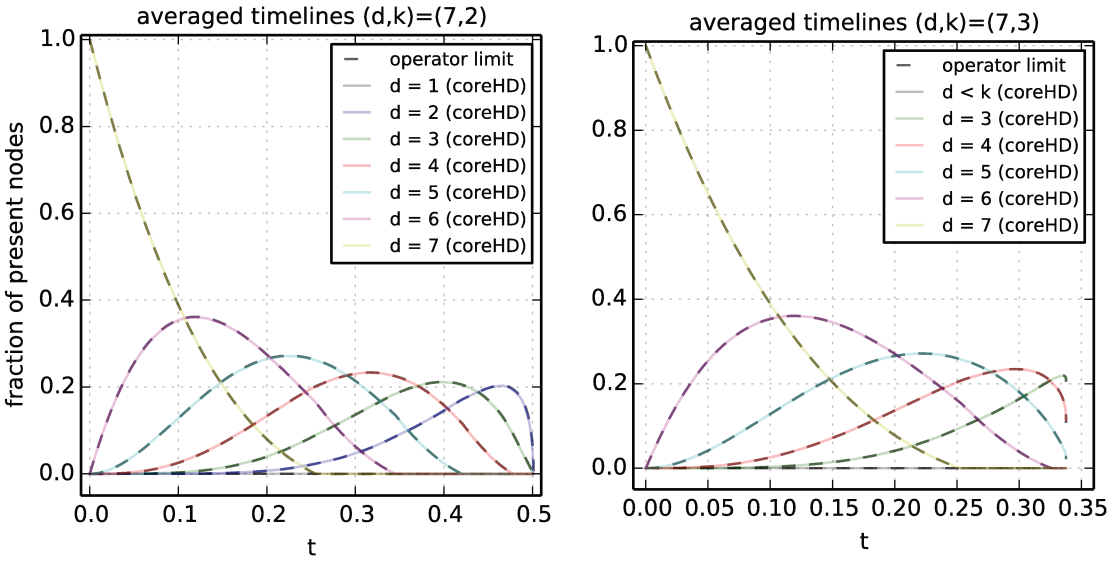

Figure 6.1: Comparison of the solution of the differential equations (cf. Alg. 3) with a direct simulations of the COREHD algorithm on several graphs of 2^{14} nodes, here with $d = 7$ and $k = 2$ (left hand side) and $k = 3$ (right hand side). The time t is directly related to the fraction of removed nodes of degree $\geq k$. For $k = 2$ the k-core disappears continuously at $t_s = 0.5006$, for $k = 3$ discontinuously at $t_s = 0.3376$.

d	k	RRG COREHD	RRG theory	final core	ERG COREHD	ERG theory	final core
3	2	.2500(0)	.25000	0	.1480(7)	.14809	0
4	2	.3462(3)	.34624	0	.2263(0)	.22634	0
	3	.0962(5)	.09623	.67349	.0388(5)	.03887	.32718
5	2	.4110(7)	.41105	0	.2924(8)	.29240	0
	3	.2083(6)	.20832	.48428	.1068(2)	.10679	.33105
	4	.0476(9)	.04764	.85960	–	–	–
6	2	.4606(2)	.46063	0	.3480(8)	.34816	0
	3	.2811(2)	.28107	.40916	.1700(1)	.16994	.34032
	4	.1401(1)	.14007	.68672	.0466(0)	.04662	.46582
	5	.0280(7)	.02809	.92410	–	–	–
7	2	.5006(0)	.50060	0	.3954(7)	.39554	0
	3	.3376(1)	.33757	.36179	.2260(0)	.22597	.32878
	4	.2115(0)	.21150	.58779	.1043(2)	.10429	.46788
	5	.1010(2)	.10100	.78903	.0088(5)	.00882	.54284
	6	.0184(6)	.01842	.95038	–	–	–

Table 6.1: Comparison between direct simulations of the COREHD Algorithm 2 and the theoretical results for random graphs of different degree d and core-index k. The left hand side of the table reports results for random regular graphs, the right hand side for Erdős-Rényi random graphs of the corresponding average degree. The simulation results are averaged over 50 runs on graphs of size $N = 2^{19}$. The fractions of removed nodes agree up to the last stable digit (digits in brackets are fluctuating due to finite size effects). We also list the size of the k-core just before it disappears (in the column "final core").

d (degree)	Bau, Wormald, Zhou	two stages	COREHD	cavity method
3	.25	.3750	.25	.25
4	.3955	.3913	.3462	.3333
5	.4731	.4465	.4110	.3785
6	.5289	.4911	.4606	.4227
7	.5717	.5278	.5006	.4602

Table 6.2: Best known upper bounds on the minimal decycling sets by [BWZ02], compared to the upper bounds obtained from our analysis when all nodes of maximum degree are removed before the graph is trimmed back to its 2-core (two stages), and to our analysis of COREHD. The last column gives the non-algorithmic cavity method results of Guggiola and Semerjian [GS15] that provide (non-rigorously) the actual optimal sizes.

Rei12; CFKR15]. In particular the authors of [BWZ02] provide upper bounds for the decycling number (2-core) that are based on an analysis similar to ours, but of a different algorithm.[1] In table 6.2 we compare the results from [BWZ02] with the ones obtained from our analysis and the presumably exact results from [GS15]. We clearly see that while COREHD is not quite reaching the optimal performance, yet the improvement over the existing upped bound is considerable.

In Tab. 6.3 we quantify the gap between our upper bound and the results of [GS15] for larger values of k. Besides its simplicity, the COREHD algorithm provides significantly better upper bounds than those known before. Clearly, we only consider a limited class of random graphs here and the bounds remain away from the conjectured optimum. However, it is worth emphasizing that previous analyses were often based on much more involved algorithms. The analysis in [CFKR15] or the procedure in [BWZ02] are both based on algorithms that are more difficult to analyze.

6.2 Improving corehd

The main focus of our work lays on the analysis of COREHD on random networks. However, rather naturally the question of how to improve over it raises. In this section the possibility of a simple local strategy that systematically improves over the COREHD performance is evaluated. We show that introducing additional local information about the direct neighbors of a node into the decision process can significantly improve the performance, while conserving the essential time complexity.

The COREHD algorithm (Alg. 2) does not take into account any information of the neighborhood of a node. The theoretical analysis in the previous section owes its simplicity to this

[1] The numerical values, provided for the bounds in [BWZ02] are actually not correct, as the authors realized and corrected in a later paper [HW08].

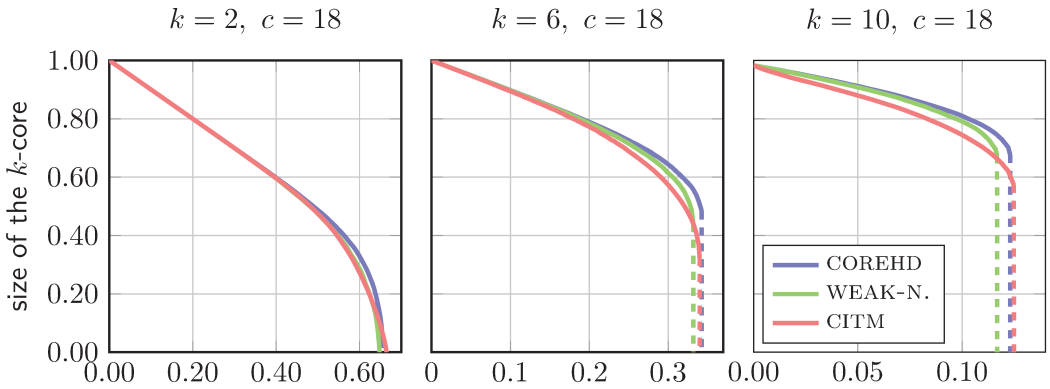

Figure 6.2: A typical transition of the three algorithms, with L for the CITM set to the size of the graph. When k is increased the gap between the CITM algorithm and the COREHD algorithm closes and they become close in terms of performance. However, there remains a gap to the WEAK-NEIGHBOR algorithm.

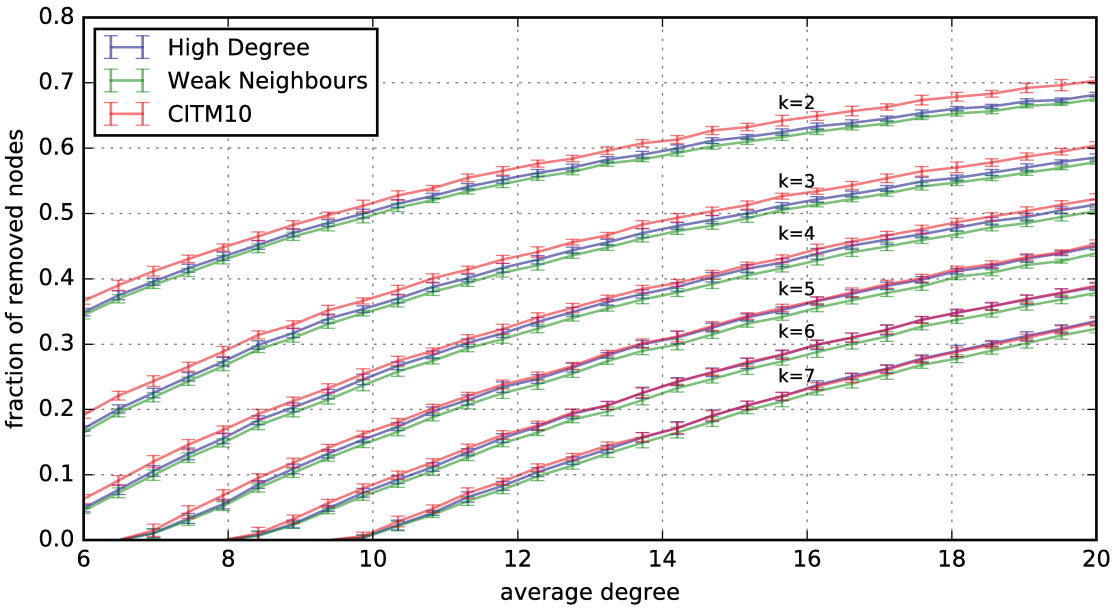

Figure 6.3: Comparison of the three algorithms (here CITM with $L = 10$) on ER graphs for different k and vary the average degree. Compare also Tab. 6.3 for results on regular random graphs.

fact. However, the idea behind COREHD can be extended to the next order by considering the structure of the nearest neighbors of a node. Once we include the nearest neighbors the number of options is large. Our aim is not to make an extensive study of all possible strategies, but rather to point out some heuristic arguments that lead to improvement.

According to the previous section, selecting high degree nodes is a good strategy. Another natural approach is a greedy strategy that, in each step, selects a node such that the caused cascade of nodes dropping out of the core is maximized [GS15; CFKR15; SXS+17]. The vanilla-greedy approach removes nodes according to the size of cascade that is caused by their removal. Nodes that cause the largest cascade in the subsequent trimming process are removed first. The high degree version of this approach additionally focuses on the nodes of maximum degree, e.g., by picking nodes according to $\arg\max_i [d_i + s_i]$ with s_i now being the size of the corona (not limited to the direct neighborhood, but the total graph). Note that these algorithms have $O(N^2)$ running time.

In the following we list some strategies that aim to improve over COREHD. On graphs with $O(1)$ maximum degree they can all be implemented in $O(N)$ running time.[2]

weak-neighbor

The WEAK-NEIGHBOR strategy aims to remove first those nodes that have high degree *and* low average degree of the neighbors, thus causing a larger trimming step on average. There are different ways to achieve this. We tried two strategies that both yield very similar results on all the cases considered. These two strategies are

- The order in which the nodes are removed is according to $\arg\max_i d_i - s_i$ with d_i being the degree of node i and s_i the average degree of the neighbors of i.
- We separate the two steps. First select the set $\mathcal{M} \leftarrow \{i \in \mathcal{V}_\mathcal{C} \mid i = \arg\max_i [d_i]\}$ and then update it according to $\mathcal{M} \leftarrow \{i \in \mathcal{M} \mid i = \arg\min_i \left[\sum_{j\in\partial i} d_j\right]\}$.

Our implementation of the WEAK-NEIGHBOR algorithm is available through our repository [Sch18] or the digital appendix of this thesis.

corehdld

This approach selects high degree nodes, but then discriminates those that have many neighbors of high degree. The idea is that nodes that have neighbors of large degree might get removed in the trimming procedure preferentially and hence the degree of the node in question will likely decrease. More specifically we implemented the following:

- First select the set $\mathcal{M} \leftarrow \{i \in \mathcal{V}_\mathcal{C} \mid i = \arg\max_i [d_i]\}$ and then update it according to $\mathcal{M} \leftarrow \{i \in \mathcal{M} \mid i = \arg\max_i \left[\sum_{j\in\partial i} \mathbb{I}[d_j < d_i]\right]\}$.

[2] Following up on [BDSZ16], we mention that, in applications of practical interest, it is possible to improve each of the mentioned algorithm by adding an additional random process that attempts to re-insert nodes. Consider the set $\mathcal{S} \subset \mathcal{G}$ of nodes that, when removed, yield an empty k-core. Then find back the nodes in $\mathcal{S}$ that can be re-inserted into the graph without causing the k-core to re-appear.

corehd-critical

The idea is slightly different in that the discriminating step aims to maximize the local cascade caused by the removal of a node. This approach is motivated by the typical greedy strategies, brought down to the level of the direct neighbors. Nodes are first selected according to their degree and subsequently among them we remove nodes first that have the largest number of *direct neighbors* that will drop out in the trimming process.

- First select the set $\mathcal{M} \leftarrow \{i \in \mathcal{V}_\mathcal{C} \mid i = \arg\max_i [d_i]\}$ and then update it according to $\mathcal{M} \leftarrow \{i \in \mathcal{M} \mid i = \arg\max_i \left[\sum_{j\in\partial i} \mathbb{I}\left[d_j \leq k\right]\right]\}$.

Finally, it is interesting to contrast this algorithm with a version in which the high degree selection step is left out, i.e., select $\mathcal{M} \leftarrow \{i \in \mathcal{V}_\mathcal{C} \mid i = \arg\max_i \left[\sum_{j\in\partial i} \mathbb{I}\left[d_j \leq k\right]\right]\}$ and then remove at random from this set.

We find that all of them improve over the COREHD algorithm (at least in some regimes). Second, we find that among the different strategies the WEAK-NEIGHBOR algorithm performs best. While we have systematic numerical evidence that the WEAK-NEIGHBOR strategy performs best, it is not clear which are the effects responsible. We also report that the greedy procedures tend to perform worse than WEAK-NEIGHBOR when $k \ll d-1$ and become comparable when $k \approx d-1$.

Next we contrast the COREHD performance and the WEAK-NEIGHBOR performance with the performance of the recently introduced [SXS+17] algorithm, CITM-L, that uses information from a neighborhood up to distance L and was shown in [SXS+17] to outperform a range of more basic algorithm. In Fig. 6.3 we compare the performances of the WEAK-NEIGHBOR, COREHD, and CITM-10 (beyond $L = 10$ resulted in negligible improvements) on Erdős-Rény graphs. For small k the COREHD algorithm outperforms the CITM algorithm but, when k is increased, the performance gap between them shrinks. For large k (e.g. in Fig. 6.2) CITM may outperform COREHD. Both COREHD and CITM are outperformed by the WEAK-NEIGHBOR algorithm in all the cases we tested. In addition to the results on Erdős-Rény graphs we summarize and compare all three algorithms on random regular graphs and put them in perspective to the cavity method results in Tab. 6.3. It can be concluded that the optimal use of information about the neighborhood is not given by the CITM algorithm. What the optimal strategy is, that only uses local information up to a given distance, remains an intriguing open problem for future work.

6.3 Conclusion

We have derived a provably exact analysis of the COREHD algorithm for random graphs from the configuration model with bounded maximum degree. To that end a deterministic description of the associated random process on sparse random graphs was derived that lead to a set of non-linear ordinary differential equations. From the stopping time of these differential equations (the time at which the k-core disappears) we extracted an upper bound

d	k	CITM-10	COREHD	WEAK-NEIGHBOR	cavity method
3	2	.254	.2500	.2500	.25000
4	2	.365	.3462	.3376	.33333
	3	.077	.0963	.0744	.04633
5	2	.437	.4111	.3965	.37847
	3	.205	.2084	.1876	.16667
	4	.032	.0477	.0277	.01326
6	2	.493	.4606	.4438	.42270
	3	.296	.2811	.2644	.25000
	4	.121	.1401	.1081	.07625
	5	.019	.0281	.0134	.00582
7	2	.540	.5006	.4831	.46023
	3	.362	.3376	.3206	.30009
	4	.207	.2115	.1813	.15006
	5	.081	.1010	.0686	.04289
	6	.013	.0185	.0077	.00317

Table 6.3: Performance results of the three algorithms (CITM (with $L = 10$), COREHD, WEAK-NEIGHBOR (with $\arg\max_i d_i - s_i$)) on random regular graphs. As well as the conjectured optimal results obtained non-constructively via the cavity method in [GS15].

on the minimal size of the contagious set of the underlying graph ensemble. The derived upper bounds are considerably better than previously known ones.

Next to the theoretical analysis of COREHD we proposed a novel algorithm, WEAK-NEIGHBOR. It is based on selecting large degree nodes from the k-core that have neighbors of low average degree. In numerical experiments on random regular and Erdős-Rény graphs it was shown that the WEAK-NEIGHBOR algorithm outperforms COREHD, as well as other scalable state-of-the-art algorithms [SXS+17]. It further exhibits a satisfying $O(N)$ running time on graphs with maximum degree of $O(1)$.

— 7 —

Matrix Estimation: A Random Matrix Problem in the Replica Space

"My work consists of two parts: of the one which is here, and of everything which I have not written. And precisely this second part is the important one." — Ludwig Wittgenstein

7.1 Prologue

In this chapter we present some work that is related to the problem of matrix factorization, when the matrices that form the product are of extensive rank. Due to finite time effects, this chapter must be considered an unfinished journey.

In **matrix factorization** the aim is as follows: from a noisy measurement of the product of two matrices the aim is to estimate the original matrices, that is, the components that formed the product. More formally, given an $N \times M$ matrix, $\boldsymbol{Y}$, that was obtained from a noisy, elementwise, measurement of the matrix $\boldsymbol{S} \coloneqq \boldsymbol{F}\boldsymbol{X}$, with $\boldsymbol{F} \in \mathbb{R}^{N\times R}$ and $\boldsymbol{X} \in \mathbb{R}^{R\times M}$, the aim is to recover $\boldsymbol{F}$ and $\boldsymbol{X}$. This problem has several important applications [KKM+16b] of practical interest; among them are dictionary learning, sparse principal component analysis, blind source separation, matrix completion and principal component analysis. Each of these problems can be mapped onto matrix factorization by either the channel $P(\boldsymbol{Y} \mid \boldsymbol{F}, \boldsymbol{X})$ and/or the priors $P(\boldsymbol{F})$, $P(\boldsymbol{X})$. To give one example, if one has an i.i.d. Gaußian prior on $\boldsymbol{F}$, an i.i.d. sparse Gaußian prior on $\boldsymbol{X}$ and i.i.d. Gaußian noise, $\boldsymbol{R}$, we are in a setting that can be used to model dictionary learning [KMZ13]: $\boldsymbol{F}$ then represents the dictionary and $\boldsymbol{X}$ the sparse features.

Thus, we are given a matrix $\boldsymbol{Y}$ that is a noisy version of a matrix $\boldsymbol{S}$, that is itself a deterministic function of the matrices $\boldsymbol{F}$ and $\boldsymbol{X}$; we write

$$\boldsymbol{Y} = \boldsymbol{F}\boldsymbol{X} + \boldsymbol{R}, \tag{7.1}$$

with $\boldsymbol{R}$ some noise matrix, not necessarily Gaußian. There is a great technical difference between the case where $R = \Theta(1)$ and $R = \Theta(N) = \Theta(M)$. The former case is the low-

rank matrix factorization that was already encountered in the chapter on crowdsourcing. The latter is much more difficult to solve, because in an expansions w.r.t. $\boldsymbol{S}$ higher order terms survive. Indeed, the statistical physics investigation of the latter problem with **extensive rank** was claimed to have been exactly solved in [KKM+16b]. But as we shall see in the following this analysis relies on the assumption that the overlap concentrates, while our investigation suggests that this is not the case.

A warning, before we move on to the analysis. The results, presented in this chapter are not final. There remain many open questions that we will summarize in the end of this chapter. Nevertheless, a significant fraction of my research time was invested into the analysis of this problem and many failed algorithmic and analytic attempts. I decided to present the following preliminary results because I think that they are a promising route towards the resolution of this problem and in the hope to gain feedback on the current status of the analysis.

7.2 A Simpler Problem

In this chapter we focus on a simplified form of the generic problem (7.1) that will allow us to grasp the important physics behind the problem, without getting lost in the formalities of different channels and priors. We study the simplest possible problem that falls into the above category: The symmetric version with a Gaußian prior on $\boldsymbol{X}$ under symmetric Gaußian noise, $\boldsymbol{W}$.

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{X}^\intercal + \sqrt{\Delta}\boldsymbol{W}\,, \qquad \boldsymbol{X} \in \mathbb{R}^{N\times R}\,, \qquad \alpha \coloneqq \frac{N}{R} \xrightarrow{N\to\infty} \Theta(1)\,. \tag{7.2}$$

It is not too difficult to repeat the computation of [KKM+16b] for this case. Carrying out the same steps as in [KKM+16b] one realizes that the key assumption in that paper is that $R^{-1}\sum_k x^a_{\mu k}x^b_{\nu k} \to \mathbb{E}\left[x^a_{\mu k}x^b_{\nu k}\right] =: Q_{ab}$, with a, b indicating two different replicas. However, this assumption does not hold by reason of higher order correlations. To be more precise, assume that the elements of $\boldsymbol{X}$ are standard Gaußians with zero mean and variance equal to one, then $\boldsymbol{S} = \boldsymbol{X}\boldsymbol{X}^\intercal$ is Wishart distributed and

$$P(\boldsymbol{S}) \propto \det(\boldsymbol{S})^{(R-N-1)/2}\mathrm{e}^{-\mathrm{Tr}\boldsymbol{S}/2}\,.$$

In this example the higher order correlations are stemming from the determinantal interactions among the elements of $\boldsymbol{S}$, as a consequence of which (as we shall see)

$$\frac{1}{R}\sum_k x^a_{\mu k}x^b_{\nu k} \not\to Q_{ab}\,. \tag{7.3}$$

The simplified model, introduced above, will prove to be very valuable for academic purposes and in order to understand the underlying mathematical difficulties with extensive rank matrix factorization. Furthermore, the Gaußian prior can be considered rather difficult for inference as it possesses a large amount of symmetry. More practically realistic priors are typically more restrictive and make inference easier.

7.3 Replica Approach

We are interested in solving the problem (7.2). In this part of the thesis we consider the problem solved, if we can compute the free energy and recover the asymptotic MMSE for the recovery of $\boldsymbol{X}$, i.e., the x-MMSE. We will not consider any algorithmic aspects.

We assume that the elements of $\boldsymbol{X}$ are i.i.d. Gaußian and the noise matrix is a Wigner matrix with variance Δ. In other words, we observe a Wishart matrix, distorted by a Wigner matrix (white Gaußian noise)

$$\boldsymbol{Y} = \frac{1}{N}\boldsymbol{X}\boldsymbol{X}^\intercal + \sqrt{\frac{\Delta}{N}}\boldsymbol{W}\,. \tag{7.4}$$

We have chosen the scaling such that the eigenvalues of $\boldsymbol{Y}$ are $O(1)$, if $x_{\mu k}$ and $w_{\mu\nu}$ are both $O(1)$. In the following we assume, without loss of generality, that the elements of $\boldsymbol{X}$ have zero mean and unity variance. The associated posterior for $\boldsymbol{X}$, given $\boldsymbol{Y}$, reads

$$P(\boldsymbol{X} \mid \boldsymbol{Y}) = \frac{1}{Z(\boldsymbol{Y})} \prod_{\mu\leq N}\prod_{k\leq R} \mathcal{N}\left(x_{\mu k}; 0, \frac{1}{\sqrt{N}}\right) \prod_{1\leq\mu\leq\nu\leq N} \mathcal{N}\left(y_{\mu\nu}; \frac{1}{N}\sum_k x_{\mu k}x_{\nu k}, \sqrt{\frac{\Delta^*}{N}}\right). \tag{7.5}$$

Here Δ^* denotes that the diagonal elements carry twice the noise of the off-diagonal elements.

We aim to compute $\mathbb{E}_{\boldsymbol{Y}} \ln Z(\boldsymbol{Y})$, with $Z(\boldsymbol{Y})$ explicitly given as

$$Z(\boldsymbol{Y}) = c \cdot \int \mathrm{d}\boldsymbol{X} \exp\left(-\frac{N}{2}\mathrm{Tr}\boldsymbol{X}\boldsymbol{X}^\intercal\right) \exp\left[-\frac{N}{4\Delta}\mathrm{Tr}\left(\boldsymbol{Y} - \frac{1}{N}\boldsymbol{X}\boldsymbol{X}^\intercal\right)^2\right]. \tag{7.6}$$

The constant is given by

$$c = \left(\frac{N}{2\pi}\right)^{\frac{NR}{2}} \left(\frac{N}{2\pi\Delta}\right)^{\frac{N(N+1)}{4}} \frac{1}{2^{\frac{N}{2}}}\,.$$

It should be noted that $\boldsymbol{X}$ can only be recovered up to a rotation, since any rotation of the $\boldsymbol{X}$, i.e., $\boldsymbol{X} \to \boldsymbol{X}\boldsymbol{O}$ with $\boldsymbol{O} \in O(R)$ results in the same observations $\boldsymbol{Y}$. The solution is therefore hugely degenerated. Later this degeneracy will allow us to reduce the problem to the eigenvalues of an auxiliary order-parameter-matrix.

7.3.1 Replica Free Energy

In this section we compute $\mathbb{E}_{\boldsymbol{Y}} \ln Z(\boldsymbol{Y})$ with the replica method. The analysis will turn out to simplify if one removes the term

$$\mathbb{E}_{\boldsymbol{Y}}[-\frac{N}{2\Delta^*}\sum_{\mu\leq\nu} y_{\mu\nu}^2] - \ln c$$

from the free energy. This term depends only on the eigenvalue spectrum of $\boldsymbol{Y}$, that can be computed independent of the ground truth $\boldsymbol{X}$. It can thus be added back at the end of the

computation. Removing this part amounts to a rescaling of the partition function with an $\boldsymbol{X}$ independent term

$$\tilde{Z}(\boldsymbol{Y}) = \int \mathrm{d}\boldsymbol{X} \mathrm{e}^{-\frac{N}{2}\mathrm{Tr}\boldsymbol{X}\boldsymbol{X}^\intercal} \mathrm{e}^{-\frac{N}{4\Delta}\mathrm{Tr}(\boldsymbol{Y}-\frac{1}{N}\boldsymbol{X}\boldsymbol{X}^\intercal)^2+\frac{N}{4\Delta}\mathrm{Tr}\boldsymbol{Y}^2} . \tag{7.7}$$

We will now compute $\mathbb{E}_{\boldsymbol{Y}} \ln \tilde{Z}(\boldsymbol{Y})$ with the replica method. The average over the logarithm of the partition function is replaced by an average of the n-fold replicated partition function. The average with respect to $\boldsymbol{Y}$ can be carried out by realizing that $\int \mathrm{d}\boldsymbol{Y} P(\boldsymbol{Y}) = \int \mathrm{d}\boldsymbol{Y}\mathrm{d}\boldsymbol{X}^0 P(\boldsymbol{Y} \mid \boldsymbol{X}^0) P(\boldsymbol{X}^0)$. Note that, due to the symmetry in $\boldsymbol{Y}$, the integration over $\boldsymbol{Y}$ is only over the elements $y_{\mu\leq\nu}$. We obtain

$$\mathbb{E}_{\boldsymbol{Y}}\left[\tilde{Z}(\boldsymbol{Y})^n\right] \propto \int \mathrm{d}\boldsymbol{Y} \int \prod_{a=0}^{n} \mathrm{d}\boldsymbol{X}^a \mathrm{e}^{-\frac{N}{2}\sum_{a=0}^{n}\mathrm{Tr}\boldsymbol{X}^a(\boldsymbol{X}^a)^\intercal} \mathrm{e}^{-\frac{N}{4\Delta}\sum_{a=0}^{n}\mathrm{Tr}(\boldsymbol{Y}-\frac{1}{N}\boldsymbol{X}^a(\boldsymbol{X}^a)^\intercal)^2+\frac{N}{4\Delta}\sum_{a=1}^{n}\mathrm{Tr}\boldsymbol{Y}^2} . \tag{7.8}$$

Expanding the square and introducing the shorthand

$$\mathrm{D}\underline{\boldsymbol{X}} \coloneqq \int \prod_{a=0}^{n} \mathrm{d}\boldsymbol{X}^a \exp\left(-\frac{N}{2}\sum_{a=0}^{n}\mathrm{Tr}\boldsymbol{X}^a(\boldsymbol{X}^a)^\intercal\right)$$

leads to

$$\mathbb{E}_{\boldsymbol{Y}}\left[\tilde{Z}(\boldsymbol{Y})^n\right] \propto \int \mathrm{D}\underline{\boldsymbol{X}} \int \mathrm{d}\boldsymbol{Y} \mathrm{e}^{-\frac{N}{4\Delta}\mathrm{Tr}\boldsymbol{Y}^2+\frac{1}{2\Delta}\mathrm{Tr}\boldsymbol{Y}\left(\sum_{a=0}^{n}\boldsymbol{X}^a(\boldsymbol{X}^a)^\intercal\right)} \mathrm{e}^{-\frac{1}{N}\frac{1}{4\Delta}\sum_{a=0}^{n}\mathrm{Tr}(\boldsymbol{X}^a(\boldsymbol{X}^a)^\intercal)^2} . \tag{7.9}$$

The integration over $\boldsymbol{Y}$ is Gaußian and can be performed:

$$\begin{aligned}
\mathbb{E}_{\boldsymbol{Y}}\left[\tilde{Z}(\boldsymbol{Y})^n\right] &\propto \int \mathrm{D}\underline{\boldsymbol{X}} \mathrm{e}^{-\frac{1}{2\Delta^*}\frac{1}{N}\sum_{a=0}^{n}\sum_{\mu\leq\nu}(\sum_k x^a_{\mu k}x^a_{\nu k})^2} \\
&\qquad \prod_{\mu\leq\nu}\left\{\int \mathrm{d}y_{\mu\nu} \mathrm{e}^{-\frac{N}{2\Delta^*}y_{\mu\nu}^2+\frac{1}{\Delta^*}y_{\mu\nu}\left(\sum_{a=0}^{n}\sum_k x^a_{\mu k}x^a_{\nu k}\right)}\right\} \\
&\propto \int \mathrm{D}\underline{\boldsymbol{X}} \exp\left(-\frac{1}{2\Delta^*}\frac{1}{N}\sum_{a=0}^{n}\sum_{\mu\leq\nu}(\sum_k x^a_{\mu k}x^a_{\nu k})^2 + \frac{1}{2\Delta^*}\frac{1}{N}\sum_{\mu\leq\nu}(\sum_{a=0}^{n}\sum_k x^a_{\mu k}x^a_{\nu k})^2\right) \\
&= \int \mathrm{D}\underline{\boldsymbol{X}} \exp\left(+\frac{1}{2\Delta^*}\frac{1}{N}\sum_{\mu\leq\nu}\sum_{a\neq b}(\sum_{k,k'} x^a_{\mu k}x^b_{\mu k'}x^a_{\nu k}x^b_{\nu k'})\right) \\
&= \int \mathrm{D}\underline{\boldsymbol{X}} \exp\left(+\frac{1}{4\Delta}\frac{1}{N}\sum_{a\neq b}\sum_{k,k'}\left(\sum_\mu x^a_{\mu k}x^b_{\mu k'}\right)^2\right) .
\end{aligned} \tag{7.10}$$

Here we see why it is convenient to remove the $\boldsymbol{X}$ independent term in the beginning: it results in a cancellation of the terms that contain contributions from only one replica, as we see in the step from the second to the third line.

The last integral contains a quartic term, that we eliminate by Hubbard-Stratonovich transform:

$$\mathrm{e}^{\frac{1}{4\Delta}\frac{1}{N}\left(\sum_\mu x^a_{\mu k}x^b_{\mu k'}\right)^2} = \sqrt{\frac{4\pi\Delta}{N}}\int \mathrm{d}Q^{ab}_{k,k'}\mathrm{e}^{-N\frac{\Delta}{4}(Q^{ab}_{k,k'})^2+\frac{1}{2}Q^{ab}_{k,k'}\sum_\mu x^a_{\mu k}x^b_{\mu k'}}. \tag{7.11}$$

Note that $\boldsymbol{Q}^{ab}$ is a symmetric $R\times R$-matrix. The symmetry follows from the assumption of an ultrametric structure, i.e., that Q^{ab} only depends on the distance between a and b. Now that the quartic term is traded for an additional Gaußian field, it is possible to carry out the integration w.r.t. $\boldsymbol{X}$

$$\begin{aligned}\mathbb{E}_{\boldsymbol{Y}}\left[\tilde{Z}(\boldsymbol{Y})^n\right] \propto \int\prod_{a\neq b}\mathrm{d}\boldsymbol{Q}^{ab}\exp\left(-N\frac{\Delta}{4}\sum_{a\neq b}\mathrm{Tr}(\boldsymbol{Q}^{ab})^2\right)\\ \underbrace{\int\prod_a\prod_{\mu,k}\mathrm{d}x^a_{\mu k}\exp\left(-\frac{1}{2}\sum_a\sum_{\mu,k}(x^a_{\mu k})^2+\frac{1}{2}\sum_{a\neq b}\sum_\mu\sum_{k,k'}Q^{ab}_{k,k'}x^a_{\mu k}x^b_{\mu k'}\right)}\\ =\prod_\mu\int\prod_a\prod_k\mathrm{d}x^a_{\mu k}\exp\left(-\frac{1}{2}\sum_a\sum_k(x^a_{\mu k})^2+\frac{1}{2}\sum_{a\neq b}\sum_{k,k'}Q^{ab}_{k,k'}x^a_{\mu k}x^b_{\mu k'}\right).\end{aligned} \tag{7.12}$$

We see that the integration factorizes over the greek indices, μ. This is a manifestation of the fact that the statistics between any two rows is identical. Note however, that the integration *does not* factorize w.r.t. both indices, unlike it was assumed in the analysis of [KKM+16a]. In the following we further denote

$$\underline{\underline{\boldsymbol{Q}}} = \begin{pmatrix}\mathbf{0} & \boldsymbol{Q}^{01} & \cdots & \boldsymbol{Q}^{0n}\\ \boldsymbol{Q}^{10} & \mathbf{0} & \ddots & \vdots\\ \vdots & \ddots & \ddots & \\ \boldsymbol{Q}^{n0} & \cdots & & \mathbf{0}\end{pmatrix}\in\mathbb{R}^{(n+1)R\times(n+1)R}\qquad\text{and}\qquad \underline{\boldsymbol{x}} = \begin{pmatrix}\boldsymbol{x}^0\\ \boldsymbol{x}^1\\ \vdots\\ \boldsymbol{x}^n\end{pmatrix}\in\mathbb{R}^{(n+1)R}, \tag{7.13}$$

i.e. the (double) underline refers to a stacked-up vector (matrix) in the replica space. We rewrite the last integral over $x_{\mu:}$:

$$\begin{aligned}\int\prod_a\prod_k\mathrm{d}x^a_{\mu k}\exp\left(-\frac{1}{2}\sum_a\sum_k(x^a_{\mu k})^2+\sum_{a\neq b}\sum_{k,k'}Q^{ab}_{k,k'}x^a_{\mu k}x^b_{\mu k'}\right)\\ =\int\mathrm{d}\underline{\boldsymbol{x}}_\mu\exp\left(-\frac{1}{2}\underline{\boldsymbol{x}}^\intercal_\mu\left(\underline{\underline{\mathbf{1}}}-\underline{\underline{\boldsymbol{Q}}}\right)\underline{\boldsymbol{x}}_\mu\right).\end{aligned} \tag{7.14}$$

Carrying out the Gaußian integral in (7.12), we are left with

$$\mathbb{E}_{\boldsymbol{Y}}\left[\tilde{Z}(\boldsymbol{Y})^n\right]\propto\int\mathrm{d}\underline{\underline{\boldsymbol{Q}}}\exp\left(-N\frac{\Delta}{4}\mathrm{Tr}\underline{\underline{\boldsymbol{Q}}}^2-\frac{N}{2}\ln\det\left(\underline{\underline{\mathbf{1}}}-\underline{\underline{\boldsymbol{Q}}}\right)\right). \tag{7.15}$$

Eventually we would like to perform this integral by a saddle point approximation. However, *we expect the saddle point to be hugely degenerated, at this point, due to the rotational symmetry in the problem*. It is hence reasonable to 'eliminate' this degeneracy by decomposition of the integration over $\boldsymbol{Q}^{ab}$ by an integration over the degenerate part (orthogonal invariance) and the non-degenerate part (the eigenvalues of $\boldsymbol{Q}^{ab}$). This is done by the following orthogonal transformation [Mui09] (recall that $\boldsymbol{Q}^{ab}$ is symmetric)

$$\boldsymbol{Q}^{ab} = \boldsymbol{O}^{ab}\boldsymbol{L}^{ab}(\boldsymbol{O}^{ab})^\intercal\,, \quad \Rightarrow \quad \mathrm{d}\boldsymbol{Q}^{ab} = \underbrace{\frac{2^R\pi^{R^2/2}}{\Gamma_R(\frac{R}{2})}(\mathrm{d}\boldsymbol{O}^{ab})}_{\mathrm{D}\boldsymbol{O}^{ab}} \prod_{i<j}\left|l_i^{ab}-l_j^{ab}\right|(\mathrm{d}\boldsymbol{L}^{ab})\,. \tag{7.16}$$

Here $\boldsymbol{L}^{ab}$ denotes the diagonal matrix of eigenvalues, l_i^{ab}, of $\boldsymbol{Q}^{ab}$ and $\boldsymbol{O}^{ab}$ is an orthogonal matrix. The product $\prod_{i<j}(l_i^{ab}-l_j^{ab})$ is the Vandermonde determinant that induces a repulsive interaction among the R eigenvalues. Note well that we are assuming that the overlap matrix has full rank. Performing the change of variables, we obtain

$$\begin{aligned}\mathbb{E}_{\boldsymbol{Y}}\left[\tilde{Z}(\boldsymbol{Y})^n\right] \propto \int \mathrm{d}\underline{\underline{\boldsymbol{L}}}\exp\left(-N\frac{\Delta}{4}\mathrm{Tr}\underline{\underline{\boldsymbol{L}}}^2+\sum_{a\neq b}\sum_{i<j}\ln\left|l_i^{ab}-l_j^{ab}\right|\right)\\ \int \mathrm{D}\underline{\underline{\boldsymbol{O}}}\exp\left(-\frac{N}{2}\ln\det\left(\underline{\underline{\mathbf{1}}}-\underline{\underline{\boldsymbol{Q}}}\right)\right)\,,\end{aligned} \tag{7.17}$$

where $\underline{\underline{\boldsymbol{Q}}}$ in the last integral should be understood as a function of $\underline{\underline{\boldsymbol{O}}}$ and $\underline{\underline{\boldsymbol{L}}}$. We write the above in the following form

$$\mathbb{E}_{\boldsymbol{Y}}\left[\tilde{Z}(\boldsymbol{Y})^n\right] \propto \int \mathrm{d}\underline{\underline{\boldsymbol{L}}}\,\mathrm{e}^{-N^2S[\underline{\underline{\boldsymbol{L}}}]} \tag{7.18}$$

with the action

$$S[\underline{\underline{\boldsymbol{L}}}] = \frac{\Delta}{4}\frac{1}{N}\sum_{a\neq b}\mathrm{Tr}(\boldsymbol{L}^{ab})^2 - \frac{1}{N^2}\sum_{a\neq b}\sum_{i<j}\ln|l_i^{ab}-l_j^{ab}| - G[\underline{\underline{\boldsymbol{L}}}] \tag{7.19}$$

and $G[\underline{\underline{\boldsymbol{L}}}]$ is defined as follows:

$$G[\underline{\underline{\boldsymbol{L}}}] = \frac{1}{N^2}\ln\{\int \mathrm{D}\underline{\underline{\boldsymbol{O}}}\exp[-\frac{N}{2}\ln\det(\underline{\underline{\mathbf{1}}}-\underline{\underline{\boldsymbol{Q}}})]\} \tag{7.20}$$

where the matrix $\underline{\underline{\boldsymbol{Q}}}$ is a function of $\{\boldsymbol{L}^{ab}\}$ and $\{\boldsymbol{O}^{ab}\}$:

$$\underline{\underline{\boldsymbol{Q}}} = \begin{pmatrix} \mathbf{0} & \boldsymbol{O}^{01}\boldsymbol{L}^{01}(\boldsymbol{O}^{01})^\intercal & \cdots & \boldsymbol{O}^{0n}\boldsymbol{L}^{0n}(\boldsymbol{O}^{0n})^\intercal \\ \boldsymbol{O}^{10}\boldsymbol{L}^{10}(\boldsymbol{O}^{10})^\intercal & \mathbf{0} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \\ \boldsymbol{O}^{n0}\boldsymbol{L}^{n0}(\boldsymbol{O}^{n0})^\intercal & \cdots & & \mathbf{0} \end{pmatrix} \tag{7.21}$$

Next, we evaluate the integral in (7.18) by a **saddle point approximation**. Extremization w.r.t. the eigenvalues, $\frac{\partial S}{\partial l_k^{ab}} \stackrel{!}{=} 0$, leads to the following set of equations

$$\frac{\Delta}{2} l_k^{ab} - \frac{1}{N} \sum_{\substack{1 \le i \le R \\ i \neq k}} \frac{1}{l_k^{ab} - l_i^{ab}} - N \frac{\partial G}{\partial l_k^{ab}} = 0 \,, \qquad \forall k, \forall a \neq b \,, \tag{7.22}$$

with

$$\frac{\partial G}{\partial l_k^{ab}} = \frac{1}{N} \frac{1}{2} \frac{\int \mathrm{D}\underline{\underline{\boldsymbol{O}}} \mathrm{Tr}[(\underline{\underline{\mathbf{1}}} - \underline{\underline{\boldsymbol{Q}}})^{-1} \underline{\underline{\boldsymbol{\delta}}}_{kk}^{ab}] \exp\left(-\frac{N}{2} \ln\det\left(\underline{\underline{\mathbf{1}}} - \underline{\underline{\boldsymbol{Q}}}\right)\right)}{\int \mathrm{D}\underline{\underline{\boldsymbol{O}}} \exp\left(-\frac{N}{2} \ln\det\left(\underline{\underline{\mathbf{1}}} - \underline{\underline{\boldsymbol{Q}}}\right)\right)} \,. \tag{7.23}$$

We have ignored the symmetry under exchange of a and b that would lead to a factor of two in front of each of the terms and

$$\underline{\underline{\boldsymbol{\delta}}}_{kk}^{ab} = \begin{pmatrix} \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \ddots & \mathbf{0} & \mathbf{0} \\ \vdots & \ddots & \mathbf{0} & \boldsymbol{O}^{ab} \boldsymbol{\delta}_{kk} (\boldsymbol{O}^{ab})^{\intercal} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \end{pmatrix}$$

where $\boldsymbol{\delta}_{kk}$ denotes the matrix that carries zeros everywhere, except from position kk where it is equal to unity. It would be difficult, if possible at all, to solve the above equations for a generic $\underline{\underline{\boldsymbol{Q}}}$. Instead we shall be looking for a solution of the saddle point equations that is replica symmetric.

Replica Symmetric Solution

Once the variational equations (7.22) are restricted to a RS solution, life becomes much easier. As we shall show next, the RS Ansatz

$$\underline{\underline{\boldsymbol{Q}}} = \begin{pmatrix} \mathbf{0} & \boldsymbol{O}\boldsymbol{L}\boldsymbol{O}^{\intercal} & \cdots & \boldsymbol{O}\boldsymbol{L}\boldsymbol{O}^{\intercal} \\ \boldsymbol{O}\boldsymbol{L}\boldsymbol{O}^{\intercal} & \mathbf{0} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \\ \boldsymbol{O}\boldsymbol{L}\boldsymbol{O}^{\intercal} & \cdots & & \mathbf{0} \end{pmatrix} , \tag{7.24}$$

will eliminate the difficulties with the integration over the orthogonal group. In order to evaluate the RS saddle point equations (7.22), the first step is to compute $(\underline{\underline{\mathbf{1}}} - \underline{\underline{\boldsymbol{Q}}})^{-1}$. Using the following Ansatz

$$(\underline{\underline{\mathbf{1}}} - \underline{\underline{\boldsymbol{Q}}})^{-1} = \begin{pmatrix} \boldsymbol{A} & \boldsymbol{B} & \cdots & \boldsymbol{B} \\ \boldsymbol{B} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{B} & \cdots & \boldsymbol{B} & \boldsymbol{A} \end{pmatrix} , \tag{7.25}$$

we obtain

$$\boldsymbol{A} = \mathbf{1} - n\boldsymbol{Q}\left[n\boldsymbol{Q}^2 + (n-1)\boldsymbol{Q} - \mathbf{1}\right]^{-1}\boldsymbol{Q} \qquad \xrightarrow{n\to 0} \mathbf{1} \tag{7.26}$$

$$\boldsymbol{B} = \left[n\boldsymbol{Q}^2 - (n-1)\boldsymbol{Q} + \mathbf{1}\right]^{-1}\boldsymbol{Q} \qquad \xrightarrow{n\to 0} (\mathbf{1}+\boldsymbol{Q})^{-1}\boldsymbol{Q}\,. \tag{7.27}$$

Since the term

$$\begin{aligned}\mathrm{Tr}[(\underline{\underline{\mathbf{1}}} - \underline{\underline{\boldsymbol{Q}}})^{-1}\underline{\underline{\boldsymbol{\delta}}}^{ab}_{kk}] &= \mathrm{Tr}\left[[n\boldsymbol{O}\boldsymbol{L}^2\boldsymbol{O}^\intercal - (n-1)\boldsymbol{O}\boldsymbol{L}\boldsymbol{O}^\intercal + \mathbf{1}]^{-1}\boldsymbol{O}\boldsymbol{L}\boldsymbol{O}^\intercal\boldsymbol{O}\boldsymbol{\delta}_{kk}\boldsymbol{O}^\intercal\right] \\ &= \mathrm{Tr}\left[\boldsymbol{O}[n\boldsymbol{L}^2 - (n-1)\boldsymbol{L} + \mathbf{1}]^{-1}\boldsymbol{O}^\intercal\boldsymbol{O}\boldsymbol{L}\boldsymbol{\delta}_{kk}\boldsymbol{O}^\intercal\right] \\ &= \mathrm{Tr}\left[[n\boldsymbol{L}^2 - (n-1)\boldsymbol{L} + \mathbf{1}]^{-1}\boldsymbol{L}\boldsymbol{\delta}_{kk}\right] \\ &= \frac{l_k}{nl_k^2 - (n-1)l_k + 1}\end{aligned}$$

is orthogonally invariant, we can pull it out of the integral. Next we investigate the determinant in the exponent. It is not too difficult to guess the eigenvalues of a $\underline{\underline{\mathbf{1}}} - \underline{\underline{\boldsymbol{Q}}}$:

$$\underline{\boldsymbol{v}}^{(0)} = \begin{pmatrix}\boldsymbol{v}_Q\\ \boldsymbol{v}_Q\\ \boldsymbol{v}_Q\\ \vdots\\ \boldsymbol{v}_Q\end{pmatrix}, \quad \underline{\boldsymbol{v}}^{(1)} = \begin{pmatrix}\boldsymbol{v}_Q\\ -\boldsymbol{v}_Q\\ \mathbf{0}\\ \vdots\\ \mathbf{0}\end{pmatrix}, \quad \underline{\boldsymbol{v}}^{(2)} = \begin{pmatrix}\boldsymbol{v}_Q\\ \mathbf{0}\\ -\boldsymbol{v}_Q\\ \mathbf{0}\\ \vdots\end{pmatrix}, \quad \text{etc.} \tag{7.28}$$

and thus we can decompose the determinant:

$$\begin{aligned}\det\left(\underline{\underline{\mathbf{1}}} - \underline{\underline{\boldsymbol{Q}}}\right) &= \det(\mathbf{1} - n\boldsymbol{Q})\det(\mathbf{1}+\boldsymbol{Q})^n \\ &= \det(\mathbf{1} - n\boldsymbol{L})\det(\mathbf{1}+\boldsymbol{L})^n\end{aligned}\,. \tag{7.29}$$

We can finally write

$$\frac{\partial G}{\partial l_k^{ab}} \overset{\mathrm{RS}}{=} \frac{1}{2}\frac{1}{N}\frac{l_k}{nl_k^2 - (n-1)l_k + 1} \tag{7.30}$$

and

$$G[\boldsymbol{L}] \overset{\mathrm{RS}}{=} -\frac{1}{2}\frac{1}{N}\left[\ln\det(\mathbf{1} - n\boldsymbol{L}) + n\ln\det(\mathbf{1}+\boldsymbol{L})\right]\,. \tag{7.31}$$

So that the RS action takes of the following form

$$\begin{aligned}S_{\mathrm{RS}}[\boldsymbol{L}] = \frac{\Delta}{4}\frac{1}{N}n(n+1)\mathrm{Tr}(\boldsymbol{L}^2) - n(n+1)\frac{1}{N^2}\frac{1}{2}\sum_{i\neq j}\ln|l_i - l_j| \\ + \frac{1}{2}\frac{1}{N}\left[\mathrm{Tr}\ln(\mathbf{1} - n\boldsymbol{L}) + n\mathrm{Tr}\ln(\mathbf{1}+\boldsymbol{L})\right]\,,\end{aligned} \tag{7.32}$$

which should be evaluated together with the the RS saddle point equations

$$\frac{\Delta}{2}l_k - \frac{1}{2}\frac{l_k}{nl_k^2-(n-1)l_k+1} = \frac{1}{N}\sum_{\substack{1\le i\le R\\ i\neq k}}\frac{1}{l_k-l_i}\,, \qquad \forall k\,. \tag{7.33}$$

Before solving the saddle point equations, we recall that we will eventually be interested in evaluating the action (7.19) in the replica limit, $n \to 0$. It therefore suffice to keep the leading order contribution in the above equations

$$S_{\rm RS}[\boldsymbol{L}] = n\cdot\left(\frac{\Delta}{4}\frac{1}{N}\mathrm{Tr}(\boldsymbol{L})^2 - \frac{1}{N^2}\frac{1}{2}\sum_{i\neq j}\ln|l_i-l_j| + \frac{1}{2}\frac{1}{N}\left[-\mathrm{Tr}\boldsymbol{L} + \mathrm{Tr}\ln\left(\mathbf{1}+\boldsymbol{L}\right)\right]\right), \tag{7.34}$$

such that

$$\frac{\Delta}{2}l_k - \frac{1}{2}\frac{l_k}{l_k+1} = \frac{1}{N}\sum_{\substack{1\le i\le R\\ i\neq k}}\frac{1}{l_k-l_i} \qquad \forall k\,. \tag{7.35}$$

Finally, it can be advantageous to solve the above equations in the continuous limit. A nice summary of the procedure can be found in the appendix of [FW07]. We obtain

$$\mathbb{E}_{\boldsymbol{Y}}\left[\tilde{Z}(\boldsymbol{Y})^n\right] \overset{\rm RS}{\propto} \int \mathcal{D}\rho(\lambda)\,\mathrm{e}^{-N^2\,n\,\mathcal{F}_{\rm RS}[\rho(\lambda)]} \tag{7.36}$$

with the free energy (up to $O(\frac{1}{N})$ corrections and constants)

$$\begin{aligned}\mathcal{F}_{\rm RS}[\rho(\lambda)] = {}& \frac{\Delta}{4}\frac{1}{\alpha}\int \mathrm{d}\lambda\rho(\lambda)\,\lambda^2 - \frac{1}{2}\frac{1}{\alpha}\int \mathrm{d}\lambda\rho(\lambda)\,\lambda + \frac{1}{2}\frac{1}{\alpha}\int \mathrm{d}\lambda\rho(\lambda)\ln\left(1+\lambda\right)\\ & - \frac{1}{\alpha^2}\frac{1}{2}\int\int \mathrm{d}\lambda\mathrm{d}\lambda'\rho(\lambda)\rho(\lambda')\ln|\lambda-\lambda'|\end{aligned} \tag{7.37}$$

and the associated saddle point equations

$$\frac{1}{2}\left(\Delta\lambda - \frac{\lambda}{\lambda+1}\right) = \frac{1}{\alpha}\fint \mathrm{d}\lambda'\frac{\rho(\lambda')}{\lambda-\lambda'}\,. \tag{7.38}$$

The x-MMSE can then be obtained as follows. Looking back at equation (7.12), we can extract the covariance between two replicas, $a \neq b$,

$$\begin{aligned}\mathbb{E}\left[\boldsymbol{x}_\mu^a(\boldsymbol{x}_\mu^b)^\intercal\right] &\propto \int \mathrm{d}\underline{\underline{\boldsymbol{Q}}}\exp\left(-N\frac{\Delta}{4}\mathrm{Tr}\underline{\underline{\boldsymbol{Q}}}^2\right)\prod_\mu\int \mathrm{d}\underline{\boldsymbol{x}}_\mu\,\boldsymbol{x}_\mu^a(\boldsymbol{x}_\mu^b)^\intercal\,\exp\left(-\frac{1}{2}\underline{\boldsymbol{x}}_\mu^\intercal\left(\underline{\underline{\mathbf{1}}}-\underline{\underline{\boldsymbol{Q}}}\right)\underline{\boldsymbol{x}}_\mu\right)\\ &\propto \mathbb{E}_{\underline{\underline{\boldsymbol{Q}}}}\left[[\left(\underline{\underline{\mathbf{1}}}-\underline{\underline{\boldsymbol{Q}}}\right)^{-1}]^{ab}\right]\end{aligned}\,. \tag{7.39}$$

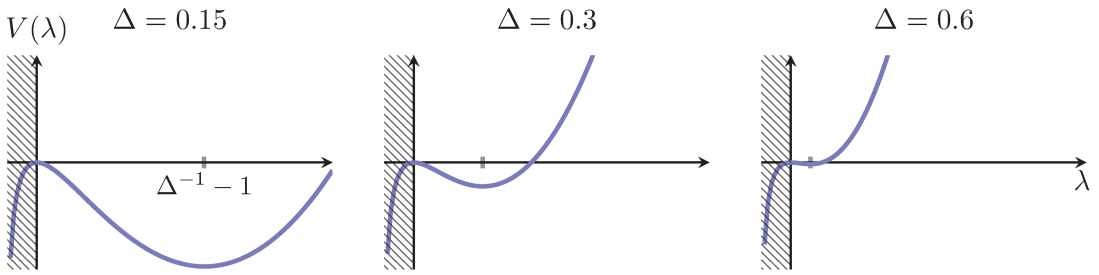

Figure 7.1: The potential, $V(\lambda)$, as seen by a single eigenvalue in blue (Here we only draw the effective, α independent, part). The R eigenvalues experience an additional logarithmic repulsion. The force equilibrium, imposed by potential and repulsion, results in the eigenvalue distribution, $\rho(\lambda)$. The negative part of the potential is not accessible, due to the positivity constraint of the overlap. For small Δ the eigenvalues find enough space around the positive minimum of the potential, while for large Δ they will be pushed against the wall that occupies the negative half-plane.

Since the latter argument in the integral does not influence the saddle point, we can later plug in the solution that we obtain from the saddle point evaluation. Together with (7.27) we can estimate the overlap between two samples from two replicas, $a \neq b$, as

$$m^{ab} := \sum_k x^a_{\mu k} x^b_{\mu k} \simeq \mathrm{Tr}\mathbb{E}_{\underline{\underline{\boldsymbol{Q}}}}\left[\boldsymbol{x}^a_\mu (\boldsymbol{x}^b_\mu)^\intercal\right] = \mathrm{Tr}\left[(\mathbf{1}+\boldsymbol{L})^{-1}\boldsymbol{L}\right] = \sum_k \frac{l_k}{1+l_k}\,. \tag{7.40}$$

The overlap between equal replicas is just one, as it should be. The MMSE on $\boldsymbol{X}$ then follows as

$$x\text{-MMSE} = 2 - 2\int_{\mathrm{supp}(\rho)} \mathrm{d}\lambda \rho(\lambda)\frac{\lambda}{1+\lambda}\,, \tag{7.41}$$

with $\rho(\lambda)$ being the solution of (7.38).

7.3.2 Solving the Replica Symmetric Saddle Point Equations

Before we move on to solve the equations, it is worth taking a look at the Coulomb gas analogy to gain some intuition. We have R eigenvalues, each of them is moving in the potential

$$\frac{1}{2\alpha}\left(\frac{\Delta}{2}\lambda^2 + \ln(1+\lambda) - \lambda\right) \tag{7.42}$$

being subject to a logarithmic repulsion, proportional to $1/\alpha^2$, due to all the other eigenvalues:

$$\frac{1}{\alpha^2}\int \mathrm{d}\lambda' \rho(\lambda') \ln|\lambda - \lambda'|\,. \tag{7.43}$$

Note that the potential has a logarithmic singularity at $\lambda = -1$. In Fig. 7.1 we plot the term in the brackets of (7.42). We ignore the prefactor: the physical information is contained in the force equilibrium $V'(\lambda) = \alpha^{-2} \fint \mathrm{d}\lambda' \rho(\lambda')/(\lambda - \lambda')$. Eliminating the overall factor of

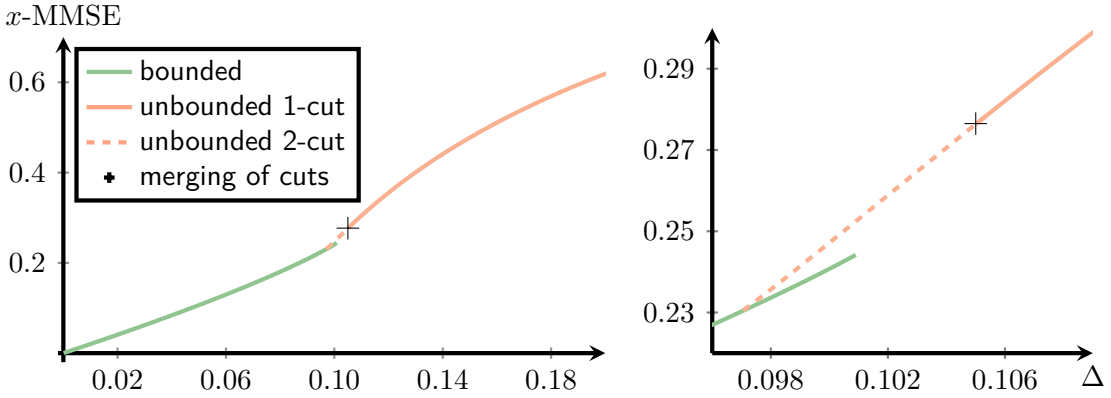

Figure 7.2: Left: the x-MMSE is reported as a function of Δ, for $\alpha = 1$. A phase transition takes place at $\Delta \approx 0.097$, where an unbounded two-cut solution becomes a feasible solution of the saddle point equations. This two-cut solution ceases to exists at $\Delta \approx 0.1055$, where it transitions into an unbounded one-cut solution. Right: a zoom around the region in which the transition takes place. Note the small region of coexistence for $\Delta \in [0.097, 0.1009]$, where we found that the bounded one-cut solution continues to exist.

α^{-1} thus yields the physical part of the potential and does not impact the location of the extrema.

One can now imagine filling up the minima of the potential in Fig. 7.1 with the repulsive eigenvalue gas. It becomes clear that for small values of Δ, the eigenvalues can equilibrate around the (non-singular) minimum on the right, for some fixed α. The induced force equilibrium (7.38), suggests that the effective repulsion is $\Theta(1/\alpha)$ and therefore becomes weaker for increasing α, i.e. the eigenvalue gas can settle more easily in the minima/saddles of the potential. In other words: the barriers, as seen by the eigenvalues, are effectively increased by a factor of α.

Assume some fixed α. When Δ is increased, this minimum wanders to the left and eventually disappears. What happens to the eigenvalues? The figure suggests that they might fall into the singular part of the potential. This, however, would suggest a singular measure and should be rejected as an unphysical solution. Instead, the solution must be restricted to the non-negative part of the real axis. This is further justified by (7.40), that should be non-negative as it corresponds to the overlap/covariance between $\boldsymbol{x}^a$ and $\boldsymbol{x}^b$. Therefore, it is necessary to seek a solution of the saddle point equations that fulfills this condition, i.e., that has non-negativ support.

Before getting into the details of solving the saddle point equations, we briefly anticipate the results, as summarized in Fig. 7.2. We find that the spectrum can be in *three* phases that can be distinguished by the nature of the solution, $\rho(\lambda)$ as follows.

- Single-cut solution, with $\rho(\lambda)$ being supported on the compact interval $[a, b]$. In such a case we find that $\rho(\lambda)$ might fall into one of the following two phases.

a) The bounded phase, in which the solution is bounded on both ends, a and b. In this case $0 < a < b$ and no zero-overlap eigenvalues are present.

b) The unbounded phase, in which $a = 0$ and the solution has an integrable singularity at this end of the support that is associated to zero-overlap eigenvalues.

- Two-cut solution, with support on the two intervals $[0, b]$ and $[c, d]$ and with $\rho(\lambda)$ unbounded at zero, where it exhibits an integrable singularity.

The unbounded two-cut phase separates the bounded one-cut phase (small Δ) from the unbounded one-cut phase (large Δ).

Riemann-Hilbert Ansatz to Solve Singular Integral Equation

In this section we solve the saddle-point equations (7.38) by the method due to Brézin, Itzykson, Parisi and Zuber [BIPZ78], outlined in chapter 3. We recall that, for $\rho(\lambda)$ compactly supported, the resolvent has the following properties:

1) It is analytic in the whole complex plane except from the cuts along the support of ρ (**consistency constraint**).

2) It behaves as $\frac{1}{z}$ for $|z| \to \infty$ (**asymptotic constraint**).

3) It is real for $z \in \mathbb{R} \backslash \mathrm{supp}(\rho)$.

4) It has a jump when approaching the support of ρ from above/below:

$$\mathfrak{g}(x - \mathrm{i}\,0) - \mathfrak{g}(x + \mathrm{i}\,0) = 2\pi\,\mathrm{i}\,\rho(x) \text{ for } x \in \mathrm{supp}(\rho)\,.$$

We have also approached the problem by direct inversion of the singular integral equation via Tricomi's relations. We found that both approaches gave the same result and merely present the former, more intuitive approach. Next to the analytic solutions, we also report the results direct Monte-Carlo simulations of (7.36).

Looking at Fig. 7.1 one might be tempted to assume a single support solution for all values of α and Δ. However, as already anticipated above, this is not the case. Instead, there exists a small region in which a **two-cut** solution exists. We refer to the phases a follows

I The solution is supported along a single cut and is bounded on both ends of it. This is the phase in which there is enough space around the minimum at $\lambda = \Delta^{-1} - 1$ to stably contain all the eigenvalues.

II The solution has support on two separate cuts; the first cut has its left end at zero and the solution is bounded everywhere, except at zero. In this phase, a fraction of the eigenvalues separates from the bulk that is supported around $\lambda = \Delta^{-1} - 1$ and settles around the maximum, close to the wall at $\lambda = 0$.

III The solution has again a single compact support. This time between zero, where it is unbounded, and some finite value to the right of zero. For large Δ (or small α), the potential pushes the right-hand bulk of the two-cut solution (regime II) into the left-hand bulk. The mutual repulsion does not suffice to split the eigenvalue spectrum among the two extrema.

Phase I The following Ansatz for the resolvent is natural

$$\mathfrak{g}(z) = V'(z) - Q(z)\sqrt{(z-a)(z-b)}\,, \tag{7.44}$$

with

$$Q(z) = \frac{c_{-1}}{z+L} + c_0 \quad \text{and} \quad V'(z) = \frac{\alpha}{2}\left(\Delta z + \frac{1}{1+z} - 1\right). \tag{7.45}$$

The resolvent is defined for z in the whole complex plane, except along the cut between $[a,b]$ on the real line. The free parameters, L, c_{-1}, c_0, a, b can now be obtained from the constraints listed above. The consistency constraint suggests that $L = 1$ in order to eliminate the pole of $V'(z)$ at $z = -1$. First we employ the asymptotic constraint: expanding (7.45) at $1/z = 0$ up to $O\left(1/z^2\right)$ leads to

$$\left(\frac{\alpha\Delta}{2} - c_0\right) z + \frac{1}{2}\left(ac_0 - \alpha + bc_0 - 2c_{-1}\right) + \frac{a^2c_0 - 2abc_0 + 4ac_{-1} + 4\alpha + b^2c_0 + 4bc_{-1} + 8c_{-1}}{8z}\,.$$

The first two terms yield

$$c_0 = \frac{\alpha\Delta}{2} \tag{7.46}$$

$$c_{-1} = \frac{\alpha}{4}((a+b)\Delta - 2) \tag{7.47}$$

and from the $O(1/z)$ term we have

$$\frac{\alpha}{4}\left(\frac{\Delta}{4}(3a^2 + 2ab + 4a + 4b + 3b^2) - (a+b)\right) = 1\,. \tag{7.48}$$

Finally, applying the consistency constraint by an expansion around $z = -1$, we obtain a further equation

$$\sqrt{(1+a)(1+b)}((a+b)\Delta - 2) + 2 = 0\,. \tag{7.49}$$

Note the positive sign in front of the square root. This is because $z = -1$ is to the left of the cut $[a,b]$ and therefore the square root picks up a minus sign. The latter two equations can now be used to solve for a and b.

In Fig. 7.3 we compare the solution with Monte Carlo simulations of (7.34) under the measure $\exp(-N^2 S[\boldsymbol{L}])$. The two are in excellent agreement.

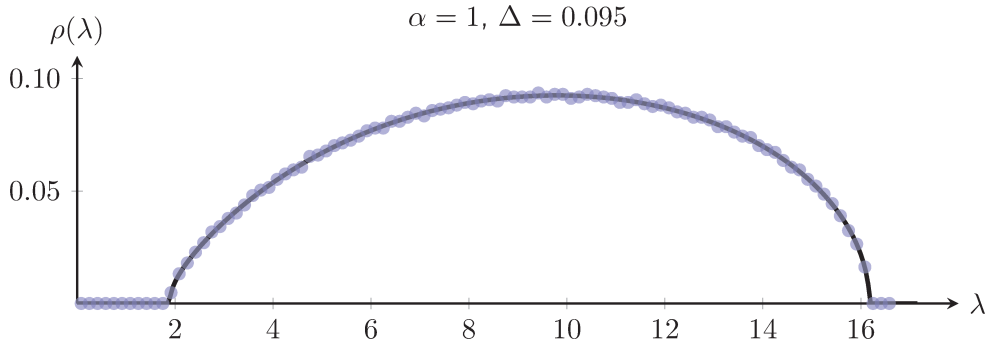

Figure 7.3: The typical spectral density in phase I, where the solution has a single support and is unbounded at both ends. We compare the analytic results with simulations that were carried out for a system with $N = R = 10^5$.

Phase III In this paragraph we investigate the second one-cut solution that appears for large Δ. We expect that the spectrum is unbounded at $\lambda = 0$, where we placed a wall to enforce the non-negativety constraint. Therefore we establish the following Ansatz for the resolvent:

$$\mathfrak{g}(z) = V'(z) - Q(z)\sqrt{\frac{z-b}{z}}\,, \tag{7.50}$$

with $V'(z)$ as before and

$$Q(z) = \frac{c_{-1}}{z+1} + c_0 + c_1 z\,. \tag{7.51}$$

The additional coefficient c_1 is necessary to eliminate the linear term in the expansion at infinity:

$$z\left(\frac{\alpha\Delta}{2} - c_1\right) + \frac{1}{2}\left(-\alpha + bc_1 - 2c_0\right) + \frac{4\alpha + b^2 c_1 + 4bc_0 - 8c_{-1}}{8z}\,.$$

The asymptotic constraint imposes

$$c_1 = \frac{\alpha\Delta}{2} \tag{7.52}$$

$$c_0 = \frac{\alpha}{4}(b\Delta - 2) \tag{7.53}$$

$$c_{-1} = \frac{\alpha}{16}(3b^2\Delta - 4b + 8) - 1\,, \tag{7.54}$$

which leaves us with b as the only free parameter. It can be set by imposing the consistency constraint, which leads to

$$\alpha - \sqrt{1+b}\left(\frac{\alpha}{8}(3b^2\Delta - 4b + 8) - 2\right) = 0\,. \tag{7.55}$$

Note that this time the square root carries a negative sign because the two phases of $z - b$ and z^{-1} cancel out on the real line to the left of the cut. Evaluating these equations leads to

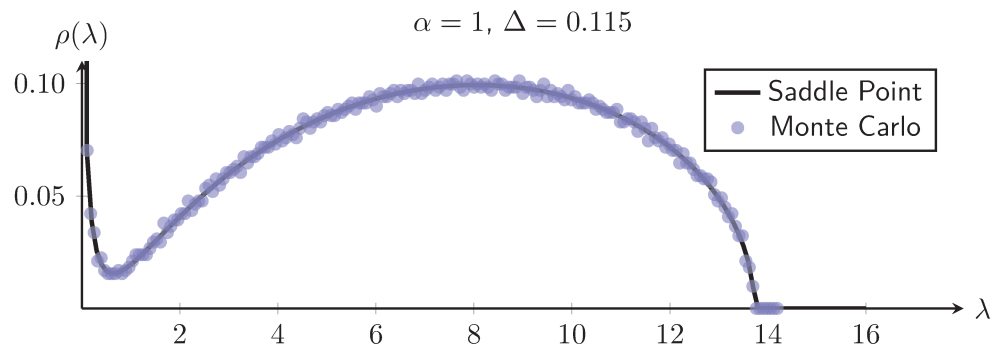

Figure 7.4: In phase III (that is: for large values of Δ) the spectrum has a single support. The non-negativity constraint causes an integrable singularity at zero. This shows up clearly in the Monte Carlo simulation ($N = R = 10^5$).

the spectral density in phase III. Once again, we compare the analytic result with simulations and find that they are in perfect agreement (cf. Fig. 7.4).

Phase II In this phase the spectrum is split and has support along the two cuts $[0, b]$ and $[c, d]$, suggesting that we should adapt our Ansatz for the resolvent as follows

$$\mathfrak{g}(z) = V'(z) - Q(z)\sqrt{\frac{(z-b)(z-c)(z-d)}{z}}\,, \tag{7.56}$$

with

$$Q(z) = \frac{c_{-1}}{z+1} + c_0\,. \tag{7.57}$$

For $b = c = 0$ this reduces to the Ansatz made for the bounded one-cut solution and for $b = c \neq 0$ it is equivalent to the Ansatz made for the one-cut solution unbounded at zero. Note that for the first time there are five parameters, instead of four, that must be determined. This time we exploit the asymptotic expansion

$$\frac{\alpha\Delta - 2c_0}{2}z - \frac{\alpha - c_0(b+c+d) + 2c_{-1}}{2} + \frac{\alpha + c_{-1}(b+c+d+2) + c_0\left(\frac{1}{4}(b+c+d)^2 - (bc+bd+cd)\right)}{2z}$$

to determine b and c (as well as c_0 evidently). After eliminating $c_0 = \alpha\Delta/2$ and enforcing the consistency constraint we obtain the following set of equations

$$\frac{\alpha}{4}\left((b+c+d)\Delta - 2\right) = c_{-1} \tag{7.58}$$

$$\frac{\alpha}{2} + \sqrt{(1+a)(1+b)(1+c)}c_{-1} = 0 \tag{7.59}$$

$$\frac{1}{2}\left(\alpha + c_{-1}(b+c+d+2) + \frac{\alpha\Delta}{2}\left(\frac{1}{4}(b+c+d)^2 - (bc+bd+cd)\right)\right) = 1\,. \tag{7.60}$$

We are facing a problem: the set of equations is underdetermined, the constraints do no longer suffice to determine a unique solution.

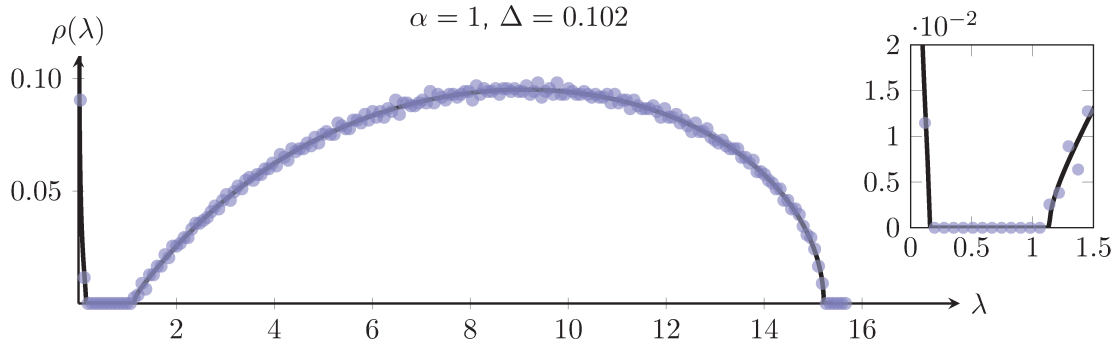

Figure 7.5: There exists an intermediate regime in which the solution of the saddle point equations is given by a two-cut spectrum, unbounded at zero. Although the gap is quite small, it appears clearly in the Monte Carlo simulations ($N = R = 10^5$) as shown in the incept.

We resolve this problem by imposing an additional **stability constraint** that guarantees the distribution to be stable against eigenvalue tunneling from one bulk to the other [Jur90; FSZ01]. This requires the following integral over the interval between the two cuts to vanish

$$\int_b^c \mathrm{d}x\, \rho(x) \stackrel{!}{=} 0\,. \tag{7.61}$$

Plugging in our Ansatz, we obtain (recall property 4 of the resolvent) the necessary additional equation

$$c_{-1}\, K(b,c,d) + \frac{\alpha\Delta}{2} J(b,c,d) = 0\,, \tag{7.62}$$

where

$$J(b,c,d) \coloneqq \int_b^c \mathrm{d}x\, \frac{1}{x}\sqrt{(x-b)(c-x)(d-x)x} \tag{7.63}$$

$$K(b,c,d) \coloneqq \int_b^c \mathrm{d}x\, \frac{1}{x(x+1)}\sqrt{(x-b)(c-x)(d-x)x}\,. \tag{7.64}$$

It is now possible to resolve all the free parameters from the equations (7.58)-(7.60) and (7.62). The comparison with Monte Carlo simulations verifies the result, as shown in Fig. 7.5.

7.3.3 Phase Diagram

We draw the α-Δ phase diagram in Fig. 7.6. There are several remarks to be made:

- There is a small region where the bounded one-cut and the two-cut distribution co-exist as feasible solutions of the saddle point equations. The similarity with the hard phase, encountered in earlier chapters, is very suggestive. However, the nature of this region is slightly different. This question will be discussed further below.

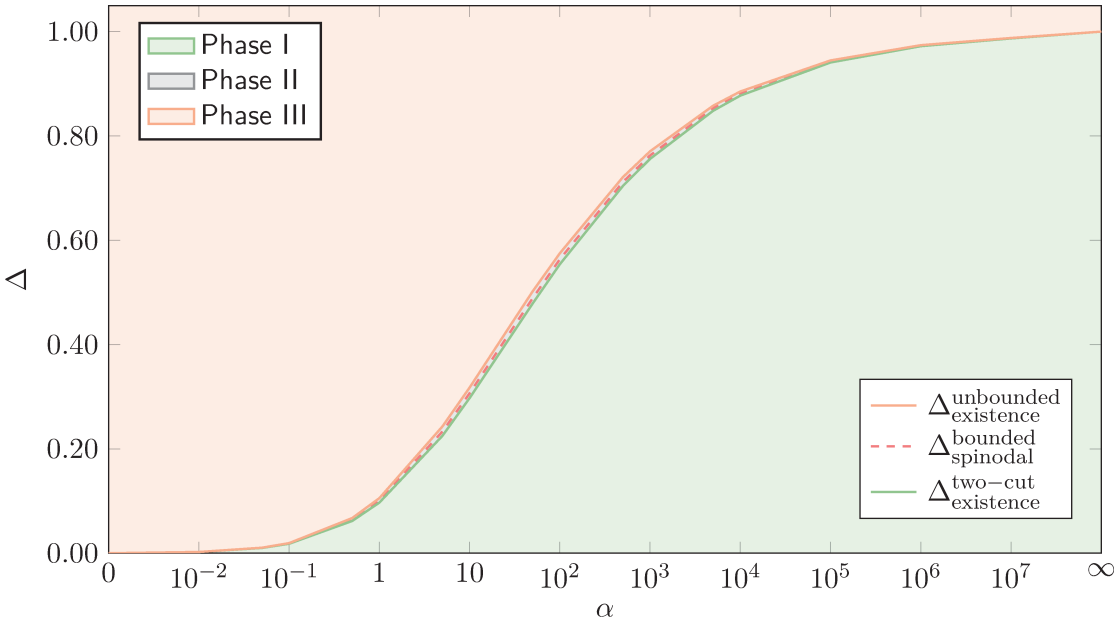

Figure 7.6: In the green-shaded region the solution has bounded positive support (Phase I). The dashed red line marks the spinodal of the bounded one-cut solution. The solid green line indicates the points where a two-cut solution can exist. The small gray-shaded region is marking the area in which a two-cut solution exists that has two bulks: one with unbounded support at zero and the other one with bounded positive support (Phase II). Note that there is a small region of co-existence where both, the two-cut solution and bounded one-cut solution co-exist. The red-orange-shaded region marks the area in which the eigenvalue distribution has a single bulk with unbounded support at zero (Phase III). The solid red line marks the transition at which the two-cut solution continuously transitions into the unbounded one-cut solution.

- The transition from the two-cut to the unbounded one-cut solution appears continuous. However, as for now, it is not clear of what order the transition is. It might be of third order, as encountered in other matrix models [Eyn06; GW80].

- When α and Δ approach the point $(0, 0)$ we find that the region of co-existence vanishes and that the region of existence of the two-cut solution vanishes. This can be explained by the strong repulsions among eigenvalues that render the unbounded part of the two-cut solution (which is located around a maximum) unstable.

- When α is increased, the area of positive overlap increases and inference becomes easier, as one should expect, as the redundancy in the data increases with α. Phase II (as well as the region of co-existence) first grow with α and then shrink again for very large α.

- For large α the area of co-existence shrinks again. And so does the whole phase II. This can be explained by the weaker interaction between the eigenvalues. They continually decouple and $\rho(\lambda)$ concentrates more sharply. The transition happens later (for larger Δ) because the eigenvalues can settle in a smaller region around the extrema.

- In the **low-rank limit**, i.e. when $\alpha \to \infty$, the interaction between the eigenvalues effectively disappears and the problem fully decouples. The saddle-point can be reduced to a single scalar equation

$$\frac{\Delta}{2}\lambda^2 + \ln(1+\lambda) - \lambda = 0\,, \tag{7.65}$$

  from which we recover the second order phase transition of low-rank matrix factorization [LKZ17]. Combining the last equation and (7.40), we obtain

$$\lim_{R\to\infty} \frac{x{-}\mathrm{MMSE}}{2} = \begin{cases} \Delta & \text{for } \Delta \le 1 \\ 1 & \text{for } \Delta > 1 \end{cases}. \tag{7.66}$$

- Finally, we remark that the MMSE only reaches its maximum for $\Delta \to \infty$ (cf. Fig. 7.2), except for $\alpha = \infty$.

The difficulties in providing an answer to the question after the meaning of the different solutions arise mostly from the impossibility to reduce the problem to a single scalar. Unlike in all other inference problems that we have encountered and for which the replica approach is well established, this problem (and therefore a whole other class of problems as well) cannot be reduced to a single scalar. The main questions that we don't consider fully answered are: (a) what is the meaning of the two-cut solution? And (b) what does the unboundedness of the spectrum correspond to?

In order to shine some light on the meaning of the different solutions we start by stating the simplest one. The bounded one-cut solution (Phase I) has a straight forward interpretation.

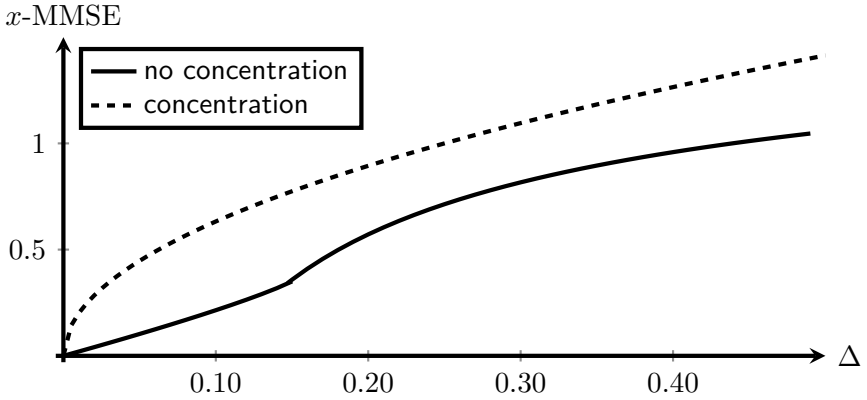

Figure 7.7: A comparison of the x-MMSE prediction for $\alpha = 2$ that results from a computation analogue to that of [KKM+16b] and the approach presented in this chapter. Our analysis predicts a significantly lower achievable x-MMSE. This may result from the correlations in $\boldsymbol{Y}$.

All associated eigenvalues and therefore overlaps are positive. They all contain information, one could talk about an *easy phase*, as encountered in 5. The only difference being that the overlap does not concentrate, even in the thermodynamic limit.

Next, let us discuss the two-cut solution (Phase II). While the bounded one-cut solution is related to purely positive overlap, the two-cut solution is not related to zero-overlap. However, one of its l.h.s. is peaked around zero. We refer to this part of the spectrum as the *less informative* bulk, as it contains eigenvalues that are less informative than those in the r.h.s. bulk (*informative bulk*). However, they are *not* uninformative as they accommodate non-zero eigenvalues. In this phase the data might contain some parts that are easier to infer than others.

The region of co-existence might be related to a similar phenomenon as the hard phase that we encountered in Chap. 5. The idea is simple: the bounded one cut solution is metastable, but can be stabilized with some additional information, just as in the hard phase of crowdsourcing. However, we have not yet investigated this analogy formally. An algorithmic approach would be necessary in order to explore this direction.

Finally, in Phase III, we are facing a similar problem: the bulk is split over a whole compact (non-negative) region and peaked at zero. It appears that still almost all eigenvalues contain some information. It is not possible to talk about in uninformative solution.

7.4 Conclusion

We finish this chapter by concluding that *extensive rank is different*. To the best of our knowledge, all inference problem that were successfully analyzed, did either decouple in some fashion or could be a priori reduced to a single scalar macroscopic order parameter. Take for example the low-rank matrix factorization: for any finite rank the problem fully decouples

(not only for the Gaußian prior). For extensive spin models, such as the above, this is no longer the case. The problem cannot be reduced to a macroscopic order parameter. Instead, it is necessary to work with the microscopic order parameters. Typically, as we have seen, the only way to treat such a problem is to go into the eigenvalue basis of the microscopic order parameters and thereby eliminate the degeneracy in the partition function. As is well established in random matrix theory, one must keep the whole spectrum of eigenvalues to express the free energy.

7.4.1 Open Problems

Many open questions remain:

- How can we verify the results? The free energy of the problem should be accessible through the results of Matytsin [Mat94]. The idea is straight forward: after performing the substitution $\boldsymbol{S} = \boldsymbol{X}\boldsymbol{X}^\intercal$ in (7.6) one has that $\mathrm{d}\boldsymbol{X} \propto (\det \boldsymbol{S})^{(R-N-1)/2}(\boldsymbol{O}^\intercal \mathrm{d}\boldsymbol{O})$ and obtains the Wishart distribution[1] under the integral (after integration over the orthogonal group)

$$Z(\boldsymbol{Y}) \propto \int \mathrm{d}\boldsymbol{S}\, (\det \boldsymbol{S})^{(R-N-1)/2} \exp\left(-\frac{1}{2}\mathrm{Tr}\boldsymbol{S}\right) \exp\left[-\frac{1}{4\Delta}\mathrm{Tr}\,(\boldsymbol{Y}-\boldsymbol{S})^2\right]$$

  The integration over $\boldsymbol{S}$ can then be carried out by substitution of $\boldsymbol{S} \to \boldsymbol{O}\boldsymbol{L}\boldsymbol{O}^\intercal$, which leads to

$$Z(\boldsymbol{Y}) \propto \exp\left(-\frac{1}{4\Delta}\mathrm{Tr}\boldsymbol{Y}^2\right) \int \mathrm{d}\boldsymbol{L} \prod_{i<j}^{N} |l_i - l_j|\, (\det \boldsymbol{L})^{\frac{R-N-1}{2}} \exp\left(-\frac{1}{2}\mathrm{Tr}\boldsymbol{L} - \frac{1}{4\Delta}\mathrm{Tr}\boldsymbol{L}^2\right)$$
$$\int \mathrm{D}\boldsymbol{O} \exp\left(\frac{1}{2\Delta}\mathrm{Tr}\boldsymbol{Y}\boldsymbol{O}\boldsymbol{L}\boldsymbol{O}^\intercal\right)$$

  The last $\boldsymbol{Y}\boldsymbol{O}\boldsymbol{L}\boldsymbol{O}$-integral is better known as the 'Harish-Chandra-Itzykson-Zuber'-(HCIZ) or 'spherical'-integral. In [Mat94] Matytsin has derived an asymptotic expression for the HCIZ integral in terms of the spectral densities of $\boldsymbol{Y}$ and $\boldsymbol{S}$. His result was used to resolve multi-matrix models in QCD. We outline some of our ideas on how this approach can be applied to Bayes-optimal inference problems in Appendix A.7. However this approach seems much more limited in applicability as compared to the replica approach if one has generalizations in mind. In particular one also loses the information about the x-MMSE and can only recover the y-MMSE in this approach. We are wondering in how far such an approach to the denoising problem is universal. The macroscopic (microscopic) eigenvalue statistics of $\boldsymbol{X}\boldsymbol{X}^\intercal$ are universal [TV12], as long as the first two (four) moments are matching. Thus we expect that the best achievable y-MMSE is independent of the prior on the elements of $\boldsymbol{X}$ as long as the first few moments are matching. However, we would not expect the x-MMSE to carry this property. Along the lines of [Ree17; RPD18] it may be possible to extend this approach to resolve the x-MMSE.

[1] For $\alpha \leq 1$, but the results can be extended to $\alpha > 1$.

- How do the results compare to the previous investigation of [KKM+16b]? In Fig. 7.7 we plot exemplarily a comparison of the predicted x-MMSE that results from an analysis based on same steps as in [KKM+16b] and the above approach.

- Related to the last question is also the following problem. From the results of the approach of [KKM+16b] we may expect that some non-trivial transition appears for $\alpha = 1$ that we do not recover in the above approach. Another issue, that may be related to the above, is that there are regions in which the *unbounded one-cut solution behaves unphysical*. If one plots the x-MMSE as a function of α for fixed Δ it grows with α for α around zero. It is possible that the full-rank assumption and/or the non-negativity constraint of the overlap are not valid and cause this unphysical behavior.

As the previous points suggest, the investigation is still rather preliminary and incomplete at this point and it has more academic than practical applicable character at the current point. Nevertheless, we consider it important to make our approach public to show that extensive spin models require quite different tools from what has been established for scalar (or sub-extensive) spins. In that sense, we hope that this analysis may be a first step towards the fusion of the replica method with random matrix theory for inference problems.

Should the above assumptions made in our analysis turn out to be all correct, there still remain several open questions that will hopefully be resolved in the future. First of all, the nature of the transition has to be studied in full detail: including the order of the transition and an investigation of the algorithmic behavior in the mixed phase. In order to resolve these questions it will be necessary to carry out some algorithmic experiments and to extend the analysis to other problems.

— A —

Appendix

A.1 Wormald's General Purpose Theorem

The method of Wormald is outlined in quite some extent in [Wor99b]. The general-purpose theorem is repeated below.

Theorem A.1 (Theorem 5.1 of [Wor99a]). *For $1 \leq l \leq a$, where a is fixed, let $y_l : S^{(N)+} \to \mathbb{R}$ and $f_l : \mathbb{R}^{a+1} \to \mathbb{R}$, such that for some constant C_0 and all l, $|y_l(h_t)| < C_0 n$ for all $h_t \in S^{(N)+}$ for all N. Let Y_l denote the random counterpart of $y_l(h_t)$. Assume the following three conditions hold, where in (ii) and (iii) D is some bounded connected open set containing the closure of*

$$\{(0, z_1, \ldots, z_a) : P\left(Y_l(0) = z_l N,\, 1 \leq l \leq a\right) \neq 0 \; for \; some \; N\}$$

(i) (Boundedness hypothesis.) For some functions $\beta = \beta(N) \geq 1$ and $\gamma = \gamma(N)$, the probability that

$$\max_{1 \leq l \leq a} |Y_l(t+1) - Y_l(t)| \leq \beta\,,$$

conditional uppon H_t, is at least $1 - \gamma$ for $t < T_D$.

(ii) (Trend hypothesis.) For some function $\lambda_1 = \lambda_1(N) = o(1)$, for all $l \leq a$

$$|\mathbb{E}[Y_l(t+1) - Y_l(t) \mid H_t] - f_l(t/N, Y_1(t)/N, \ldots, Y_a(t)/N)| \leq \lambda_1$$

for $t < T_D$.

(iii) (Lipschitz hypothesis.) Each function f_l is continuous, and satisfies a Lipschitz condition, on

$$D \cap \{(t, z_1, \ldots, z_a) : t \geq 0\}\,,$$

with the same Lipschitz constant for each l.

Then the following are true

(a) For $(0, \hat{z}_1, \ldots, \hat{z}_a) \in D$ the system of differential equations

$$\frac{\mathrm{d}z_l}{\mathrm{d}x} = f_l(x, z_1, \ldots, z_a), \quad l = 1, \ldots, a$$

has a unique solution in D for $z_l : \mathbb{R} \to \mathbb{R}$ passing through

$$z_l(0) = \hat{z}_l\,,$$

$1 \leq l \leq a$ and which extends to points arbitrarily close to the boundary of D;

(b) Let $\lambda > \lambda_1 + C_0 N\gamma$ with $\lambda = o(1)$. For a sufficiently large constant C, with probability $1 - O(N\gamma + \frac{\beta}{\lambda}\exp(-N\frac{\lambda^3}{\beta^3}))$,

$$Y_l(t) = N z_l(t/N) + O(\lambda N)$$

uniformly for $0 \leq t \leq \sigma N$ and for each l, where $z_l(x)$ is the solution in (a) with $\hat{z}_l = Y_l(0)/N$, and $\sigma = \sigma(N)$ is the supremum of those x to which the solution can be extended before reaching within l^∞-distance $C\lambda$ of the boundary of D.

Here t indicates the discrete time steps in the random process and $Y_1, \ldots, Y_a$ are the variables of the discrete random process and h_t denotes the history of them, up to time t and $S^{(N)+}$ denotes the set of all h_t. The stopping time $T_D = T_D(Y_1, \ldots, Y_a)$ to be the minimum t such that $(t/N, Y_1(t)/N, \ldots, Y_a(t)/N) \notin D$ for $D \subseteq \mathbb{R}^{a+1}$.

A.2 Proofs for chapter 6

A.2.1 Proof of Lemma 6.1

Based on some worst-case assumptions, we can analyze the evolution of the number of degree-$k_<$ edges. In particular, in the worst case, the initial removal of a degree-d node can generate $d(k-1)$ degree-$k_<$ edges (i.e., all of its edges were attached to degree-k nodes). During removal, the number of degree-k edges can also increase by at most dk. Though both of these events cannot happen simultaneously, our worst-case analysis includes both of these effects.

During the m-th trimming step, a random edge is chosen uniformly from the set of edges adjacent to nodes of degree $k_<$ (i.e., degree less than k). Let the random variable $Z_m \in \{k_<, k, k+1, \ldots, d\}$ equal the degree of the node adjacent to the other end and the random variable X_m equal the overall change in the number of degree-$k_<$ edges.

For $m = 1$, this edge is distributed according to

$$\Pr(Z_1 = z) = \begin{cases} \frac{h'_z}{M-1} & \text{if } z \geq k \\ \frac{h'_{k_<}-1}{M-1} & \text{if } z = k_<, \end{cases}$$

where h'_z denotes the number of edges of degree z before trimming. If $Z_1 = k_<$, then the edge connects two degree-$k_<$ nodes and removal reduces the number of $k_<$ edges by

$X_1 = -2$. If $Z_1 = k$, then the edge connects a degree-k node with a degree-$k_<$ node and removal reduces the number of degree-k edges by k and increases the number of degree-$k_<$ edges by $X_1 = k - 2$. If $Z_1 > k$, then removal decreases the number of degree-Z_1 edges by Z_1 and increases the number of degree-$(Z_1 - 1)$ edges by $Z_1 - 1$. In this case, the number of degree-$k_<$ edges is changed by $X_1 = -1$.

The crux of the analysis below is to make additional worst-case assumptions. One can upper bound the probability of picking degree-k edges because, after m steps of trimming, the number of degree-k edges can be at most $h_k + dk + mk$ (i.e., if initial node removal generates dk edges of degree-k and $Z_m = k + 1$ during each step of trimming). Also, one can upper bound the number of degree-$k_<$ edges because at least one is removed during each step. To upper bound the random variable T, we use the following worst-case distribution for X_m,

$$\Pr(X_m = x) = \begin{cases} \frac{h_k+dk+mk}{M-2d-2m} & \text{if } x = k-2 \\ \frac{M-2d-h_k-dk-mk}{M-2d-2m} & \text{if } x = -1. \end{cases} \tag{A.1}$$

In this formula, the $2d$ term represents initial removal of a degree-d node and the $2m$ term represents the edge removal associated with m steps of trimming. We note that, since edges attach two nodes of different degrees, all edges are counted twice in $\boldsymbol{h}$. Similarly, the dk term represents the worst-case increase in the number of degree-k edges during the initial removal.

Now, we can upper bound the number of degree-$k_<$ edges after m steps of trimming by the random sum

$$S_m = d(k-1) + X_1 + X_2 + \cdots + X_m,$$

where each X_i is drawn independently according to the worst-case distribution (A.1). The term $d(k-1)$ represents the worst-case event that the initial node removal generates $d(k-1)$ edges of degree-$k_<$. By choosing $M_0(\delta)$ appropriately, our initial assumption about $\boldsymbol{h}$ implies that $E[X_i] \le -2\delta$ for all $M > M_0(\delta)$. Thus, the random variable S_m will eventually become zero with probability 1. Since S_m upper bounds the number of degree-$k_<$ edges after m steps of trimming, when it becomes zero (say at time m'), it follows that the stopping time satisfies $T \le m'$.

To complete the proof, we focus on the case of $m' = (k-1)^2\delta^{-2}\ln M$. Since the distribution of X_m gradually places more weight on the event $X_m = k - 2$ as m increases, we can also upper bound $S_{m'}$ by replacing $X_1, \ldots, X_{m'-1}$ by i.i.d. copies of $X_{m'}$. From now on, let Y denote a random variable with the same distribution as $X_{m'}$ and define $S' = Y_1 + Y_2 + \cdots + Y_{m'}$ to be a sum of m' i.i.d. copies of Y. Then, we can write

$$\Pr(T \le m') \ge \Pr(S' \le -d(k-1)) = 1 - \Pr(S' > -d(k-1)).$$

Next, we observe that (A.1) converges in distribution (as $M \to \infty$) to a two-valued random variable Y^* that places probability at most $\frac{1-3\delta}{k-1}$ on the point $k-2$ and the remaining probability on -1. Since $E[Y^*] = -3\delta$, there is an $M_0(\delta)$ such that $E[Y] = -2\delta$ for all

$M > M(\delta)$ Finally, Hoeffding's inequality implies that

$$\Pr(S' > -\delta m') \leq e^{-2\delta^2 m'/(k-1)^2} = M^{-2}.$$

Putting these together, we see that

$$\Pr(T > m') \leq \Pr(S' > -d(k-1)) \leq M^{-2}$$

for $m' > d(k-1)/\delta$. As $m' = (k-1)^2\delta^{-2}\ln M$, the last condition can be satisfied by choosing $M(\delta) > e^{\delta^2 d/(k-1)}$. Since this condition is trivial as $\delta \to 0$, the previous convergence condition will determine the minimum value for $M_0(\delta)$ when δ is sufficiently small.

A.2.2 Proof Sketch for Lemma 6.2

We note that graph processes like COREHD often satisfy the conditions necessary for fluid limits of this type. The only challenge for the COREHD process is that the $\boldsymbol{h}(n)$ process may have unbounded jumps due to the trimming operation. To handle this, we use Lemma 6.1 to show that, with high probability, the trimming process terminates quickly when $\boldsymbol{h}(n)$ satisfies the stated conditions. Then, the fluid limit of the COREHD process follows from [Wor99a, Theorem 5.1].

To apply [Wor99a, Theorem 5.1], we first fix some $\delta > 0$ and choose the valid set D to contain normalized degree distributions satisfying $\eta_k(t) \leq \frac{1-4\delta}{k-1}|\boldsymbol{\eta}(t)|$ and $|\boldsymbol{\eta}(t)| \geq \delta$. This choice ensures that Lemma 6.1 can be applied uniformly for all $\boldsymbol{\eta}(t) \in D$. Now, we describe how the constants are chosen to satisfy the necessary conditions of the theorem and achieve the stated result. For condition (i) of [Wor99a, Theorem 5.1], we will apply Lemma 6.1 but first recall that the initial number of nodes in the graph is denoted by N and hence $|\boldsymbol{\eta}(t)| \geq \delta$ implies $M = |\boldsymbol{h}(n)| \geq BN$ with high probability for some $B > 0$. Thus, condition (i) can be satisfied by choosing[1] $\beta(N) = (k-1)^2\delta^{-2}d\ln N$ applying Lemma 6.1 with $M = BN$ to see that $\gamma(N) = O(N^{-2})$. To verify condition (ii) of the theorem, we note that (6.20) is derived from the large-system limit of the drift and the error $\lambda(n)$ can be shown to be $O(N^{-1})$. To verify condition (iii) of the theorem, we note that that (6.20) is Lipschitz on D. Finally, we choose $\lambda(N) = N^{-1/4} \geq \lambda_1(N) + C_0 N\gamma(N)$ for large enough N. Since

$$\beta(N)/\lambda(N)e^{-N\lambda(N)^3/\beta(N)^3} = O\left(\frac{\ln N}{N^{-1/4}}e^{-AN^{1/4}/(\ln N)^3}\right) = O\left(N^{-1}\right),$$

these choices imply that the COREHD process concentrates around $\boldsymbol{\eta}(t)$ as stated. We only sketch this proof because very similar arguments have used previously for other graph processes [Ach00; Ach01; CM01].

[1] The additional factor of d is required because each trimming step can change the number of degree-j edges by at most d.

A.3 Population Dynamics

Algorithm 4: Population dynamics algorithm for the 1RSB cavity equations for pairwise Markovian random fields.

input : m, $Q(d)$, M, N, T

output: An approximation of $\mathcal{P}(\boldsymbol{m})$ in terms of the messages $\{\boldsymbol{m}_{i,j}\}_{j=1,\dots,N;\, i=1,\dots,M}$

// M is the number of populations and N is the size of each one of them. $Q(d)$ is the excess distribution and m the Parisi-reweighting paramter.

Initialize: $\{\{\boldsymbol{m}\}\}$;

// A single message $\boldsymbol{m}$ is a q component vector.

for $t \leftarrow 1$ **to** T **do**

  for $i \leftarrow 1$ **to** M **do**

    Draw d from $Q(d)$;

    Draw $j_1, \dots, j_d$ uniformly from $\{1, \dots, M\}$;

    for $j \leftarrow 1$ **to** N **do**

      Draw $s_1, \dots, s_d$ uniformly in $\{1, \dots, N\}$;

      $\boldsymbol{m}_{i,j} \leftarrow \mathcal{F}(\boldsymbol{m}_{j_1,s_1}, \dots, \boldsymbol{m}_{j_d,s_d})$;

      $w_{i,j} \leftarrow$ Normalizatin of (1.69) ;

      $w_{i,j} \leftarrow (w_{i,j})^m$;

    $\boldsymbol{m}_{i,1,\dots,N} \leftarrow$ Reweight($\boldsymbol{m}_{i,1,\dots,N}$, $w_{i,1,\dots,N}$) ;

    // The function Reweight() reweights the messages $\boldsymbol{m}_{i,1,\dots,N}$ according to the weights $w_{i,1,\dots,N}$.

A.4 AMP for Crowdsourcing

Owing to the $1/\sqrt{N}$ scaling we may expand $P(L_{ij} \mid \theta_i, v_j)$ in (1.88) up to second order. Denoting

$$S_{ij} := \left.\frac{\partial g(L_{ij}, w)}{\partial w}\right|^{w=0}, \tag{A.2}$$

$$R_{ij} := \left(\left.\frac{\partial g(L_{ij}, w)}{\partial w}\right|^{w=0}\right)^2 + \left.\frac{\partial^2 g(L_{ij}, w)}{\partial w^2}\right|^{w=0}, \tag{A.3}$$

we obtain

$$\mathrm{e}^{g(L_{ij}, w_{ij})} = \mathrm{e}^{g(L_{ij},0)}\left[1 + S_{ij}w_{ij} + \frac{R_{ij}w_{ij}^2}{2} + O(N^{-\frac{3}{2}})\right] = \mathrm{e}^{g(L_{ij},0) + S_{ij}w_{ij} + \frac{1}{2}(R_{ij} - S_{ij}^2)w_{ij}^2} + O(N^{-\frac{3}{2}}). \tag{A.4}$$

For crowdsourcing (1.89) we have

The messages (1.88) can then be written in a simplified Gaußian form

$$\begin{aligned} n_{i\to ij}(\theta_i) &= \frac{1}{Z_\theta^{i\to ij}}\, P_\theta(\theta_i) \exp\left(B_{\theta,\, i\to ij}\, \theta_i - \frac{1}{2} A_{\theta,\, i\to ij}\, \theta_i^2\right) \\ m_{k\to ik}(v_k) &= \frac{1}{Z_v^{k\to ik}}\, P_v(v_k) \exp\left(B_{v,\, k\to ik}\, v_k - \frac{1}{2} A_{v,\, k\to ik}\, v_k^2\right) , \end{aligned} \tag{A.5}$$

where the new variables $A_{\theta,\, v}$ and $B_{\theta,\, v}$ follow the equations

$$\begin{aligned} B^t_{\theta,\, i\to ij} &= \frac{1}{\sqrt{N}} \sum_{k=1,\, k\neq j}^{M} S_{ik} \hat{v}^t_{k\to ik} \\ A^t_{\theta,\, i\to ij} &= \frac{1}{N} \sum_{k=1,\, k\neq j}^{M} \left[S^2_{ik} (\hat{v}^t_{k\to ik})^2 - R_{ik} \left((\hat{v}^t_{k\to ik})^2 + \sigma^t_{v,\, k\to ik}\right)\right] \\ B^t_{v,\, k\to ik} &= \frac{1}{\sqrt{N}} \sum_{l=1,\, l\neq i}^{N} S_{lk} \hat{\theta}^t_{l\to lk} \\ A^t_{\theta,\, k\to ik} &= \frac{1}{N} \sum_{l=1,\, l\neq k}^{N} \left[S^2_{lk} (\hat{\theta}^t_{l\to lk})^2 - R_{lk} \left((\hat{\theta}^t_{l\to lk})^2 + \sigma^t_{\theta,\, l\to lk}\right)\right] . \end{aligned} \tag{A.6}$$

The equations now close on the means and variances of the messages (1.93).

$$\begin{aligned} \hat{\theta}^t_{i\to ij} &= f_\theta\left(A^t_{\theta,\, i\to ij},\ B^t_{\theta,\, i\to ij}\right) \\ \sigma^t_{\theta,\, i\to ij} &= \frac{\partial f_\theta}{\partial B}\left(A^t_{\theta,\, i\to ij},\ B^t_{\theta,\, i\to ij}\right) \\ \hat{v}^{t+1}_{k\to ik} &= f_v\left(A^t_{v,\, k\to ik},\ B^t_{v,\, k\to ik}\right) \\ \sigma^{t+1}_{v,\, k\to ik} &= \frac{\partial f_v}{\partial B}\left(A^t_{v,\, k\to ik},\ B^t_{v,\, k\to ik}\right) , \end{aligned} \tag{A.7}$$

where we have introduced the input functions

$$f_x\,(A,\, B) = \frac{1}{Z_x(A,B)} \int \mathrm{d}x P_x(x) \mathrm{e}^{-\frac{1}{2}Ax^2 + Bx}\, x\,. \tag{A.8}$$

With x indicating either θ or v. These are the relaxed BP (rBP) equations.

The rBP equations are a direct consequence of the CLT. Each message is a random variable and since the $O(N)$ incoming messages are only weakly correlated they result in an effective Gaußian field, acting on each variable node. This field is additionally weighted with the prior

on each of the sides. The outgoing messages also only weakly depend on the target node:

$$\begin{aligned}
B^t_{\theta,\,i} &\coloneqq \frac{1}{\sqrt{N}}\sum_{k=1}^{M} S_{ik}\hat{v}^t_{k\to ik} && = B^t_{\theta,\,i\to ij} + \frac{1}{\sqrt{N}} S_{ij}\hat{v}^t_{j\to ij}\\
A^t_{\theta,\,i} &\coloneqq \frac{1}{N}\sum_{k=1}^{M}\left[S^2_{ik}(\hat{v}^t_{k\to ik})^2 - R_{ik}\left((\hat{v}^t_{k\to ik})^2 + \sigma^t_{v,\,k\to ik}\right)\right] && = A^t_{\theta,\,i\to ij} + O(\frac{1}{N})\\
B^t_{v,\,k} &\coloneqq \frac{1}{\sqrt{N}}\sum_{l=1}^{N} S_{lk}\hat{\theta}^t_{l\to lk} && = B^t_{v,\,k\to ik} + \frac{1}{\sqrt{N}} S_{ik}\hat{\theta}^t_{i\to ik}\\
A^t_{\theta,\,k} &\coloneqq \frac{1}{N}\sum_{l=1}^{N}\left[S^2_{lk}(\hat{\theta}^t_{l\to lk})^2 - R_{lk}\left((\hat{\theta}^t_{l\to lk})^2 + \sigma^t_{\theta,\,l\to lk}\right)\right] && = A^t_{\theta,\,k\to ik} + O(\frac{1}{N})\,.
\end{aligned} \tag{A.9}$$

We can now express the marginals in terms of the messages

$$\begin{aligned}
\hat{\theta}^t_i &= f_\theta\left(A^t_{\theta,\,i},\, B^t_{\theta,\,i}\right) = \hat{\theta}^t_{i\to ij} + \sigma^t_{v,\,i\to ij}\frac{1}{\sqrt{N}} S_{ij}\hat{v}^t_{j\to ij} + O(\frac{1}{N})\\
&= \hat{\theta}^t_{i\to ij} + \sigma^t_{v,\,i}\frac{1}{\sqrt{N}} S_{ij}\hat{v}^t_{j} + O(\frac{1}{N})\\
\hat{v}^t_k &= f_v\left(A^{t-1}_{v,\,k},\, B^{t-1}_{v,\,k}\right) = \hat{v}^t_{k\to ik} + \sigma^t_{\theta,\,k\to ik}\frac{1}{\sqrt{N}} S_{ik}\hat{\theta}^{t-1}_{i\to ik} + O(\frac{1}{N})\\
&= \hat{v}^t_{k\to ik} + \sigma^t_{\theta,\,k}\frac{1}{\sqrt{N}} S_{ik}\hat{\theta}^{t-1}_{i} + O(\frac{1}{N})\,.
\end{aligned} \tag{A.10}$$

Similarly $\sigma^t_{\theta,\,i} = \sigma^t_{\theta,\,i\to ij} + O(1/N)$ and $\sigma^t_{v,\,k} = \sigma^t_{v,\,k\to ik} + O(1/N)$. This process is sometimes referred to as TAPyfication. We finally obtain a set of equations that is independent of the messages and only depends on the marginals by plugging (A.10) back into (A.9)

$$\begin{aligned}
B^t_{\theta,\,i} &= \frac{1}{\sqrt{N}}\sum_{k=1}^{M} S_{ik}\hat{v}^t_k - \left(\frac{1}{N}\sum_{k=1}^{M} S^2_{ik}\sigma^t_{v,\,k}\right)\hat{\theta}^{t-1}_i\\
A^t_{\theta,\,i} &= \frac{1}{N}\sum_{k=1}^{M}\left[S^2_{ik}(\hat{v}^t_k)^2 - R_{ik}\left((\hat{v}^t_k)^2 + \sigma^t_{v,\,k}\right)\right]\\
\hat{\theta}^t_i &= f_\theta\left(A^t_{\theta,\,i},\, B^t_{\theta,\,i}\right)\\
\sigma^t_{\theta,\,i} &= \frac{\partial f_\theta}{\partial B}\left(A^t_{\theta,\,i},\; B^t_{\theta,\,i}\right)\\
B^t_{v,\,k} &= \frac{1}{\sqrt{N}}\sum_{l=1}^{N} S_{lk}\hat{\theta}^t_l - \left(\frac{1}{N}\sum_{l=1}^{N} S^2_{lk}\sigma^t_{\theta,\,l}\right)\hat{v}^t_k\\
A^t_{\theta,\,k} &= \frac{1}{N}\sum_{l=1}^{N}\left[S^2_{lk}(\hat{\theta}^t_l)^2 - R_{lk}\left((\hat{\theta}^t_l)^2 + \sigma^t_{\theta,\,l}\right)\right]\\
\hat{v}^{t+1}_k &= f_v\left(A^t_{v,\,k},\, B^t_{v,\,k}\right)\\
\sigma^{t+1}_{v,\,k} &= \frac{\partial f_v}{\partial B}\left(A^t_{v,\,k},\; B^t_{v,\,k}\right)\,.
\end{aligned} \tag{A.11}$$

These are the AMP equations.

A.5 State Evolution for Crowdsourcing

The AMP equations depend explicitly on the realization of the disorder (for the crowdsourcing this is L_{ij}) through S_{ij} and possibly R_{ij}. Therefore the $B^t_{\theta,\,i}$, $B^t_{v,\,j}$ and $A^t_{\theta,\,i}$, $A^t_{v,\,j}$ are random variables. Let us consider equations (A.9) in order to derive their distributions. Recalling that the different messages, incoming to one node are independent by BP assumption we can apply the CLT to the sums on the r.h.s of the equations for $B^t_{\theta,\,i}$, $B^t_{v,\,j}$ in (A.9). The mentioned independence holds only approximately because the underlying graph is not a tree, but on account of the $O(1/\sqrt{N})$ scaling this suffices in the $N \to \infty$ limit. Thus we have

$$B^t_{\theta,\,i} \sim \mathcal{N}\left(\mathbb{E}B^t_{\theta,\,i},\, \mathbb{E}(B^t_{\theta,\,i})^2 - (\mathbb{E}B^t_{\theta,\,i})^2\right) \tag{A.12}$$

$$B^t_{v,\,j} \sim \mathcal{N}\left(\mathbb{E}B^t_{v,\,j},\, \mathbb{E}(B^t_{v,\,j})^2 - (\mathbb{E}B^t_{v,\,j})^2\right)\,. \tag{A.13}$$

Further more, by the law of large numbers, the r.h.s. of the equation for $A^t_{\theta,\,i}$, $A^t_{v,\,j}$ can be replaced by its average; we obtain

$$A^t_{\theta,\,i} \overset{N\to\infty}{\to} \mathbb{E}A^t_{\theta,\,i} \tag{A.14}$$

$$A^t_{v,\,j} \overset{N\to\infty}{\to} \mathbb{E}A^t_{v,\,j}\,. \tag{A.15}$$

It remains to compute the first two moments of $B^t_{\theta,\,i}$, $B^t_{v,\,j}$ and the first moment of $A^t_{\theta,\,i}$, $A^t_{v,\,j}$. We introduce the following order parameters that will turn up naturally during the computation

$$\begin{aligned}
M^t_\theta &= \frac{1}{N}\sum_{i=1}^{N}\hat{\theta}^t_i\theta^0_i\,, & M^t_v &= \frac{1}{M}\sum_{j=1}^{M}\hat{v}^t_j v^0_j\,,\\
Q^t_\theta &= \frac{1}{N}\sum_{i=1}^{N}(\hat{\theta}^t_i)^2\,, & Q^t_v &= \frac{1}{M}\sum_{j=1}^{M}(\hat{v}^t_j)^2\,,\\
\Sigma^t_\theta &= \frac{1}{N}\sum_{i=1}^{N}\sigma^t_{\theta,\,i}\,, & \Sigma^t_v &= \frac{1}{M}\sum_{j=1}^{M}\sigma^t_{v,\,j}\,.
\end{aligned} \tag{A.16}$$

We start by consideration of the first moment of $B^t_{\theta,\,i}$

$$\mathbb{E}\left[B^t_{\theta,\,i}\right] = \frac{1}{\sqrt{N}}\sum_{k=1}^{M}\int \mathrm{d}L_{ik}P_0\left(L_{ik}\mid w^0_{ik}\right)S_{ik}\hat{v}^t_{k\to ik}\,. \tag{A.17}$$

Expansion w.r.t. w^0_{ik} leads to

$$\mathbb{E}\left[B^t_{\theta,\,i}\right] = \frac{1}{\sqrt{N}}\sum_{k=1}^{M}\int \mathrm{d}L_{ik}P_0\left(L_{ik}\mid 0\right)\left[1 + \frac{\theta^0_i v^0_j}{\sqrt{N}}\left.\frac{\partial \log P_0(L_{ik}\mid w)}{\partial w}\right|_0 + O(\frac{1}{N})\right]S_{ik}\hat{v}^t_{k\to ik}\,, \tag{A.18}$$

which can be further simplified because the first order term vanishes

$$\int \mathrm{d}L_{ik} P_0\left(L_{ik} \mid 0\right) S_{ik} = 0\,. \tag{A.19}$$

In the Bayes optimal setting this follows from

$$\int \mathrm{d}L_{ik} P(L_{ik} \mid w) = 1 \Rightarrow \int \mathrm{d}L_{ik} \frac{\partial}{\partial w} P(L_{ik} \mid w) = 0\,, \tag{A.20}$$

because

$$\frac{\partial}{\partial w} P(L_{ik} \mid w)|_{w=0} = P(L_{ik} \mid 0) \frac{\partial}{\partial w} \log P(L_{ik} \mid w)|_{w=0} \tag{A.21}$$

and thus

$$\int \mathrm{d}L_{ik} P(L_{ik} \mid 0) S_{ik} = \int \mathrm{d}L_{ik} P(L_{ik} \mid 0) \log P(L_{ik} \mid 0)\,. \tag{A.22}$$

If we are not in the Bayes optimal setting, then this still holds, as long as S_{ik} has mean $o(1/\sqrt{N})$, which will be the case in the example of crowdsourcing. Finally, replacing

$$\hat{v}^t_{k\to ik} = \hat{v}^t_k + O(1/\sqrt{N})$$

yields

$$\mathbb{E}\left[B^t_{\theta,\,i}\right] = \frac{\alpha}{\hat{\Delta}} M^t_v \theta^0_i\,, \tag{A.23}$$

where

$$\hat{\Delta}^{-1} = \mathbb{E}_{P_0(L|w^0=0)}\left[\left(\frac{\partial \log P_0(L_{ik} \mid w)}{\partial w}\right)^2\right], \tag{A.24}$$

which will be equal to Δ^{-1} from (1.101) in the Bayes optimal case. The second moment can be computed straight forwardly. To leading order one finds

$$\mathbb{E}\left[(B^t_{\theta,\,i})^2\right] = \frac{\alpha}{\hat{\Delta}} Q^t_v\,. \tag{A.25}$$

An analogue computation can be carried out to obtain

$$\mathbb{E}\left[B^t_{v,\,j}\right] = \frac{1}{\hat{\Delta}} M^t_u v^0_j\,, \tag{A.26}$$

and

$$\mathbb{E}\left[(B^t_{v,\,j})^2\right] = \frac{1}{\hat{\Delta}} Q^t_\theta v^0_j\,. \tag{A.27}$$

After the introduction of

$$\hat{R} = \mathbb{E}_{P_0(L|w^0=0)}\left[\left(\frac{\partial \log P_0(L_{ik} \mid w)}{\partial w}\right)^2 + \frac{\partial^2 \log P_0(L_{ik} \mid w)}{\partial w^2}\right] \tag{A.28}$$

similar arguments lead to an expression for the averages of A^t_θ, A^t_v in terms of order parameters:

$$\begin{aligned}\mathbb{E}\left[A^t_{\theta,\,i}\right] &= \frac{\alpha}{\hat{\Delta}} Q^t_v - \alpha \hat{R}(Q^t_v + \Sigma^t_v) \\ \mathbb{E}\left[A^t_{v,\,j}\right] &= \frac{1}{\hat{\Delta}} Q^t_\theta - \hat{R}(Q^t_\theta + \Sigma^t_\theta)\,.\end{aligned} \tag{A.29}$$

The estimators $\hat{\theta}^t_i$ and $\hat{v}^t_j$ are functions of the random variables $B^t_{\theta,\,i}$ and $B^t_{v,\,j}$ respectively, which distribution is now known in terms of the order parameters. Therefore we do now also know the distributions of $\hat{\theta}^t_i$ and $\hat{v}^t_j$

$$\begin{aligned}\hat{\theta}^t_i &= f_\theta\left(\frac{\alpha\, Q^t_v}{\hat{\Delta}} - \alpha \hat{R}(Q^t_v + \Sigma^t_v),\; \frac{\alpha\, M^t_v}{\hat{\Delta}}\theta^0 + \sqrt{\frac{\alpha\, Q^t_v}{\hat{\Delta}}} W\right) \\ \hat{v}^{t+1}_j &= f_v\left(\frac{Q^t_\theta}{\hat{\Delta}} - \hat{R}(Q^t_\theta + \Sigma^t_\theta),\; \frac{M^t_\theta}{\hat{\Delta}} v^0 + \sqrt{\frac{Q^t_\theta}{\hat{\Delta}}} W\right)\,,\end{aligned} \tag{A.30}$$

where W is a standard normal distributed random variable. The equations can now be closed on the order parameters (A.16) as summarized in the following equations

$$\begin{aligned}M^t_\theta &= \mathbb{E}_{\theta^0,W}\left[f_\theta\left(\frac{\alpha\, Q^t_v}{\hat{\Delta}} - \alpha \hat{R}(Q^t_v + \Sigma^t_v),\; \frac{\alpha\, M^t_v}{\hat{\Delta}}\theta^0 + \sqrt{\frac{\alpha\, Q^t_v}{\hat{\Delta}}} W\right)\,\theta^0\right] \\ Q^t_\theta &= \mathbb{E}_{\theta^0,W}\left[\left(f_\theta\left(\frac{\alpha\, Q^t_v}{\hat{\Delta}} - \alpha \hat{R}(Q^t_v + \Sigma^t_v),\; \frac{\alpha\, M^t_v}{\hat{\Delta}}\theta^0 + \sqrt{\frac{\alpha\, Q^t_v}{\hat{\Delta}}} W\right)\right)^2\right] \\ \Sigma^t_\theta &= \mathbb{E}_{\theta^0,W}\left[\frac{\partial f_\theta}{\partial B}\left(\frac{\alpha\, Q^t_v}{\hat{\Delta}} - \alpha \hat{R}(Q^t_v + \Sigma^t_v),\; \frac{\alpha\, M^t_v}{\hat{\Delta}}\theta^0 + \sqrt{\frac{\alpha\, Q^t_v}{\hat{\Delta}}} W\right)\right] \\ M^{t+1}_v &= \mathbb{E}_{v^0,W}\left[f_v\left(\frac{Q^t_\theta}{\hat{\Delta}} - \hat{R}(Q^t_\theta + \Sigma^t_\theta),\; \frac{M^t_\theta}{\hat{\Delta}} v^0 + \sqrt{\frac{Q^t_\theta}{\hat{\Delta}}} W\right)\, v^0\right] \\ Q^{t+1}_v &= \mathbb{E}_{v^0,W}\left[\left(f_v\left(\frac{Q^t_\theta}{\hat{\Delta}} - \hat{R}(Q^t_\theta + \Sigma^t_\theta),\; \frac{M^t_\theta}{\hat{\Delta}} v^0 + \sqrt{\frac{Q^t_\theta}{\hat{\Delta}}} W\right)\right)^2\right] \\ \Sigma^{t+1}_v &= \mathbb{E}_{v^0,W}\left[\frac{\partial f_\theta}{\partial B}\left(\frac{Q^t_\theta}{\hat{\Delta}} - \hat{R}(Q^t_\theta + \Sigma^t_\theta),\; \frac{M^t_\theta}{\hat{\Delta}} v^0 + \sqrt{\frac{Q^t_\theta}{\hat{\Delta}}} W\right)\right]\,.\end{aligned} \tag{A.31}$$

These equations track the evolution of the AMP equations (1.97).

In the Bayes optimal setting $\hat{\Delta} = \Delta$ and $\hat{R} = R$, given in (1.101). Further more, the set of order parameters can be reduced because $M^t_x = Q^t_x$ and $\Sigma^t_x = \mathbb{E}_{x^0}[(x^0)^2] - Q^t_x$, where x

stands for either θ or v. The above equations simplify to

$$M_\theta^t = \mathbb{E}_{\theta^0,W}\left[f_\theta\left(\frac{\alpha\, M_v^t}{\Delta},\ \frac{\alpha\, M_v^t}{\Delta}\theta^0 + \sqrt{\frac{\alpha\, M_v^t}{\Delta}}W\right)\ \theta^0\right] \tag{A.32}$$

$$M_v^t = \mathbb{E}_{v^0,W}\left[f_v\left(\frac{M_\theta^t}{\Delta},\ \frac{M_\theta^t}{\Delta}v^0 + \sqrt{\frac{M_\theta^t}{\Delta}}W\right)\ v^0\right]. \tag{A.33}$$

A.6 The Free Energy of Crowdsourcing

In this appendix we derive the *Bayes optimal* replica free energy. As a consequence of the Bayes optimal setting, the replica symmetry will hold.

A.6.1 Channel Universality

The $g(\boldsymbol{L}, \boldsymbol{w}^a)$ can be expanded elementwise w.r.t. $w_{ij}^a = O(1/\sqrt{N})$. This is known as the **channel universality** [LKZ15] because the resulting equations are always Gaußian with the effective noise, Δ, given by (1.101). The expansion yields

$$\begin{aligned}\mathbb{E}_{\boldsymbol{L}}\left[\tilde{Z}(\boldsymbol{L})^n\right] &\simeq \int \prod_{a=0}^{n} \mathrm{D}\boldsymbol{\theta}^a \mathrm{D}\boldsymbol{v}^a \int \mathrm{d}\boldsymbol{L} \prod_{i,j} \mathrm{e}^{g(L_{ij},0)+\sum_{a=0}^{n}\left[\partial_w g(L_{ij},0)\cdot w_{ij}^a + \frac{1}{2}\partial_w^2 g(L_{ij},0)\cdot (w_{ij}^a)^2\right]} \\ &= \int \prod_{a=0}^{n} \mathrm{D}\boldsymbol{\theta}^a \mathrm{D}\boldsymbol{v}^a \int \mathrm{d}\boldsymbol{L} \prod_{i,j} P(L_{ij} \mid 0)\, \mathrm{e}^{\sum_{a=0}^{n}\left[\partial_w g(L_{ij},0)\cdot w_{ij}^a + \frac{1}{2}\partial_w^2 g(L_{ij},0)\cdot (w_{ij}^a)^2\right]}\end{aligned} \tag{A.34}$$

Expanding the exponential we can carry out the integration over $\boldsymbol{L}$. Using the definitions, presented in (1.101) and recalled here for convenience:

$$S_{ij} \coloneqq \left.\frac{\partial g(L_{ij}, w)}{\partial w}\right|^{w=0}, \tag{A.35}$$

$$R_{ij} \coloneqq \left(\left.\frac{\partial g(L_{ij}, w)}{\partial w}\right|^{w=0}\right)^2 + \left.\frac{\partial^2 g(L_{ij}, w)}{\partial w^2}\right|^{w=0} \tag{A.36}$$

and

$$\begin{aligned}\Delta^{-1} &= \mathbb{E}_{P(L_{ij}|w_{ij}=0)}\left[S_{ij}^2\right] \\ R &= \mathbb{E}_{P(L_{ij}|w_{ij}=0)}\left[R_{ij}\right].\end{aligned}$$

Since $R = 0$ in the Bayes optimal setting, we obtain

$$\begin{aligned}
\mathbb{E}_{\boldsymbol{L}}\left[\tilde{Z}(\boldsymbol{L})^n\right] &= \int \prod_{a=0}^{n} \mathrm{D}\boldsymbol{\theta}^a \mathrm{D}\boldsymbol{v}^a \int \mathrm{d}\boldsymbol{L} \prod_{i,j} P(L_{ij} \mid 0) \\
&\quad \left[1 + \partial_w g(L_{ij}, 0) \sum_{a=0}^{n} w_{ij}^a + \right. \\
&\quad \left. + \frac{1}{2}\partial_w^2 g(L_{ij}, 0) \sum_{a=0}^{n} (w_{ij}^a)^2 + \frac{1}{2}(\partial_w g(L_{ij}, 0))^2 \sum_{a,b=0}^{n} w_{ij}^a w_{ij}^b + O(N^{-3/2})\right] \\
&= \int \prod_{a=0}^{n} \mathrm{D}\boldsymbol{\theta}^a \mathrm{D}\boldsymbol{v}^a \prod_{i,j} \left[1 + \frac{1}{2\Delta} \sum_{a \neq b=0}^{n} w_{ij}^a w_{ij}^b + O(N^{-3/2})\right] \\
&= \int \prod_{a=0}^{n} \mathrm{D}\boldsymbol{\theta}^a \mathrm{D}\boldsymbol{v}^a \mathrm{e}^{\frac{1}{2\Delta}\frac{1}{N}\sum_{a\neq b}^{n}\sum_{i,j}\theta_i^a v_j^a \theta_i^b v_j^b} + O(N^{-1/2})\,.
\end{aligned} \tag{A.37}$$

A.6.2 Replica Action

The two order parameters are emerging

$$q_\theta^{ab} = \frac{1}{N}\sum_{i=1}^{N} \theta_i^a \theta_i^b \quad \text{and} \quad q_v^{ab} = \frac{1}{M}\sum_{j=1}^{M} v_j^a v_j^b\,. \tag{A.38}$$

They can be introduced in the integral by a delta function to obtain

$$\mathbb{E}_{\boldsymbol{L}}\left[\tilde{Z}(\boldsymbol{L})^n\right] = \int \prod_{0\leq a\leq b\leq n} \mathrm{d}q_\theta^{ab} \mathrm{d}q_v^{ab} \mathrm{e}^{\frac{1}{2\Delta}\frac{MN}{N}\sum_{0\leq a\leq n,\, b\neq a} q_\theta^{ab} q_v^{ab}} I_\theta(\{q_\theta^{ab}\}) I_v(\{q_v^{ab}\})\,. \tag{A.39}$$

Here we absorbed all the contribution from the priors in the two functions $I_\theta(\{q_\theta^{ab}\})$, $I_v(\{q_v^{ab}\})$. We have

$$I_\theta(\{q_\theta^{ab}\}) = N^{\frac{(n+1)(n+2)}{2}} \int \prod_{a=0}^{n}\prod_{i=1}^{N} \mathrm{D}\theta_i^a \prod_{0\leq a<b\leq n} \delta\left(N q_\theta^{ab} - \sum_{i=1}^{N} \theta_i^a \theta_i^b\right) \tag{A.40}$$

$$= \left(\frac{N}{2\pi}\right)^{\frac{(n+1)(n+2)}{2}} \int \prod_{0\leq a<b\leq n} \mathrm{d}\hat{q}_\theta^{ab} \int \prod_{a=0}^{n}\prod_{i=1}^{N} \mathrm{D}\theta_i^a \, \mathrm{e}^{\mathrm{i}\sum_{a<b}\hat{q}_\theta^{ab}\left(N q_\theta^{ab} - \sum_{i=1}^{N}\theta_i^a\theta_i^b\right)} \tag{A.41}$$

$$= \left(\frac{N}{2\pi}\right)^{\frac{(n+1)(n+2)}{2}} \int \prod_{0\leq a<b\leq n} \mathrm{d}\hat{q}_\theta^{ab} \mathrm{e}^{\mathrm{i}\,N\sum_{a<b}\hat{q}_\theta^{ab} q_\theta^{ab}} \left[\int \prod_{a=0}^{n} \mathrm{D}\theta^a \, \mathrm{e}^{-\mathrm{i}\sum_{a<b}\hat{q}_\theta^{ab}\theta^a\theta^b}\right]^N, \tag{A.42}$$

where the Fourier representation of the delta function was used. Similarly:

$$I_v(\{q_v^{ab}\}) = (\frac{M}{2\pi})^{\frac{(n+1)(n+2)}{2}} \int \prod_{0\le a\le b\le n} \mathrm{d}\hat{q}_v^{ab} \mathrm{e}^{\mathrm{i}\, M \sum_{a<b} \hat{q}_v^{ab} q_v^{ab}} \left[\int \prod_{a=0}^{n} \mathrm{D}v^a \, \mathrm{e}^{-\mathrm{i} \sum_{a<b} \hat{q}_v^{ab} v^a v^b} \right]^M \tag{A.43}$$

We can now write

$$\mathbb{E}_{\boldsymbol{L}} \left[\tilde{Z}(\boldsymbol{L})^n \right] = \int \prod_{0\le a<b\le n} \mathrm{d}q_\theta^{ab} \mathrm{d}q_v^{ab} \mathrm{d}\hat{q}_\theta^{ab} \mathrm{d}\hat{q}_v^{ab} \, \mathrm{e}^{NS(\{q_\theta^{ab}, q_v^{ab}, \hat{q}_\theta^{ab}, \hat{q}_v^{ab}\})} \,. \tag{A.44}$$

In the thermodynamic limit this integral will be solved by the saddle point method. It is convenient to rotate the integration over the auxiliary variables $\hat{q}_x^{ab}$ onto the imaginary axis and write

$$\tilde{q}_x^{ab} = -\,\mathrm{i}\, \hat{q}_x^{ab} \,. \tag{A.45}$$

The "action" reads

$$\begin{aligned} S(\{q_\theta^{ab}, q_v^{ab}, \tilde{q}_\theta^{ab}, \tilde{q}_v^{ab}\}) = \frac{\alpha}{\Delta} \sum_{a<b} q_\theta^{ab} q_v^{ab} - \sum_{a<b} (\tilde{q}_\theta^{ab} q_\theta^{ab} + \alpha \tilde{q}_v^{ab} q_v^{ab}) + \\ \ln \left[\int \prod_{a=0}^{n} \mathrm{D}\theta^a \mathrm{e}^{\sum_{a<b} \tilde{q}_\theta^{ab} \theta^a \theta^b} \right] + \alpha \ln \left[\int \prod_{a=0}^{n} \mathrm{D}v^a \mathrm{e}^{\sum_{a<b} \tilde{q}_v^{ab} v^a v^b} \right] . \end{aligned} \tag{A.46}$$

It simplifies in the replica symmetric setting, where ($x \in \{\theta, v\}$)

$$q_x^{ab} = q_x \tag{A.47}$$

We obtain

$$\begin{aligned} S(q_\theta, q_v, \tilde{q}_\theta, \tilde{q}_v) = \frac{\alpha}{\Delta} \frac{n(n+1)}{2} q_\theta q_v + \\ - \frac{n(n+1)}{2} (\tilde{q}_\theta q_\theta + \alpha \tilde{q}_v q_v) \\ + \ln \left[\int \prod_{a=0}^{n} \mathrm{D}\theta^a \mathrm{e}^{-\tilde{q}_\theta \sum_{a>0}^{n} (\theta^a)^2 + \tilde{q}_\theta \left(\sum_{a>0}^{n} \theta^a \right)^2} \right] \\ + \alpha \ln \left[\int \prod_{a=0}^{n} \mathrm{D}v^a \mathrm{e}^{-\tilde{q}_v \sum_{a>0}^{n} (v^a)^2 + \tilde{q}_v \left(\sum_{a>0}^{n} v^a \right)^2} \right] . \end{aligned} \tag{A.48}$$

The terms that contain $(\sum_a x^a)^2$ can be eliminated by a Hubbard-Stratonovich transformation,

$$\mathrm{e}^{\left(\sum_{a=1}^{n} \sqrt{\tilde{q}_x} x^a\right)^2} = \frac{1}{\sqrt{2\pi}} \int \mathrm{d}W \mathrm{e}^{-\frac{W^2}{2} + \sqrt{\tilde{q}_x} W \sum_{a=1}^{n} x^a} \,, \tag{A.49}$$

such that

$$\begin{aligned} S(q_\theta, q_v, \tilde{q}_\theta, \tilde{q}_v) = \frac{\alpha}{\Delta}\frac{n(n+1)}{2} q_\theta q_v - \frac{n(n+1)}{2}(\tilde{q}_\theta q_\theta + \alpha \tilde{q}_v q_v) \\ + \ln\left[\int \mathrm{D}W \mathrm{D}\theta \left[\int \mathrm{D}\theta \mathrm{e}^{-\tilde{q}_\theta \theta^2 + \sqrt{\tilde{q}_\theta} W\theta}\right]^n\right] \\ + \alpha \ln\left[\int \mathrm{D}W \mathrm{D}v \left[\int \mathrm{D}v \mathrm{e}^{-\tilde{q}_v v^2 + \sqrt{\tilde{q}_v} W v}\right]^n\right]. \end{aligned} \tag{A.50}$$

Where $\mathrm{D}W = \mathrm{d}W \exp(-w^2/2)/\sqrt{2\pi}$ is the standard Gaußian measure.

Extremizing w.r.t. q_x ($x \in \{\theta, v\}$) leads to

$$\tilde{q}_\theta = \frac{\alpha}{\Delta} q_v \tag{A.51}$$

$$\tilde{q}_v = \frac{1}{\Delta} q_\theta \tag{A.52}$$

and the replicated partition function becomes

$$\mathbb{E}_{\boldsymbol{L}}\left[\tilde{Z}(\boldsymbol{L})^n\right] = \int \mathrm{d}q_\theta \mathrm{d}q_v \mathrm{e}^{N S_n(q_\theta, q_v)}, \tag{A.53}$$

with

$$\begin{aligned} S_n(q_\theta, q_v) = -\frac{\alpha}{2}\frac{1}{\Delta} n(n+1) q_\theta q_v + \ln\left[\int \mathrm{D}W \mathrm{D}\theta \left[\int \mathrm{D}\theta \mathrm{e}^{-\frac{\alpha}{\Delta} q_v \theta^2 + \sqrt{\frac{\alpha}{\Delta} q_v} W\theta}\right]^n\right] \\ + \alpha \ln\left[\int \mathrm{D}W \mathrm{D}v \left[\int \mathrm{D}v \mathrm{e}^{-\frac{q_\theta}{\Delta} v^2 + \sqrt{\frac{q_\theta}{\Delta}} W v}\right]^n\right]. \end{aligned} \tag{A.54}$$

In the limit $N \to \infty$ the equation (A.53) can be approximated by steepest descent

$$\mathbb{E}_{\boldsymbol{L}}\left[\tilde{Z}(\boldsymbol{L})^n\right] \backsimeq \mathrm{e}^{N S_n(q_\theta, q_v)}. \tag{A.55}$$

According to the replica trick (2.38), i.e., $\ln Z = \lim_{n\to 0} \partial_n Z^n$, it is necessary to compute the derivative of the above equations. Since we will be taking the limit $n \to 0$, we can replace $\alpha/(2\Delta) n(n+1) q_\theta q_v \to \alpha/(2\Delta) n q_\theta q_v$. Carrying out the derivative w.r.t. n and taking the $n \to 0$ now leads to

$$\begin{aligned} \phi(q_\theta, q_v) = \mathbb{E}_{\boldsymbol{L}}\left[\ln \tilde{Z}(\boldsymbol{L})\right] = \lim_{n\to 0} \frac{\partial}{\partial n} \mathrm{e}^{N S_n(q_\theta, q_v)} \\ = -\frac{\alpha}{2\Delta} q_\theta q_v + \int \mathrm{D}W \mathrm{D}\theta \ln\left[\int \mathrm{D}\theta \mathrm{e}^{-\frac{\alpha}{\Delta} q_v \theta^2 + \sqrt{\frac{\alpha}{\Delta} q_v} W\theta}\right] \\ + \alpha \int \mathrm{D}W \mathrm{D}v \ln\left[\int \mathrm{D}v \mathrm{e}^{-\frac{1}{\Delta} q_\theta v^2 + \sqrt{\frac{1}{\Delta} q_\theta} W v}\right] \end{aligned} \tag{A.56}$$

A.7 Spherical Integrals and Matytsin's Approach

Here we sketch some of our ideas of how to approach the problem from a different angle that is mostly due to the work of Matytsin [Mat94] in combinations with some ideas from free probability. The idea is to study the denoising problem of (7.2). More generally, let us assume that

$$\boldsymbol{Y} = \boldsymbol{S} + \sqrt{\frac{\Delta}{N}}\boldsymbol{W}$$

with $\boldsymbol{Y},\ \boldsymbol{S},\ \boldsymbol{W} \in \mathbb{R}^{N\times N}$ and $\boldsymbol{W}$ a Wigner matrix.

The partition function of this model reads

$$Z(\boldsymbol{Y}) \propto \int \mathrm{d}\boldsymbol{S} P(\boldsymbol{S}) \mathrm{e}^{-\frac{N}{4\Delta}\mathrm{Tr}(\boldsymbol{Y}-\boldsymbol{S})^2}\,. \tag{A.57}$$

For rotationally invariant $P(\boldsymbol{S})$, the transformation

$$\boldsymbol{S} = \boldsymbol{O}\boldsymbol{L}\boldsymbol{O}^\intercal \quad \Rightarrow \quad (\mathrm{d}\boldsymbol{S}) = \underbrace{\frac{2^N \pi^{N^2/2}}{\Gamma_N(\frac{N}{2})}(\mathrm{d}\boldsymbol{O})}_{\mathrm{D}\boldsymbol{O}} \prod_{i<j}(l_i - l_j)(\mathrm{d}\boldsymbol{L}) \tag{A.58}$$

induces a measure on the space of eigenvalues, $\boldsymbol{L}$, and orthogonal matrices, $\boldsymbol{O}$,

$$Z(\boldsymbol{Y}) \propto \mathrm{e}^{-\frac{N}{4\Delta}\mathrm{Tr}\boldsymbol{Y}^2} \int \mathrm{d}\boldsymbol{L} \prod_{i<j}^{N}(l_i - l_j)\, P(\boldsymbol{L}) \mathrm{e}^{-\frac{N}{4\Delta}\mathrm{Tr}\boldsymbol{L}^2} \int \mathrm{D}\boldsymbol{O} \mathrm{e}^{\frac{N}{2\Delta}\mathrm{Tr}\boldsymbol{Y}\boldsymbol{O}\boldsymbol{L}\boldsymbol{O}^\intercal}\,. \tag{A.59}$$

The HCIZ integral over the orthogonal group

$$I_1[\boldsymbol{Y}, \boldsymbol{L}] \coloneqq \int \mathrm{D}\boldsymbol{O}\, \mathrm{e}^{\frac{N}{2\Delta}\mathrm{Tr}\boldsymbol{Y}\boldsymbol{O}\boldsymbol{L}\boldsymbol{O}^\intercal} \tag{A.60}$$

can be resolved, in the limit where $N \to \infty$, by combination of Matytsin's equations [Mat94; BBMP14] and Zuber's $1/2$-rule [Zub08]

$$I_1[\boldsymbol{Y}, \boldsymbol{L}] \xrightarrow{N\to\infty} \exp\left(\frac{N^2}{2} F_2[\rho_{\boldsymbol{S}}, \rho_{\boldsymbol{Y}}; \Delta] + O(1)\right)\,. \tag{A.61}$$

With

$$\begin{aligned} F_2[\rho_{\boldsymbol{S}}, \rho_{\boldsymbol{Y}}; \Delta] = c_\Delta &+ \frac{1}{2\Delta}\left(\int \mathrm{d}y \rho_{\boldsymbol{Y}}(y) y^2 + \int \mathrm{d}s \rho_{\boldsymbol{S}}(s) s^2\right) \\ &- \frac{1}{2}\left(\int \mathrm{d}y \mathrm{d}y' \rho_{\boldsymbol{Y}}(y)\rho_{\boldsymbol{Y}}(y') \ln|y - y'| \int \mathrm{d}s \mathrm{d}s' \rho_{\boldsymbol{S}}(s)\rho_{\boldsymbol{S}}(s') \ln|s - s'|\right) \\ &- S[\rho_{\boldsymbol{S}}, \rho_{\boldsymbol{Y}}; \Delta]\,, \end{aligned} \tag{A.62}$$

where

$$S[\rho_{\boldsymbol{S}}, \rho_{\boldsymbol{Y}}; \Delta] = \frac{1}{2}\int_0^\Delta \mathrm{d}t \int \mathrm{d}x \rho(x,t) \left(v(x,t)^2 + \frac{\pi^2}{3}\rho(x,t)^2 \right) \tag{A.63}$$

and $\rho(x,t)$ obeys the Euler-Matytsin equations:

$$\partial_t \rho(x,t) + \partial_x[\rho(x,t)v(x,t)] = 0 \tag{A.64}$$

$$\partial_t v(x,t) + v(x,t)\partial_x v(x,t) = \pi^2 \rho(x,t)\partial_x \rho(x,t) \tag{A.65}$$

such that the following boundary conditions hold

$$\rho(x, t=0) = \rho_{\boldsymbol{S}}(x) \tag{A.66}$$

$$\rho(x, t=\Delta) = \rho_{\boldsymbol{Y}}(x)\,. \tag{A.67}$$

Substituting $f(x,t) \coloneqq v(x,t) + \mathrm{i}\,\pi\rho(x,t)$ leads to

$$\frac{\partial f}{\partial t} + f\frac{\partial f}{\partial x} = 0\,, \tag{A.68}$$

which can be solved by the method of characteristics which yields

$$f(x,t) = f_0\Big(x - t f(x,t)\Big)\,. \tag{A.69}$$

Where $f_0(x) = f(x, t=0)$.

At the same time, free probability leads to a functional fixed point equation for the resolvent of

$$\boldsymbol{Y}(t) \coloneqq \boldsymbol{S} + \sqrt{\frac{t}{N}}\boldsymbol{W}\,. \tag{A.70}$$

For generic $\boldsymbol{S}$ and $\boldsymbol{W}(t) \coloneqq \sqrt{t/N}\boldsymbol{W}$, with $\boldsymbol{W}$ a standard Wigner matrix, one obtains [BBP17]

$$\mathfrak{g}_{\boldsymbol{Y}(t)}(z) = \mathfrak{g}_{\boldsymbol{S}}\left(z - \mathfrak{R}_{\boldsymbol{W}(t)}\Big(\mathfrak{g}_{\boldsymbol{Y}(t)}(z)\Big)\right)\,.$$

Here $\mathfrak{g}_{\boldsymbol{M}}(z)$ and $\mathfrak{R}_{\boldsymbol{M}}(z)$ are the resolvent and $\mathfrak{R}$-transforms of $\boldsymbol{M}$. Since the $\mathfrak{R}$-transform of a Wigner matrix is simply $\mathfrak{R}_{\boldsymbol{W}(t)}(z) = tz$, these equations simplify and lead to

$$\mathfrak{g}_{\boldsymbol{Y}(t)}(z) = \mathfrak{g}_{\boldsymbol{S}}\Big(z - t\mathfrak{g}_{\boldsymbol{Y}(t)}(z)\Big)\,. \tag{A.71}$$

Therefore, solving the Euler-Matytsin equation reduced to solving the above fixed point equation for the resolvent. In the *Bayes-optimal setting* the Euler-Matytsin equations are thus solved by the resolvent (assuming that the analytical continuation to the complex plane is valid)

$$v(x,t) = \mathrm{Re}\lim_{\eta\to 0^+} \mathfrak{g}_{\boldsymbol{Y}(t)}(x - \mathrm{i}\,\eta) \tag{A.72}$$

$$\rho(x,t) = \frac{1}{\pi}\mathrm{Im}\lim_{\eta\to 0^+} \mathfrak{g}_{\boldsymbol{Y}(t)}(x - \mathrm{i}\,\eta) \tag{A.73}$$

of the matrix $\boldsymbol{Y}(t)$.

A.7.1 Shooting Birds with Cannons

Before we move to the analysis of (7.2), we take a step back and start with an even simpler model, in order to introduce the methods.

We assume that

$$\boldsymbol{S} \in \mathbb{R}^{N\times N}\,, \qquad s_{\mu\nu} = s_{\nu\mu} \sim \begin{cases} \mathcal{N}(0, \frac{\sigma^2}{N}) & \text{if} \mu < \nu \\ \mathcal{N}(0, 2\frac{\sigma^2}{N}) & \text{if} \mu = \nu \end{cases}. \tag{A.74}$$

That is, we observe a Wigner matrix $\boldsymbol{Y}$, that is itself a sum of a Wigner matrix $\boldsymbol{S}$ of variance σ^2, perturbed by another Wigner matrix $\boldsymbol{W}$ of variance Δ; we have

$$\boldsymbol{Y} = \boldsymbol{S} + \sqrt{\frac{\Delta}{N}}\boldsymbol{W}\,. \tag{A.75}$$

The associated rescaled partition function of this model reads

$$\tilde{Z}(\boldsymbol{Y}) = \int \mathrm{d}\boldsymbol{S} \mathrm{e}^{-\frac{N}{4\sigma^2}\mathrm{Tr}\boldsymbol{S}^2} \mathrm{e}^{-\frac{N}{4\Delta}\mathrm{Tr}(\boldsymbol{Y}-\boldsymbol{S})^2} \tag{A.76}$$

Note that we are considering the special partition function, that does not contain the normalizations of the likelihood and prior, as is necessary on order to extract the y-MMSE.

We can evaluate the y-MMSE (which coincides with the x-MMSE in this case) from the I-MMSE relations [GSV05] by simple Gaussian integration.

$$y\text{−MMSE} = \frac{\Delta}{1+\Delta/\sigma^2}\,. \tag{A.77}$$

In this setting, denoising and reconstruction become one and the same problem. It is not difficult to see that the $N(N+1)/2$ independent elements (due to the symmetry in the matrices) fully decouple and we could have written down the result straight away. However, it is instructive to forget about this fact for a moment and to consider it, instead, as a playground to introduce the Matytsin approach to the free energy of this problem. Although this seems like shooting a bird with a cannon, the advantage is that the same cannon may be used in order to attack the more difficult problem (7.2).

Matytsin's Method

The starting point is (A.76). After expanding the square, we have (for the rescaled partition function)

$$\tilde{Z}(\boldsymbol{Y}) = \mathrm{e}^{-\frac{N}{4\Delta}\mathrm{Tr}\boldsymbol{Y}^2} \int \mathrm{d}\boldsymbol{S} \mathrm{e}^{-\frac{N}{4}(\frac{1}{\sigma^2}+\frac{1}{\Delta})\mathrm{Tr}\boldsymbol{S}^2} \mathrm{e}^{-\frac{N}{2\Delta}\mathrm{Tr}\boldsymbol{Y}\boldsymbol{S}}\,. \tag{A.78}$$

In order to demonstrate Matytsin's method, we *ignore that this integral factorizes over the elements of the matrices* $\boldsymbol{S}$ *and* $\boldsymbol{Y}$. Instead, we decompose the integration over $\boldsymbol{S}$ into

an integration over its eigenvalues and eigenvectors. Since $\boldsymbol{S}$ is symmetric we may use the substitution (A.58), which leads to

$$\tilde{Z}(\boldsymbol{Y}) = \mathrm{e}^{-\frac{N}{4\Delta}\mathrm{Tr}\boldsymbol{Y}^2} \int \mathrm{d}\boldsymbol{L} \prod_{i<j}(l_i - l_j)\mathrm{e}^{-\frac{N}{4}(\frac{1}{\sigma^2}+\frac{1}{\Delta})\mathrm{Tr}\boldsymbol{L}^2} \int \mathrm{D}\boldsymbol{O}\mathrm{e}^{-\frac{N}{2\Delta}\mathrm{Tr}\boldsymbol{Y}\boldsymbol{O}\boldsymbol{L}\boldsymbol{O}^\intercal} . \tag{A.79}$$

We know from the addition of two (free) Wigner matrices with variances σ_1^2 and σ_2^2 respectively, results in another Wigner matrix of variance $\sigma_1^2 + \sigma_2^2$ and thus

$$\lim_{\eta\to 0^+} g_{\boldsymbol{Y}(t)}(x - \mathrm{i}\,\eta) = \frac{1}{2(\sigma^2+t)}\left(x - \mathrm{i}\sqrt{|x^2 - 4(\sigma^2+t)|}\right) . \tag{A.80}$$

From which we can extract the density and velocity fields:

$$v(x,t) = \frac{x}{2(\sigma^2+t)} \tag{A.81}$$

$$\rho(x,t) = \frac{\sqrt{|x^2 - 4(\sigma^2+t)|}}{2\pi(\sigma^2+t)} . \tag{A.82}$$

It is now possible to evaluate all terms in the free energy

$$\begin{aligned} F[\rho_{\boldsymbol{S}}, \rho_{\boldsymbol{Y}}; \Delta] = \frac{1}{4\Delta}\int \mathrm{d}y \rho_{\boldsymbol{Y}}(y) y^2 + \frac{1}{4}\left(\frac{1}{\sigma^2}+\frac{1}{\Delta}\right)\int \mathrm{d}s \rho_{\boldsymbol{S}}(s) s^2 \\ - \frac{1}{2}\int \mathrm{d}s\mathrm{d}s' \rho_{\boldsymbol{S}}(s)\rho_{\boldsymbol{S}}(s')\ln|s-s'| \\ - \frac{1}{2}F_2[\rho_{\boldsymbol{S}}, \rho_{\boldsymbol{Y}}; \Delta] \end{aligned} . \tag{A.83}$$

Plugging in the asymptotic expression for the HCIZ-Integral, (A.62), we obtain

$$\begin{aligned} F[\rho_{\boldsymbol{S}}, \rho_{\boldsymbol{Y}}; \Delta] = \frac{1}{4}\frac{1}{\sigma^2}\int \mathrm{d}s \rho_{\boldsymbol{S}}(s) s^2 - \frac{1}{4}\int \mathrm{d}s\mathrm{d}s' \rho_{\boldsymbol{S}}(s)\rho_{\boldsymbol{S}}(s')\ln|s-s'| \\ + \frac{1}{2}S[\rho_{\boldsymbol{S}}, \rho_{\boldsymbol{Y}}; \Delta] + \frac{1}{4}\int \mathrm{d}y\mathrm{d}y' \rho_{\boldsymbol{Y}}(y)\rho_{\boldsymbol{Y}}(y')\ln|y-y'| - \frac{1}{2}c_\Delta . \end{aligned} \tag{A.84}$$

However, our final goal is to evaluate its derivative w.r.t. Δ^{-1}, from which we obtain the MMSE by the I-MMSE relation [GSV05]. Therefore, it is not necessary to evaluate the first parts (first line in the previous equation) of the free energy, but instead only the three last parts. We start with the term independent of the spectrum of $\boldsymbol{Y}$

$$\frac{\mathrm{d}}{\mathrm{d}\Delta^{-1}} c_\Delta = -\frac{\Delta}{2} . \tag{A.85}$$

The derivative of (A.63) is also straight forward

$$\frac{\mathrm{d}}{\mathrm{d}\Delta^{-1}} S[\rho_{\boldsymbol{S}}, \rho_{\boldsymbol{Y}}; \Delta] = -\Delta^2 \int \mathrm{d}y \rho_{\boldsymbol{Y}}(y)\left(v_{\boldsymbol{Y}}(y)^2 + \frac{\pi^2}{3}\rho_{\boldsymbol{Y}}(y)^2\right) . \tag{A.86}$$

In the following we denote

$$\Sigma \coloneqq \sqrt{\sigma^2 + \Delta}$$

to lighten the notation. We plug in (A.82) to evaluate the two terms in the above integral explicitly.

- $\frac{\pi^2}{3} \int \mathrm{d}y \rho_{\boldsymbol{Y}}(y, \Delta)^3 = \frac{\pi^2}{3} \left(\frac{1}{2\pi\Sigma^2}\right)^3 \int_{-2\Sigma}^{2\Sigma} \mathrm{d}y \left(4\Sigma^2 - y^2\right)^{\frac{3}{2}} = \frac{1}{4(\sigma^2+\Delta)}$
- $\int \mathrm{d}y \rho_{\boldsymbol{Y}}(y) v_{\boldsymbol{Y}}(y)^2 = \frac{1}{2\pi\Sigma^2} \left(\frac{1}{2\Sigma}\right)^2 \int_{-2\Sigma}^{2\Sigma} \mathrm{d}y y^2 \sqrt{4\Sigma^2 - y^2} = \frac{1}{4(\sigma^2+\Delta)}$

And thus

$$\frac{\mathrm{d}}{\mathrm{d}\Delta^{-1}} S[\rho_{\boldsymbol{S}}, \rho_{\boldsymbol{Y}}; \Delta] = -\frac{1}{4} \frac{\Delta^2}{\sigma^2 + \Delta} \,. \tag{A.87}$$

It remains to evaluate the integral that results from the logarithmic repulsion between eigenvalues $\int \mathrm{d}y \mathrm{d}y' \rho_{\boldsymbol{Y}}(y) \rho_{\boldsymbol{Y}}(y') \ln|y - y'|$. This integral can be carried out by the following trick. Instead of direct integration, we use the saddle point equations for the eigenvalues of $\boldsymbol{Y}$:

$$\int \mathrm{d}y' \frac{\rho_{\boldsymbol{Y}}(y')}{y - y'} = \frac{y}{2\Sigma^2} \tag{A.88}$$

and integrate both sites w.r.t. y from 0 to y, which leads to

$$\int \mathrm{d}y' \rho(y') \left\{\ln|y - y'| - \ln|y'|\right\} = \frac{y^2}{4\Sigma^2} \,. \tag{A.89}$$

Multiplying both sides with $\rho_{\boldsymbol{Y}}(y)$ and integrating over y yields the result:

$$\int \mathrm{d}y \mathrm{d}y' \rho_{\boldsymbol{Y}}(y) \rho_{\boldsymbol{Y}}(y') \ln|y - y'| = \frac{1}{4} + \int \mathrm{d}y \rho_{\boldsymbol{Y}}(y) \ln|y| \tag{A.90}$$

$$= \frac{1}{4} + \frac{1}{2}(\ln(\sigma^2 + \Delta) - 1) \,. \tag{A.91}$$

The derivative w.r.t. Δ^{-1} can now be easily evaluated:

$$\frac{\mathrm{d}}{\mathrm{d}\Delta^{-1}} \int \mathrm{d}y \mathrm{d}y' \rho_{\boldsymbol{Y}}(y) \rho_{\boldsymbol{Y}}(y') \ln|y - y'| = -\frac{1}{2} \frac{\Delta^2}{\sigma^2 + \Delta} \,. \tag{A.92}$$

Collecting terms leads to

$$\frac{\mathrm{d}F[\rho_{\boldsymbol{S}}, \rho_{\boldsymbol{Y}}; \Delta]}{\mathrm{d}\Delta^{-1}} = -\frac{1}{4} \frac{\Delta^2}{\sigma^2 + \Delta} + \frac{\Delta}{4} = \frac{1}{4} \frac{\Delta}{1 + \Delta/\sigma^2} \tag{A.93}$$

and from the I-MMSE relation, i.e., $y\text{−MMSE}/4 = \mathrm{d}F[\rho_{\boldsymbol{S}}, \rho_{\boldsymbol{Y}}; \Delta]/\mathrm{d}\Delta^{-1}$, we obtain

$$y\text{−MMSE} = \frac{\Delta}{1 + \Delta/\sigma^2} \tag{A.94}$$

as it should be, according to (A.77).

No sane person should carry out the calculation this way, as the integral factorizes in the first place. However, it is satisfying to see that the two approaches lead to the same result. Next, we will use it to approach the more difficult problem (7.2).

A.7.2 Shooting Beasts with Cannons

If $\boldsymbol{S}$ is not a Wigner matrix, the computation is not much different. One must merely adopt the distribution on $\boldsymbol{S}$ to a Wishart distribution. There is the slight catch that the Wishart distribution is singular for $\alpha > 1$. However, this issue can be resolved as shown in [Uhl94; Sri+03]. For $\alpha \leq 1$ one has

$$P(\boldsymbol{S}) = \frac{N^{\frac{NR}{2}}}{2^{\frac{NR}{2}}\Gamma_N(\frac{R}{2})}(\det \boldsymbol{S})^{\frac{R-N-1}{2}}\mathrm{e}^{-\frac{N}{2}\mathrm{Tr}\boldsymbol{S}} \tag{A.95}$$

and for $\alpha > 1$ one must merely swap N for R and vice versa. As a consequence, the fixed point equations (A.71) are more involved. They lead to the following algebraic equation for $\mathfrak{g}_{\boldsymbol{Y}(t)}$ (that is valid for arbitrary α)

$$\mathfrak{g}^3_{\boldsymbol{Y}(t)}(z) - \Big(1+\frac{z}{t}\Big)\mathfrak{g}^2_{\boldsymbol{Y}(t)}(z) + \frac{1}{t}\Big(z + (1-\frac{1}{\alpha})\Big)\mathfrak{g}_{\boldsymbol{Y}(t)}(z) - \frac{1}{t} = 0\,. \tag{A.96}$$

Following the same steps as before, we find that

$$\frac{y\text{-MMSE}}{4} = \frac{\mathrm{d}}{\mathrm{d}\Delta^{-1}}\frac{\mathrm{d}F[\rho_{\boldsymbol{S}},\rho_{\boldsymbol{Y}};\Delta]}{\mathrm{d}\Delta^{-1}} \tag{A.97}$$

$$= \frac{\Delta}{4} - \frac{\Delta}{4}r(\Delta) + \frac{\mathrm{d}}{\mathrm{d}\Delta^{-1}}\int \mathrm{d}y\mathrm{d}y'\rho_{\boldsymbol{Y}}(y)\rho_{\boldsymbol{Y}}(y')\ln|y-y'|\,. \tag{A.98}$$

Here

$$r(\Delta) = \int \mathrm{d}y\rho_{\boldsymbol{Y}}(y)\left(v^2_{\boldsymbol{Y}}(y) + \frac{\pi^2}{3}\rho^2_{\boldsymbol{Y}}(y)\right) \tag{A.99}$$

and $\rho_{\boldsymbol{Y}}(y)$ and $v_{\boldsymbol{Y}}(y)$ follow respectively from the imaginary and real parts of the solution of (A.96) at $t = \Delta$ when approaching the real axis. This formula is valid for arbitrary, but rotationally invariant, $P(\boldsymbol{S})$.

Bibliography

[ABW10] E. Ackerman, O. Ben-Zwi, and G. Wolfovitz, "Combinatorial model and bounds for target set selection," *Theoretical Computer Science*, vol. 411, no. 44–46, pp. 4017–4022, 2010.

[Ach00] D. Achlioptas, "Setting 2 variables at a time yields a new lower bound for random 3-SAT," in *Proceedings of the thirty-second annual ACM symposium on Theory of Computing*, ACM, 2000, pp. 28–37.

[Ach01] ——, "Lower bounds for random 3-SAT via differential equations," *Theoretical Computer Science*, vol. 265, no. 1-2, pp. 159–185, 2001.

[BBMP14] J. Bun, J.-P. Bouchaud, S. N. Majumdar, and M. Potters, "Instanton approach to large n harish-chandra-itzykson-zuber integrals," *Physical review letters*, vol. 113, no. 7, p. 070 201, 2014.

[BBP17] J. Bun, J.-P. Bouchaud, and M. Potters, "Cleaning large correlation matrices: Tools from random matrix theory," *Physics Reports*, vol. 666, pp. 1–109, 2017.

[BCH+16] V. Bapst *et al.*, "The condensation phase transition in random graph coloring," *Communications in Mathematical Physics*, no. 341, pp. 543–606, 2016.

[BDM+16] J. Barbier *et al.*, "Mutual information for symmetric rank-one matrix estimation: A proof of the replica formula," in *Advances in Neural Information Processing Systems 29*, 2016, pp. 424–432.

[BDSZ16] A. Braunstein, L. Dall'Asta, G. Semerjian, and L. Zdeborová, "Network dismantling," *Proceedings of the National Academy of Sciences*, vol. 113, no. 44, pp. 12 368–12 373, 2016.

[BIPZ78] E. Brézin, C. Itzykson, G. Parisi, and J. B. Zuber, "Planar diagrams," *Comm. Math. Phys.*, vol. 59, no. 1, pp. 35–51, 1978.

[BMP+03] A. Braunstein *et al.*, "Polynomial iterative algorithms for coloring and analyzing random graphs," *Physical Review E*, vol. 68, no. 3, p. 036 702, 2003.

[BMZ05] A. Braunstein, M. Mézard, and R. Zecchina, "Survey propagation: An algorithm for satisfiability," *Random Struct. Algorithms*, vol. 27, no. 2, pp. 201–226, 2005.

[Bol01] B. Bollobás, *Random Graphs*, 2nd ed., ser. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2001.

[Bol14] E. Bolthausen, "An iterative construction of solutions of the TAP equations for the Sherrington–Kirkpatrick model," *Communications in Mathematical Physics*, vol. 325, no. 1, pp. 333–366, 2014.

[Bro41] R. L. Brooks, "On colouring the nodes of a network," in *Mathematical Proceedings of the Cambridge Philosophical Society*, Cambridge University Press, vol. 37, 1941, pp. 194–197.

[BWZ02] S. Bau, N. C. Wormald, and S. Zhou, "Decycling numbers of random regular graphs," *Random Structures and Algorithms*, vol. 21, no. 3-4, pp. 397–413, 2002.

[CFKR15] A. Coja-Oghlan, U. Feige, M. Krivelevich, and D. Reichman, "Contagious sets in expanders," in *Proceedings of the Twenty-Sixth Annual ACM-SIAM Symposium on Discrete Algorithms*, 2015, pp. 1953–1987.

[CK94] L. Cugliandolo and J. Kurchan, "On the out of equilibrium dynamics of the long-range sherrington-kirkpatrick model," *Journal of Physics A Mathematical General*, vol. 27, no. 17, pp. 5749–5772, 1994.

[CKP+17] P. Charbonneau *et al.*, "Glass and jamming transitions: From exact results to finite-dimensional descriptions," *Annual Review of Condensed Matter Physics*, vol. 8, no. 1, pp. 265–288, 2017.

[CLM17] F. Caltagirone, M. Lelarge, and L. Miolane, "Recovering asymmetric communities in the stochastic block model," *IEEE Transactions on Network Science and Engineering*, vol. PP, no. 99, pp. 1–1, 2017.

[CLR79] J. Chalupa, P. L. Leath, and G. R. Reich, "Bootstrap percolation on a bethe lattice," *Journal of Physics C: Solid State Physics*, vol. 12, no. 1, p. L31, 1979.

[CM01] S. Cocco and R. Monasson, "Statistical physics analysis of the computational complexity of solving random satisfiability problems using backtrack algorithms," vol. 22, pp. 505–531, 2001.

[Coj13] A. Coja-Oghlan, "Upper-bounding the k-colorability threshold by counting covers," *The Electronic Journal of Combinatorics*, vol. 20, no. 3, P32, 2013.

[CP16] A. Coja-Oghlan and K. Panagiotou, "The asymptotic k-sat threshold," *Advances in Mathematics*, no. 288, pp. 985–1068, 2016.

[CZ12] A. Coja-Oghlan and L. Zdeborová, "The condensation transition in random hypergraph 2-coloring," in *Proceedings of the twenty-third annual ACM-SIAM symposium on Discrete Algorithms*, SIAM, 2012, pp. 241–250.

[DAM16] Y. Deshpande, E. Abbe, and A. Montanari, "Asymptotic mutual information for the binary stochastic block model," in *IEEE International Symposium on Information Theory*, 2016, pp. 185–189.

[DM06] C. Deroulers and R. Monasson, "Criticality and universality in the unit-propagation search rule," *The European Physical Journal B-Condensed Matter and Complex Systems*, vol. 49, no. 3, pp. 339–369, 2006.

[DM14] Y. Deshpande and A. Montanari, "Information-theoretically optimal sparse PCA," in *IEEE International Symposium on Information Theory*, 2014, pp. 2197–2201.

[DR09] P. A. Dreyer and F. S. Roberts, "Irreversible k-threshold processes: Graph-theoretical threshold models of the spread of disease and of opinion," *Discrete Applied Mathematics*, vol. 157, no. 7, pp. 1615–1627, 2009.

[DS79] A. P. Dawid and A. M. Skene, "Maximum likelihood estimation of observer error-rates using the EM algorithm," *Journal of the Royal Statistical Society. Series C*, vol. 28, no. 1, pp. 20–28, 1979.

[Dys62] F. J. Dyson, "Statistical theory of the energy levels of complex systems. i," *Journal of Mathematical Physics*, vol. 3, no. 1, pp. 140–156, 1962.

[EKR15] B. Eynard, T. Kimura, and S. Ribault, "Random matrices," *ArXiv e-prints arXiv:1510.04430*, 2015.

[ER60] P. Erdös and A. Rényi, "On the evolution of random graphs," *Publ. Math. Inst. Hung. Acad. Sci*, vol. 5, pp. 17–61, 1960.

[Eyn06] B. Eynard, "Universal distribution of random matrix eigenvalues near the 'birth of a cut' transition," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2006, no. 07, P07005, 2006.

[FLM+13] U. Ferrari *et al.*, "Finite-size corrections to disordered systems on erdös-rényi random graphs," *Phys. Rev. B*, vol. 88, p. 184 201, 18 2013.

[FSZ01] J. Feinberg, R. Scalettar, and A. Zee, "'single ring theorem" and the disk-annulus phase transition," *Journal of Mathematical Physics*, vol. 42, no. 12, pp. 5718–5740, 2001.

[FW07] Y. V. Fyodorov and I. Williams, "Replica symmetry breaking condition exposed by random matrix calculation of landscape complexity," *Journal of Statistical Physics*, vol. 129, no. 5-6, pp. 1081–1116, 2007.

[Gak90] F. Gakhov, *Boundary Value Problems*, ser. Dover books on advanced mathematics. Dover, 1990.

[GDSZ17] M. Gabrié, V. Dani, G. Semerjian, and L. Zdeborová, "Phase transitions in the q -coloring of random hypergraphs," *Journal of Physics A: Mathematical and Theoretical*, vol. 50, no. 50, p. 505 002, 2017.

[Gra78] M. Granovetter, "Threshold models of collective behavior," *American journal of sociology*, vol. 83, no. 6, pp. 1420–1443, 1978.

[GS15] A. Guggiola and G. Semerjian, "Minimal contagious sets in random regular graphs," *Journal of Statistical Physics*, vol. 158, no. 2, pp. 300–358, 2015.

[GSV05] D. Guo, S. Shamai, and S. Verdú, "Mutual information and minimum mean-square error in gaussian channels," *IEEE Transactions on Information Theory*, vol. 51, no. 4, pp. 1261–1282, 2005.

[GW80] D. J. Gross and E. Witten, "Possible third-order phase transition in the large-n lattice gauge theory," *Phys. Rev. D*, vol. 21, pp. 446–453, 2 1980.

[Hat05] H. Hatami, "Random cubic graphs are not homomorphic to the cycle of size 7," *Journal of Combinatorial Theory Series B*, vol. 93, no. 2, pp. 319–325, 2005.

[HC71] J. Hammersley and P. Clifford, "Markov fields on finite graphs and lattices," Unpublished manuscript, 1971.

[HW08] C. Hoppen and N. Wormald, "Induced forests in regular graphs with large girth," *Comb. Probab. Comput.*, vol. 17, no. 3, pp. 389–410, 2008.

[Jur90] J. Jurkiewicz, "Regularization of one-matrix models," *Physics Letters B*, vol. 245, no. 2, pp. 178–184, 1990.

[KKM+16a] Y. Kabashima *et al.*, "Phase transitions and sample complexity in Bayes-optimal matrix factorization," *"IEEE" Transactions on Information Theory*, vol. 62, no. 7, pp. 4228–4265, 2016.

[KKM+16b] ——, "Phase transitions and sample complexity in bayes-optimal matrix factorization," *"IEEE Transactions on Information Theory"*, vol. 62, no. 7, pp. 4228–4265, 2016.

[KMR+07] F. Krzakala *et al.*, "Gibbs states and the set of solutions of random constraint satisfaction problems," *Proc. Natl. Acad. Sci. U.S.A*, vol. 104, p. 10 318, 2007.

[KMZ13] F. Krzakala, M. Mézard, and L. Zdeborová, "Phase diagram and approximate message passing for blind calibration and dictionary learning," in *"2013 IEEE International Symposium on Information Theory Proceedings (ISIT)""*, "IEEE", 2013, pp. 659–663.

[KNS01] A. Kostochka, J. Nešetřil, and P. Smolíková, "Colorings and homomorphisms of degenerate and bounded degree graphs," *Discrete Mathematics*, vol. 233, no. 1, pp. 257–276, 2001.

[KOS11] D. R. Karger, S. Oh, and D. Shah, "Iterative learning for reliable crowdsourcing systems," in *Advances in Neural Information Processing Systems 24*, 2011, pp. 1953–1961.

[KZ09] F. Krzakala and L. Zdeborová, "Hiding quiet solutions in random constraint satisfaction problems," *Phys. Rev. Lett.*, vol. 102, p. 238 701, 23 2009.

[LKZ15] T. Lesieur, F. Krzakala, and L. Zdeborová, "MMSE of probabilistic low-rank matrix estimation: Universality with respect to the output channel," in *53rd Annual Allerton Conference on Communication, Control, and Computing*, 2015, pp. 680–687.

[LKZ17] ——, "Constrained low-rank matrix estimation: Phase transitions, approximate message passing and applications," *Journal of Statistical Mechanics: Theory and Experiment*, no. 7, p. 073 403, 2017.

[LPI12] Q. Liu, J. Peng, and A. T. Ihler, "Variational inference for crowdsourcing," in *Advances in Neural Information Processing Systems 25*, 2012, pp. 692–700.

[Mat94] A. Matytsin, "On the large-n limit of the itzykson-zuber integral," *Nuclear Physics B*, vol. 411, no. 2, pp. 805–820, 1994.

[Mio17] L. Miolane, "Fundamental limits of low-rank matrix estimation: the non-symmetric case," *ArXiv preprint arXiv:1702.00473*, 2017.

[MM09] M. Mézard and A. Montanari, *Information, Physics, and Computation*, ser. Oxford Graduate Texts. OUP Oxford, 2009.

[Mon98] R. Monasson, "Optimization problems and replica symmetry breaking in finite connectivity spin glasses," *Journal of Physics A: Mathematical and General*, vol. 31, no. 2, p. 513, 1998.

[MP01] M. Mézard and G. Parisi, "The bethe lattice spin glass revisited," *Eur. Phys. J. B*, vol. 20, p. 217, 2001.

[MP03] ——, "The cavity method at zero temperature," *J. Stat. Phys.*, vol. 111, pp. 1–34, 2003.

[MPR05] M. Mézard, M. Palassini, and O. Rivoire, "Landscape of solutions in constraint satisfaction problems," *Physical review letters*, vol. 95, no. 20, p. 200 202, 2005.

[MPV87] M. Mezard, G. Parisi, and M. Virasoro, *Spin Glass Theory and Beyond*, ser. Lecture Notes in Physics. World Scientific, 1987.

[MPWZ02] R. Mulet, A. Pagnani, M. Weigt, and R. Zecchina, "Coloring random graphs," *Phys. Rev. Lett.*, vol. 89, p. 268 701, 2002.

[MPZ02] M. Mézard, G. Parisi, and R. Zecchina, "Analytic and algorithmic solution of random satisfiability problems," *Science*, vol. 297, pp. 812–815, 2002.

[MR08] N. Muskhelishvili and J. Radok, *Singular Integral Equations: Boundary Problems of Function Theory and Their Application to Mathematical Physics*, ser. Dover Books on Physics Series. Dover Publications, 2008.

[MR95] M. Molloy and B. Reed, "A critical point for random graphs with a given degree sequence," *Random Structures & Algorithms*, vol. 6, pp. 161–180, 1995.

[MT13] R. Matsushita and T. Tanaka, "Low-rank matrix reconstruction and clustering via approximate message passing," in *Advances in Neural Information Processing Systems 26*, 2013, pp. 917–925.

[Mui09] R. J. Muirhead, *Aspects of multivariate statistical theory*. John Wiley & Sons, 2009, vol. 197.

[Neš13] J. Nešetřil, "A combinatorial classic—sparse graphs with high chromatic number," in *Erdős Centennial*, Springer, 2013, pp. 383–407.

[Nis80] H. Nishimori, "Exact results and critical properties of the ising model with competing interactions," *Journal of Physics C: Solid State Physics*, vol. 13, no. 21, p. 4071, 1980.

[OOSY16a] J. Ok, S. Oh, J. Shin, and Y. Yi, "Optimal inference in crowdsourced classification via belief propagation," *ArXiv preprint arXiv:1602.03619*, 2016.

[OOSY16b] ——, "Optimality of belief propagation for crowdsourced classification," in *International Conference on Machine Learning*, 2016, pp. 535–544.

[Pfi14] H. D. Pfister, *The analysis of graph peeling processes*, course notes available online: `http://pfister.ee.duke.edu/courses/ece590_gmi/peeling.pdf`, 2014.

[PSW96a] B. Pittel, J. Spencer, and N. Wormald, "Sudden emergence of a giant k-core in a random graph," *Journal of Combinatorial Theory-Series B*, vol. 67, no. 1, 1996.

[PSW96b] B. Pittel, J. Spencer, and N. Wormald, "Sudden emergence of a giantk-core in a random graph," *Journal of Combinatorial Theory, Series B*, vol. 67, no. 1, pp. 111–151, 1996.

[Ree17] G. Reeves, "Additivity of information in multilayer networks via additive gaussian noise transforms," in *2017 55th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, 2017, pp. 1064–1070.

[Rei12] D. Reichman, "New bounds for contagious sets," *Discrete Mathematics*, vol. 312, no. 10, pp. 1812–1814, 2012.

[RF12] S. Rangan and A. K. Fletcher, "Iterative estimation of constrained rank-one matrices in noise," in *IEEE International Symposium on Information Theory Proceedings*, 2012, pp. 1246–1250.

[RPD18] G. Reeves, H. D. Pfister, and A. Dytso, "Mutual information as a function of matrix snr for linear gaussian channels," *IEEE International Symposium on Information Theory*, 2018.

[Sch18] C. Schmidt, `https://github.com/hcmidt/corehd`, 2018.

[SGZ16] C. Schmidt, N.-E. Guenther, and L. Zdeborová, "Circular coloring of random graphs: Statistical physics investigation," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2016, no. 8, p. 083 303, 2016.

[SPZ18] C. Schmidt, H. D. Pfister, and L. Zdeborová, "On minimal sets to destroy the k-core in random networks," *ArXiv e-prints arXiv:1806.03134*, 2018.

[Sri+03] M. S. Srivastava *et al.*, "Singular wishart and multivariate beta distributions," *The Annals of Statistics*, vol. 31, no. 5, pp. 1537–1560, 2003.

[SW99] A. Steger and N. Wormald, "Generating random regular graphs quickly," *Combinatorics, Probability and Computing*, vol. 8, pp. 377–396, 1999.

[SXS+17] P. Sen *et al.*, "Efficient collective influence maximization in cascading processes with first-order transitions," *Scientific Reports*, vol. 7, 2017.

[SZ18] C. Schmidt and L. Zdeborová, "Dense limit of the Dawid-Skene model for crowdsourcing and regions of sub-optimality of message passing algorithms," *ArXiv e-prints arXiv:1803.04924*, 2018.

[TAP77] D. J. Thouless, P. W. Anderson, and R. G. Palmer, "Solution of 'solvable model of a spin glass'," *The Philosophical Magazine: A Journal of Theoretical Experimental and Applied Physics*, vol. 35, no. 3, pp. 593–601, 1977.

[Tho86] D. J. Thouless, "Spin-glass on a Bethe lattice," *Phys. Rev. Lett.*, vol. 56, pp. 1082–1085, 1986.

[Tri85] F. G. Tricomi, *Integral equations*. Courier Corporation, 1985, vol. 5.

[TV12] T. Tao and V. Vu, "Random covariance matrices: Universality of local statistics of eigenvalues," *The Annals of Probability*, vol. 40, no. 3, pp. 1285–1315, 2012.

[Uhl94] H. Uhlig, "On singular wishart and singular multivariate beta distributions," *The Annals of Statistics*, pp. 395–405, 1994.

[VB85] L. Viana and A. J. Bray, "Phase diagrams for dilute spin-glasses," *J. Phys. C*, vol. 18, pp. 3037–3051, 1985.

[Was13] L. Wasserman, *All of Statistics: A Concise Course in Statistical Inference*, ser. Springer Texts in Statistics. Springer New York, 2013.

[WBPB10] P. Welinder, S. Branson, P. Perona, and S. J. Belongie, "The multidimensional wisdom of crowds," in *Advances in Neural Information Processing Systems 23*, 2010, pp. 2424–2432.

[Wig51] E. P. Wigner, "On the statistical distribution of the widths and spacings of nuclear resonance levels," in *Mathematical Proceedings of the Cambridge Philosophical Society*, Cambridge University Press, vol. 47, 1951, pp. 790–798.

[Wis28] J. Wishart, "The generalized product moment distribution in samples from a normal multivariate population," *Biometrika*, vol. A, no. 20, pp. 32–52, 1928.

[Wor95] N. C. Wormald, "Differential equations for random processes and random graphs," *Ann. Appl. Probab.*, vol. 5, no. 4, pp. 1217–1235, 1995.

[Wor99a] N. C. Wormald, "The differential equation method for random graph processes and greedy algorithms," *Lectures on approximation and randomized algorithms*, pp. 73–155, 1999.

[Wor99b] N. C. Wormald, "The differential equation method for random graph processes and greedy algorithms," *Lectures on approximation and randomized algorithms*, vol. 73, p. 155, 1999.

[WW01] I. M. Wanless and N. C. Wormald, "Regular graphs with no homomorphisms onto cycles," *Journal of Combinatorial Theory, Series B*, vol. 82, no. 1, pp. 155–160, 2001.

[Zam10] F. Zamponi, "Mean field theory of spin glasses," *ArXiv preprint arXiv:1008.4844*, 2010.

[ZB10] L. Zdeborová and S. Boettcher, "A conjecture on the maximum cut and bisection width in random regular graphs," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2010, no. 02, P02020, 2010.

[Zde09] L. Zdeborová, "Statistical physics of hard optimization problems," *Acta physica slovaca*, vol. 59, no. 3, pp. 169–303, 2009.

[Zho13] H.-J. Zhou, "Spin glass approach to the feedback vertex set problem," *The European Physical Journal B*, vol. 86, no. 11, p. 455, 2013.

[Zhu06] X. Zhu, "Recent developments in circular colouring of graphs," in *Topics in discrete mathematics*, Springer, 2006, pp. 497–550.

[ZK07] L. Zdeborová and F. Krzakala, "Phase transitions in the coloring of random graphs," *Phys. Rev. E*, vol. 76, p. 031 131, 2007.

[ZK11] L. Zdeborová and F. Krzakala, "Quiet planting in the locked constraint satisfaction problems," *SIAM Journal on Discrete Mathematics*, vol. 25, no. 2, pp. 750–770, 2011.

[ZK16] ——, "Statistical physics of inference: Thresholds and algorithms," *Advances in Physics*, vol. 65, pp. 453–552, 2016.

[Zub08] J.-B. Zuber, "The large-n limit of matrix integrals over the orthogonal group," *Journal of Physics A: Mathematical and Theoretical*, vol. 41, no. 38, p. 382 001, 2008.

[ZZZ16] L. Zdeborová, P. Zhang, and H.-J. Zhou, "Fast and simple decycling and dismantling of networks," *Scientific Reports*, vol. 6, no. 37954, 2016.

Synthèse de la thèse

À première vue, les sujets traités dans cette thèse ne semblent pas des problèmes de physique. Parmi eux se trouvent un problème de satisfaction de contrainte, un processus dynamique sur un graphe et deux problèmes d'inférence. Cependant, un examen plus attentif révèle que tous ces problèmes sont liés par le fait même qu'ils contiennent du *désordre* d'une manière ou d'une autre et qu'ils peuvent être formulés sous la forme d'un *modèle graphique probabiliste*. Cette connexion révèle le lien étroit qui existe entre la physique statistique et les systèmes désordonnés, qui peut être exploité pour examiner les problèmes par une approche de la physique statistique. Une telle approche est la manière appropriée d'étudier des systèmes contenant trop de constituants, d'interactions complexes ou, pour une autre raison, sont trop compliqués pour être décrits de manière microscopique.

Ce lien est plus facile à décrire pour les problèmes d'inférence, dans lesquels on fournit des données et tente de déduire quelque chose sur le processus de génération de données sous-jacent. Ces ensembles de données sont généralement bruyants et / ou volumineux; les deux sont de arguments solides en faveur d'une approche de la physique statistique. Les données jouent le rôle du désordre et le théorème de Bayes nous fournit le lien formel avec la physique statistique. Dénotez les données $\boldsymbol{D}$ et laissez le processus génératif être décrit par les variables $\boldsymbol{G}$. Puis le théorème de Bayes fournit le lien entre la *probabilité postérieure* pour $\boldsymbol{G}$, à partir de $\boldsymbol{D}$, écrit $P(\boldsymbol{G} \mid \boldsymbol{D})$, et la *fonction de partition*, $Z(\boldsymbol{Y})$:

$$P(\boldsymbol{G} \mid \boldsymbol{D}) = \frac{P(\boldsymbol{D} \mid \boldsymbol{G})P(\boldsymbol{G})}{Z(\boldsymbol{D})} \,.$$

Les problèmes d'inférence considérés seront définis dans le cadre *Bayes optimal*, dans lequel on connaît la forme fonctionnelle de $P(\boldsymbol{D} \mid \boldsymbol{G})$ et $P(\boldsymbol{G})$.

Tout au long de cette thèse, il y a deux questions récurrentes: (I) quelle est la performance optimale pouvant être atteinte? Et (II) quelles sont les stratégies algorithmiques qui conduisent à une performance optimale. Le logarithme moyenné de la fonction de partition, $\mathbb{E} \ln Z(\boldsymbol{D})$, l'énergie libre, est la fonction génératrice qui permet de déduire les propriétés physiques du système. Il est lié aux deux questions ci-dessus. La méthode de *réplique* facilite le calcul de l'énergie libre en fonction de certains paramètres d'ordre macroscopiques liés à la meilleure erreur possible dans l'inférence de $\boldsymbol{G}$ et répond donc à la première question. La méthode *cavité* permet de formuler une énergie libre microscopique dont la minimisation (par rapport aux variables microscopiques) conduit à un algorithme de transmission de messages exploitable pour aborder la deuxième question. Les problèmes d'inférence considérés sont tous deux liés à *factorisation matricielle (ou estimation)*.

Tout d'abord, le modèle de Dawid-Skene est étudiée dans la limite où le modèle graphique bipartite sous-jacent est densément connecté. Ce qui est montré de réduire à un problème de *factorisation matricielle bas rang*. Et deuxièmement, nous étudions un problème particulier pour *estimation matricielle symétrique à rangs élevé* pour lequel il est nécessaire de combiner la méthode de réplique avec des outils empruntés de la *théorie des matrices aléatoires*.

Les *problèmes de satisfaction de contraintes* entrent dans un cadre mathématique très similaire. Cependant, le problème considéré n'est pas défini dans le cadre optimal de Bayes. Les implications physiques sont que des états vitreux peuvent être présents. L'espace des solutions peut présenter une structure difficile qui empêche la solution symétrique de réplique simple de l'énergie libre de réplique d'être correcte. Au lieu de cela, le schéma de brisure de symétrie de réplique de Parisi doit être appliqué pour un traitement complet de ces problèmes. Cela a également des conséquences immédiates pour l'approche algorithmique qui doivent également être étendue.

Le lien profond avec la physique statistique depasse l'analogie aux problèmes de satisfaction d'inférence et de contrainte. Les modèles de seuil sont une approche courante pour modéliser des processus dynamiques collectifs sur des réseaux et constituent un autre exemple étroitement lié au processus de percolation par bootstrap en physique statistique. Une question d'importance majeure concerne les ensembles minimaux contagieux dans de tels processus dynamiques. Ce problème peut être considéré comme un problème d'optimisation par rapport aux conditions initiales du processus dynamique. Ou, également, un problème de verre de spin avec des interactions particulières. Bien que très différent des deux modèles précédents, il peut être démontré qu'il est accessible par les mêmes méthodes [GS15]. Cependant, les aspects algorithmiques de ce problème ne sont pas encore bien élaborés et laissent des questions ouvertes (a) après de bons algorithmes locaux et (b) leurs performances.

Contributions et organisation du manuscrit

Dans la première partie, le contexte nécessaire est présenté et les méthodes susmentionnées sont décrites. Au chapitre 1, une introduction aux modèles graphiques probabilistes est donnée. Ensuite, au chapitre 2, les concepts de systèmes désordonnés sont introduits, nécessaires pour développer le lien entre les problèmes d'inférence et la physique statistique. Enfin, au chapitre 3, certains éléments de la théorie des matrices aléatoires sont collectés qui sont nécessaires pour résoudre le problème de l'estimation matricielle par rangs élevé. La deuxième partie de cette thèse contient mes contributions qui sont décrites par ordre d'apparition dans ce qui suit.

Coloration Circulaire

Le problème de coloration circulaire est un problème de satisfaction de contraintes qui appartient à une ensemble des problèmes plus large qui généralise le problème de coloration de graphes canoniques. Alors que dans la coloration canonique, deux nœuds d'un graphe reliés par une arête doivent avoir des couleurs différentes, dans une coloration circulaire, les couleurs sont ordonnées dans un cercle et deux nœuds adjacents doivent avoir deux couleurs adjacentes.

Ma contribution principale consiste à analyser le problème dans le cadre de briser la symétrie des réplique en une étape. Cela a révélé plusieurs caractéristiques intéressantes, qui ne sont pas communes à d'autres problèmes de satisfaction de contraintes

discrètes. De plus, l'investigation établit du point de vue de la physique statistique une conjecture de Nešetřil sur la colorabilité circulaire des graphes sous-cubiques, pour les graphes aléatoires.

Ce travail a été publié dans le *Journal de mécanique statistique: théorie et expériences, volume 2016, août 2016* [SGZ16]:

- Christian Schmidt, Nils-Eric Guenther and Lenka Zdeborová, "Circular coloring of random graphs: statistical physics investigation"

Modèle Dawid-Skene dans la limite dense (Crowdsourcing)

Considérons la catégorisation des M points de données par la contribution de N individus et supposons que chaque individu puisse être caractérisé par un seul paramètre de fiabilité (la probabilité qu'un réponse correct est donner). C'est ce qu'on appelle le modèle de Dawid et Skene (DS). Dans le modèle DS, deux régimes peuvent être distingués. Le régime épars, dans lequel chaque individu n'est assigné qu'un nombre infime, $\Theta(1)$, de questions. Et le régime dense, dans lequel chaque individu se voit attribuer un grand nombre, $\Theta(N)$ de questions. La plupart des études théoriques précédentes sur le crowdsourcing se sont concentrées sur la limite éparse du modèle DS.

Au contraire, ma contribution est l'analyse de forme fermée du modèle DS dans la limite dense qui révèle le diagramme de phase du problème. Ceci est fait en montrant que le modèle DS dense appartient à une classe plus large de problèmes d'estimation de matrice de bas rang pour lesquels il est possible d'exprimer la performance asymptotique, optimale de Bayes, dans une forme fermée simple. De plus, des résultats numériques pour l'extrapolation dans le régime clairsemé sont obtenus et sur des données du monde réel.

Le travail a été soumis aux *IEEE Transactions on Information Theory*

- Christian Schmidt, Lenka Zdeborová, "Dense limit of the Dawid-Skene model for crowdsourcing and regions of sub-optimality of message passing algorithms"

et peut également être trouvé sur le arxiv [SZ18].

Ensembles contagieux dans des graphes aléatoires

Considérons un modèle dans lequel les nœuds du graphe peuvent être dans un état actif ou inactif et dans lequel un nœud v passe de l'état inactif à l'état actif si plus de $t_v = d_v - k$ de ses voisins sont actifs. Le nombre t_v est appelé le seuil pour le nœud v. Détruire le k-noyau équivaut à l'activation de l'intégralité du graphique sous la dynamique de seuil ci-dessus. Le plus petit de ces ensembles est l'ensemble minimal contagieux.

Ma contribution principale est de montrer que l'algorithme COREHD (une stratégie gloutonne locale) se traduit par un processus aléatoire sur la distribution des degrés du graphe qui peut être suivi exactement par la dérivation de la limite continue du processus aléatoire. Cela conduit à de nouvelles limites supérieures rigoureuses sur la taille des ensembles contagieux minimaux pour les graphes aléatoires à degré lié, tirés

du modèle de configuration. J'ai également proposé de nouvelles méthodes heuristiques qui attaquent le problème de manière algorithmique et surperforment les meilleures algorithmes locaux connus.

Les contributions ont été soumises

- Christian Schmidt, Henry D. Pfister and Lenka Zdeborová, "On Minimal Sets to Destroy the k-Core in Random Networks"

en *Physical Review E* et un eprint peut être trouvé sous [SPZ18].

L'Estimation Matricielle par Rangs Élevé

Considérons le problème le plus simple possible qui entre dans la catégorie des problèmes de factorisation matricielle à rangs élevé. La version symétrique $\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{X}^\intercal + \boldsymbol{W}$ avec un probabilité a priori Gaußian sur $\boldsymbol{X}$ sous un bruit symétrique Gaußian, $\boldsymbol{W}$. Si $\boldsymbol{X} \in \mathbb{R}^{N\times R}$ et $N/R = O(1)$, dans la limite où $N \to \infty$, on dit que la matrice a rang élevé.

Je calcule l'erreur optimale de Bayes sur $\boldsymbol{X}$. Au cours de ce calcul, je montre qu'il n'est pas possible de réduire l'énergie libre de la réplique à un seul paramètre d'ordre scalaire sous la forme précédemment assumée dans d'autres travaux. Au lieu de cela, il est nécessaire de conserver la densité spectrale de toute une matrice de chevauchement et l'énergie libre devient fonctionnelle de cette densité. Pour résoudre l'énergie libre, il est nécessaire de résoudre une équation de point-selle de matrice aléatoire qui mène à une équation d'intégrale singulière dans la limite $N \to \infty$. Je résous les équations et exprime l'erreur asymptotique sur $\boldsymbol{X}$ en termes de cette solution.

Cette partie de mon travail est actuellement en préparation pour être soumise.

université PARIS-SACLAY
ÉCOLE DOCTORALE
Physique en Île-de-France (EDPIF)

Titre: Physique statistique des modèles épars et denses en optimisation et inférence

Mots clés: Systèmes désordonnés, verres de spin, inférence bayésienne, graphe aléatoire, problème de satisfaction de contraintes

Résumé: Une donnée peut avoir diverses formes et peut provenir d'un large panel d'applications. Habituellement, une donnée possède beaucoup de bruit et peut être soumise aux effets du hasard. Les recents progrès en apprentissage automatique ont relancé les recherches théoriques sur les limites des différentes méthodes probabilistes de traitement du signal. Dans cette thèse, nous nous intéressons aux questions suivantes : quelle est la meilleure performance possible atteignable ? Et comment peut-elle être atteinte, i.e., quelle est la stratégie algorithmique optimale ? La réponse dépend de la forme des données. Les sujets traités dans cette thèse peuvent tous être représentés par des modèles graphiques. Les propriétés des données determinent la structure intrinsèque du modéle graphique correspondant. Les structures considérées ici sont soit éparses, soit denses.
Les questions précédentes peuvent être étudiées dans un cadre probabiliste, qui permet d'apporter des réponses typiques. Un tel cadre est naturel en physique statistique et crée une analogie formelle avec la physique des systèmes désordonnés. En retour, cela permet l'utilisation d'outils spécifiques à ce domaine et de résoudre des problèmes de satisfaction de contraintes et d'inférence statistique. La problématique de performance optimale est directement reliée à la structure des extrema de la fonction d'énergie libre macroscopique, tandis que les aspects algorithmiques proviennent eux de la minimisation de la fonction d'énergie libre microscopique (c'est-à-dire, dans la forme de Bethe).
Cette thèse est divisée en quatre parties. Premièrement, nous aborderons par une approche de physique statistique le problème de la coloration de graphes aléatoires et mettrons en évidence un certain nombre de caractéristiques. Dans un second temps, nous calculerons une nouvelle limite supérieure de la taille de l'ensemble contagieux. Troisièmement, nous calculerons le diagramme de phase du modèle de Dawid et Skene dans la région dense en modélisant le problème par une factorisation matricielle de petit rang. Enfin, nous calculerons l'erreur optimale de Bayes pour une classe restreinte de l'estimation matricielle de rang élevé.

Title: Statistical physics of sparse and dense models in optimization and inference

Keywords: Disordered systems, spin glasses, Bayesian inference, random graphs, constraint satisfaction problems

Abstract: Datasets come in a variety of forms and from a broad range of different applications. Typically, the observed data is noisy or in some other way subject to randomness. The recent developments in machine learning have revived the need for exact theoretical limits of probabilistic methods that recover information from noisy data. In this thesis we are concerned with the following two questions: what is the asymptotically best achievable performance? And how can this performance be achieved, i.e., what is the optimal algorithmic strategy? The answer depends on the properties of the data. The problems in this thesis can all be represented as probabilistic graphical models. The generative process of the data determines the structure of the underlying graphical model. The structures considered here are either sparse random graphs or dense (fully connected) models.
The above questions can be studied in a probabilistic framework, which leads to an average (or typical) case answer. Such a probabilistic formulation is natural to statistical physics and leads to a formal analogy with problems in disordered systems. In turn, this permits to harvest the methods developed in the study of disordered systems, to attack constraint satisfaction and statistical inference problems. The formal analogy can be exploited as follows. The optimal performance analysis is directly related to the structure of the extrema of the macroscopic free energy. The algorithmic aspects follow from the minimization of the microscopic free energy (that is, the Bethe free energy in this work) which is closely related to message passing algorithms.
This thesis is divided into four contributions. First, a statistical physics investigation of the circular coloring problem is carried out that reveals several distinct features. Second, new rigorous upper bounds on the size of minimal contagious sets in random graphs, with bounded maximum degree, are obtained. Third, the phase diagram of the dense Dawid-Skene model is derived by mapping the problem onto low-rank matrix factorization. The associated approximate message passing algorithm is evaluated on real-world data. Finally, the Bayes optimal denoising mean square error is derived for a restricted class of extensive rank matrix estimation problems.

Université Paris-Saclay
Espace Technologique / Immeuble Discovery
Route de l'Orme aux Merisiers RD 128 / 91190 Saint-Aubin, France