

POLITECNICO DI TORINO



MASTER THESIS

Disorder Chaos in Diluted Spin Glass models and Constraint Satisfaction Problems

Author:
Alessandro ZORZAN

Supervisors:
Prof. Noela MÜLLER
Prof. Andrea PAGNANI

*A thesis submitted in fulfillment of the requirements
for the Master degree in*

Physics of Complex Systems

October 20, 2025

“The mathematician’s patterns, like the painter’s or the poet’s must be beautiful; the ideas like the colours or the words, must fit together in a harmonious way. Beauty is the first test: there is no permanent place in the world for ugly mathematics”

G.H. Hardy, A Mathematician’s Apology

Acknowledgements

In physics, we often try to describe systems made up of millions of participants: particles, spins, molecules, quarks.

At first, to make things simpler, we assume they do not influence one another, that none of them cares about the others.

Not the slightest exchange of information, not the tiniest or most minimal interaction.

Thus, many of these small, singular, complex contributions are lost every day, and yet in everyday stories it is precisely these interactions that matter most.

The interactions I have had over these years are decidedly non-negligible, and it is to them that the acknowledgements of this thesis are addressed.

First of all, I thank Professor Noela Müller, without whom this thesis would not exist. Thank you for your superhuman patience and your way of encouraging research independence. Thank you for the commitment in supervising me and for the incredibly humane way of taking a student under your wing. Thank you for making my brief but intense stay in Eindhoven possible and for encouraging me to attend the YEP workshop. Thank you for giving me the opportunity to experience first-hand what it means to do research and, thank you, for the kindness with which you urged me to “be more aggressive.”

Thanks to my Mom Paola and my Dad Diego, for always making me feel the unconditional love that parents give, even in moments of discouragement when I didn’t know what was hardest to do. Thank you for (almost) never holding against me my being phlegmatic and limping along. For supporting and contributing by every possible means to my education and academic path.

Thanks to Fabi, who with every passing day makes me the proudest brother in the world, the older brother who will always have something to learn from his little sister.

Thanks to my Uncle Aldo and my Aunt Tizi for having been a loving extension of this family for as long as I can remember.

And thanks to Ari, for being more an older sister than a cousin, for all you’ve passed on to me, directly and indirectly—thank you for the company at the little concerts.

Thanks to Grandma Elsa and Grandpa Bruno, for being the best place in which to grow up.

Thanks to Fede, for seeing who-knows-what in me, for putting up with me while I prepared my last exams, and for welcoming all the unnecessary and unsolicited info dumps about this thesis. Thank you for choosing, and for choosing again, to stand by me.

Thanks to Giovanni, Matteo, and Salvatore for the continual exchange and comfort over the years of this Master’s. I could not have asked for better people by my side to face this path, certainly not free of obstacles, seemingly insurmountable.

Thanks to Lorenzo, Kai-Chun, and Wei-Ting for being patient enough to listen to me ramble about the topic of this thesis, and just as curious in asking me questions about it. Thanks for being the toughest examiners I could have faced on that June afternoon, in the Metaforum office.

And thanks to Marco for the warm welcome in Eindhoven, the productive climbing sessions, the disastrous chess matches with a considerable alcohol content, and the snacks offered at the Spar under the mathematics department.

Thanks to Stefano (Dr.) Piva, for the trust you placed in me, which then became a great friendship.

Thanks to Mr. Michele and Mr. Luigi, for their impassioned enthusiasm for Science, for their faith in outreach and sharing, for everyone, with no barriers of any kind. For being the people who best look “from the shoulders of giants,” and perhaps a little further.

Finally, thanks to lifelong friends.

Thanks to Podo, for his extraordinary kindness. For his disarming reliability and readiness, without compromise.

Thanks to Luca, for his brotherly presence and for friendship in its most sincere sense.

Thanks to Bobo, for his cheerfulness, contagious laughter, and the “business-man” moments on the subway.

Thanks to Ste, for being there “since day zero,” for making the early years of university light, for being a sure and steady presence between one lecture and the next.

Thanks to Bera, for his quiet presence in rough moments, for his selfless gestures of care, for his patience during the multi-pitch climb.

Thanks to Fedo, for being the quiet friend, yet with an ear always pricked and ready to listen.

Thanks to Simo, for the therapeutic moments shared on rock or at the climbing gym.

Thanks to Gaia, for being Gaia. For her contagious happiness, for the hours spent in the library between exams and thesis, for the thousand words of comfort and the hugs when things weren’t going well.

Thanks to Elisa for helping me understand that “questo eterno mio incespicare” is not necessarily always a bad thing. For putting up with my ramblings of every kind on those walks where we gave I-don’t-know-how-many steps to the road and hours to the night.

Thanks to Euge for her explosive enthusiasm and great drive, for being the one true testimonial of the “cioccolatlattecioccolato” moment.

Thanks to Nico, for the sincerity of the “how are you?” every time we see each other.

To this sea of affection, gratitude as my poor, only response.

In fisica, cerchiamo spesso di descrivere sistemi composti da milioni di partecipanti: particelle, spin, molecole, quark.

All'inizio, per rendere le cose più semplici, supponiamo che non si influenzino a vicenda, che a nessuno importi nulla dell'altro.

Né il più piccolo scambio di informazione, né la più piccola o minima interazione.

Così, tanti di questi piccoli e singolari, complessi contributi si perdono ogni giorno, eppure nel racconto di tutti i giorni, sono queste interazioni a contare di più.

Le interazioni che ho avuto in questi anni sono decisamente non trascurabili ed è a queste che sono rivolte i ringraziamenti di questa tesi.

Prima di tutto, ringrazio la Professoressa Noela Müller, senza la quale questa tesi non esisterebbe. Grazie per la sua pazienza sovrumana e il suo modo di incoraggiare l'indipendenza della ricerca. Grazie per l'impegno nel seguirmi e nel modo incredibilmente umano di prendere uno studente sotto la propria ala. Grazie per aver reso possibile la mia breve, ma intensa, permanenza a Eindhoven e avermi incoraggiato a partecipare al workshop YEP. Grazie per avermi dato l'opportunità di vivere in prima persona cosa significa fare ricerca e grazie, per la gentilezza con cui mi ha spronato a "essere più aggressivo".

Grazie a mamma Paola e papà Diego, per avermi fatto sempre sentire incondizionatamente l'amore che i genitori danno, anche nei momenti di sconforto in cui non conoscevo cosa più difficile da fare. Grazie per non avermi fatto (quasi) mai pesare questo mio essere flemmatico e claudicante. Per aver supportato e contribuito con ogni mezzo possibile alla mia istruzione e al mio percorso accademico.

Grazie a Fabi, che ogni giorno che passa mi rende il fratello più orgoglioso del mondo, il fratello maggiore che avrà sempre da imparare dalla sorellina.

Grazie a zio Aldo e zia Tizi per essere stati un'estensione amorevole di questa famiglia da che ne ho memoria.

E grazie ad Ari, per essere stata una sorella maggiore più che una cugina, per avermi trasmesso molto direttamente e indirettamente, grazie per la compagnia ai concerti.

Grazie a nonna Elsa e nonno Bruno, per essere stati il miglior posto in cui crescere.

Grazie a Fede, per aver visto chissà cosa in me, per avermi supportato durante la preparazione degli ultimi esami e per aver accolto tutti gli info dump non necessari e non richiesti su questa tesi. Grazie per aver scelto e scegliere di starmi accanto.

Grazie a Giovanni, Matteo e Salvatore per il continuo confronto e conforto, scambiato negli anni di questa magistrale. Non avrei potuto chiedere al mio fianco persone migliori per affrontare questo percorso non certo privo di ostacoli, apparentemente insormontabili.

Grazie a Lorenzo, Kai-Chun e Wei-Ting per essere stati così pazienti da ascoltarmi farneticare sull'argomento di questa tesi e altrettanto curiosi da avermi fatto domande a riguardo. Grazie per essere stati gli esaminatori più severi che potessi affrontare in quel pomeriggio di giugno, nell'ufficio del Metaforum.

E grazie a Marco, per la calorosa accoglienza ad Eindhoven, le proficue sessioni di arrampicata, le disastrose partite a scacchi dal considerevole tasso alcolemico e le merende offerte allo Spar sotto al dipartimento di matematica.

Grazie a Stefano (Dottor) Piva, per la fiducia che mi ha concesso, trasformatasi poi

in grande amicizia.

Grazie al signor Michele e al signor Luigi, per il loro accalorato entusiasmo nella Scienza, per la fiducia nella divulgazione e nella condivisione, per tutte e tutti, senza alcun tipo di barriera. Per essere le persone che meglio osservano "sulle spalle dei giganti" e forse un po' più in là.

Infine, grazie agli amici di una vita.

Grazie a Pode, per la sua gentilezza fuori dal normale. Per la sua affidabilità e disponibilità che disarmano, senza compromessi.

Grazie a Luca, per la presenza fraterna e per l'amicizia nel suo significato più sincero.

Grazie a Bobo, per la sua allegria, le risate contagiose, per i momenti "business-man" in metropolitana.

Grazie a Ste, per esserci stato "dal giorno zero", per aver reso leggeri i primi anni di università, per essere stato una presenza certa e sicura tra una lezione e l'altra.

Grazie a Bera, per la sua presenza silenziosa nei momenti no, per i suoi gesti di cura disinteressati, per la pazienza durante la scalata in multi-pitch.

Grazie a Fedo, per essere l'amico silenzioso, ma con l'orecchio sempre teso e pronto ad ascoltare.

Grazie a Simo, per i momenti terapeutici condivisi su roccia o in palestra d'arrampicata.

Grazie a Gaia, per essere Gaia. Per la sua felicità contagiosa, per le ore passate in biblioteca tra esami e tesi, per le mille parole di conforto spese e gli abbracci quando le cose non andavano.

Grazie a Elisa per avermi fatto capire che "questo eterno mio icespicare" non è sempre necessariamente un male. Per aver sopportato i miei sproloqui di ogni genere nelle camminate in cui abbiamo regalato non so quanti passi alla strada e ore alla notte.

Grazie a Euge per il suo entusiasmo esplosivo e la sua grande intraprendenza, per essere l'unica vera testimonial del momento "cioccolatlattateecioccolato".

Grazie a Nico, per la sincerità dei "come stai?" ogni volta che ci vediamo.

A questo mare di affetto, la gratitudine come povera, unica risposta.

Abstract

Disorder chaos refers to the extreme sensitivity of the equilibrium states of a system to small perturbations of the underlying disorder. While this effect has been extensively studied and rigorously characterized in fully connected models, its understanding in diluted systems remains incomplete. We begin by introducing the framework of spin glass models, reviewing the Sherrington–Kirkpatrick (SK) model and its generalization to mixed p -spin models. These fully connected models serve as paradigmatic setting where the mathematical theory is well developed: the presence of disorder chaos have been rigorously established. We then move to diluted models, in which each variable interacts with only finitely many others. In this regime, we focus on several Random Constraint Satisfaction Problems: Random k -SAT, Random k -NAESAT, Random Hypergraph 2-coloring and Random k -XORSAT, and interpret them with the spin glass formalism by defining suitable Hamiltonians with random interactions. This analogy highlights the deep structural connection between Spin Glasses and Random Constraint Satisfaction Problems, allowing us to apply techniques originally developed for fully connected systems. Within this framework, we analyze different perturbation schemes: such as resampling a fraction of clauses or flipping random signs, and investigate their effect on the structure of the solution space. Our goal is to quantify how correlation between two instances depend on the degree of the disorder perturbation and to identify the onset of chaos in these diluted models.

Finally, this thesis aims to establish a rigorous foundation of disorder chaos in diluted Random Constraint Satisfaction Problems, extending the mathematical theory of Spin Glass chaos beyond the fully connected setting. By bridging methods from Probability Theory and Statistical Physics we show that the same principles governing disorder chaos in the Sherrington-Kirkpatrick model underline the geometry and stability of solutions in Random Constraint Satisfaction Problems.

Contents

Acknowledgements	iii
Abstract	vii
1 The Sherrington-Kirkpatrick model and Random k-SAT	1
1.1 The Sherrington-Kirkpatrick model	1
1.2 Random k -SAT	3
1.3 Other Models	4
1.3.1 Random k -NAESAT	4
1.3.2 Random hypergraph 2-coloring	4
1.3.3 Random k -XORSAT	4
2 Disorder chaos in diluted models	7
2.1 Disorder chaos in the SK model	7
2.1.1 p -spin models	9
2.2 Disorder Chaos in the Diluted Random k -SAT model	10
2.2.1 Resampling clauses	12
2.2.2 Resampling random signs	12
2.3 Outline of the proof	13
3 Proof of Disorder Chaos in the random k-SAT model	15
3.1 Approximation to the fully connected model	15
3.2 Guerra-Toninelli interpolation method	19
3.3 Derivative of the Gaussian term	19
Gaussian integration by parts	20
3.4 Derivative of the Poissonian term	23
3.5 Bounds for the energy	27
4 Other Models	31
4.1 Disorder Chaos in random k -NAESAT	31
4.1.1 The k -NAESAT Hamiltonian	32
4.1.2 The k -NAESAT fully connected approximation	32
4.1.3 Covariance computation	34
4.2 Disorder Chaos in hypergraph 2-coloring	34
4.3 The Hypergraph 2-col Hamiltonian	35
4.3.1 The Hypergraph 2-coloring fully connected approximation	35
4.3.2 Covariance computation	37
4.4 Disorder Chaos in random k -XORSAT	37
4.4.1 The k -XORSAT Hamiltonian	37
4.4.2 The k -XORSAT fully connected approximation	38
4.4.3 Covariance computation	39

5	Simulations	41
5.1	Random k -SAT	41
5.1.1	Pseudo-code for the random k -SAT simulation	43
5.2	Random k -NAESAT	45
5.3	Random hypergraph 2-coloring	45
5.3.1	Pseudo-code for the hypergraph 2-col simulation	49
5.4	Random k -XORSAT	49
5.4.1	Pseudo-code for the simulated annealing (XORSAT)	52
6	Conclusions and Open Problems	53
6.1	k -XORSAT: Temperature Chaos	53
6.2	Conclusions	54
	Bibliography	55

Chapter 1

The Sherrington-Kirkpatrick model and Random k -SAT

The central aim of this thesis is to investigate the chaotic nature of the random constraint satisfaction problems, exploring how ideas from spin glass theory, and in particular disorder chaos in the SK model, can shed light on the structure and stability of the solution space.

In this introductory chapter we provide a basic description of what a **spin glass** is, with particular attention to the **Sherrington-Kirkpatrick** (SK) model, and its connection to the **random k -SAT model**, which belongs to the family of **random constraint satisfaction problems** (CSP).

In addition, we discuss other models, such as **random k -NAESAT**, **random k -XORSAT** and **random hypergraph 2-coloring**, that also play a role in this thesis.

Spin glasses are disordered magnetic systems that exhibit highly non-trivial collective behavior. Unlike conventional magnets, where the interactions between spins could be either **ferromagnetic**, meaning that the energy of the system is minimized by parallel spin configurations, or **antiferromagnetic**, where the energy of the system is minimized when spins are anti-parallel to each other, in spin glasses the spin interactions can be both ferromagnetic and antiferromagnetic. This feature, leads spin glasses to be characterized by **frustration** (Figure 1.1), the impossibility to simultaneously minimize all local interactions, and **disorder**, which comes from the random nature of the couplings between spins.

As a consequence, spin glasses typically display a complex energy landscape with a huge number of nearly degenerate states.

1.1 The Sherrington-Kirkpatrick model

A paradigmatic example of a spin glass is the **Sherrington-Kirkpatrick** (SK) model, introduced in 1975 in *Solvable Model of a Spin Glass* [1]. Its Hamiltonian H is given by

$$H(\sigma) = \frac{1}{\sqrt{n}} \sum_{i,j=1}^n J_{ij} \sigma_i \sigma_j, \quad \sigma \in \{-1, +1\}^n, \quad (1.1)$$

where the random variables J_{ij} , $1 \leq i, j \leq n$, are i.i.d. standard Gaussian random variables. In this model, every pair of spins (σ_i, σ_j) interacts through the random coupling J_{ij} that is equally likely to be positive (ferromagnetic) or negative (antiferromagnetic). The $1/\sqrt{n}$ prefactor ensures the extensivity of the model.

The simplification that the SK model introduces with respect to an actual spin glass,

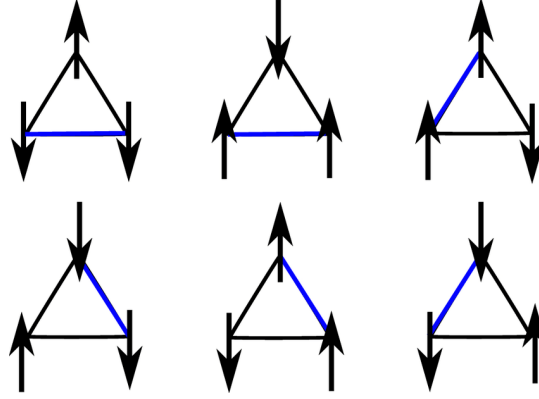


FIGURE 1.1: "Ising spins on triangular lattice. Six possible configurations for ground state are shown. Arrows indicate the spin direction. Blue lines denote the frustrated bonds, along which spins are parallel (Shaginyan, Vasily & Msezane, Alfred Z & Japaridze, G. & Clark, J. & Amusia, M. & Kirichenko, E.. (2020). Theoretical and experimental developments in quantum spin liquid in geometrically frustrated magnets: a review. *Journal of Materials Science*. 55. 1-34. 10.1007/s10853-019-04128-w.")

where interactions are usually short-ranged, is the fully connected hypothesis: each spin interacts with each other spin. This assumption corresponds to abandoning the underlying spatial structure in favor of a fully connected description: there is no notion of Euclidean distance between the positions of the spins, it is also called a fully connected model, since each spin interacts directly with all the others.

One central question is how to describe the complex structure of the energy landscape of this model. In the SK model this is captured by the **overlap** R between two configurations defined as

$$R(\sigma^1, \sigma^2) = \frac{1}{n} \sum_{i=1}^N \sigma_i^1 \sigma_i^2, \quad \sigma^1, \sigma^2 \in \{-1, +1\}^n. \quad (1.2)$$

Specifically, later, we will use the probabilistic properties of the overlap of two independent equilibrium configurations σ^1 and σ^2 .

In their original solution, Sherrington and Kirkpatrick assumed that the overlap distribution was trivial: all pairs of equilibrium states have the same overlap. This hypothesis is known as the **replica symmetric ansatz**, and it leads to an explicit formula for the free energy. However, it was soon discovered that this solution is unstable at low temperatures and leads to unphysical results, such as negative entropy.

A breakthrough came in 1979, when Giorgio Parisi proposed a new type of solution in his paper *Infinite Number of Order Parameters for Spin-Glasses*[2] involving a hierarchical organization of the energy landscape and a continuous distribution of the overlaps. This theory, known as **Replica Symmetry Breaking** (RSB), leads to a physically consistent and remarkably rich picture of the spin glass phase. Parisi's solution was later proven to be exact, and has become a cornerstone of statistical mechanics of disordered systems.

The SK model thus provided not only a mathematical tractable approximation of spin glasses, but also a conceptual framework for understanding complexity in high-dimensional disordered systems.

Another remarkable property of spin glasses, and in particular the SK model, is their sensitivity to perturbations: Already small changes in the disorder, for instance a small modification of the coupling parameters J_{ij} , can lead to drastic rearrangement of the equilibrium states.

This phenomenon is known as **disorder chaos** and reflects the rugged structure of the energy landscape: as the disorder is perturbed, the equilibrium configurations of the system reorganize in such a way that typical states in the original system may bear almost no similarity. In terms of the overlap: equilibrium configurations under one realization of the disorder may have vanishing overlap with respect to those equilibrium configurations of a slightly perturbed realization.

1.2 Random k -SAT

Formally a **Constraint Satisfaction Problems** is defined by: a set of n variables $\{x_1, \dots, x_n\}$ each taking values in a finite domain (e.g. $\{0, 1\}$ or a set of colors), and by a set of m constraints, each acting on a subset of the variables, specifying which combinations of values are allowed. The goal is to determine whether there exists at least one assignment of the variables that satisfies all the constraints.

If such an assignment exists, the instance is said to be **satisfiable** (SAT). Otherwise it is **unsatisfiable** (UNSAT).

The **k -SAT problem**, which lies at the heart of computational complexity theory, belongs to the class of decision problems. These are problems with a binary answer: True or False.

An instance of the k -SAT problem is given by a set of n Boolean variables $\{x_1, \dots, x_n\}$ each of which can take the value True or False, and a set of m constraints in the form of clauses. Each clause is a logical OR of k literals, where a literal is either a variable x_i or its negation $\neg x_i$. A formula Φ is a conjunction (logic AND) of clauses.

$$\Phi = C_1 \wedge \dots \wedge C_m \quad (1.3)$$

$$C_i = (\ell_1 \vee \dots \vee \ell_k) \quad (1.4)$$

$$\ell_i = x_i, \neg x_i \quad (1.5)$$

This particular form, a conjunction of disjunctions, is called **Conjunctive Normal Form** (CNF). A formula Φ is said to be satisfiable (SAT) if there exists an assignment of truth values to the variables that makes the entire formula true.

We can define a formula for the **random k -SAT problem** $\Phi = \Phi(k, n, m \sim Po(\lambda n))$ by taking $m \sim Po(\lambda n)$ clauses independently where each clause has size k and is created by selecting randomly uniformly from $\{x_1, \neg x_1, \dots, x_n, \neg x_n\}$. One of the most basic, yet extremely challenging, questions about random k -SAT is to determine the range of λ for which a solution exists with high probability (with probability tending to one as $n \rightarrow \infty$).

Interestingly, the notion of disorder chaos extends beyond the SK model and has a natural counterpart in CSPs. Indeed, we can see that there is some analogy between these two systems: In CSPs, instead of spins interacting through random couplings, we have variables constrained by randomly chosen clauses (spin-coupling $\longleftrightarrow$ variable-clause). The analogue of changing the disorder on a spin glass corresponds to modifying the set of constraints in the CSP instance.

Even small perturbations, such as **flipping the sign of a clause** or **resampling a**

small fraction of constraints can lead to dramatic changes in the structure of the solution space.

In this thesis we will investigate the chaotic nature of the random k -SAT problem, and related models, focusing on how small perturbations in the disorder affect the correlations between solutions, and how this phenomenon parallels the notion of disorder chaos in spin glass models.

1.3 Other Models

Beside the random k -SAT problem, several related random constraint satisfaction problems play an important role in this thesis. These models share structural similarities with k -SAT but introduce different types of constraints or symmetries, which make them valuable both for comparison and for highlighting the phenomena of disorder chaos.

In particular we will consider **random k -NAESAT**, the **random Hypergraph 2-col problem** and **random k -XORSAT**.

1.3.1 Random k -NAESAT

The random k -NAESAT problem is again defined on n Boolean variables $x_1, \dots, x_n$. We generate $m \sim Po(\lambda n)$ clauses, each consisting of k literals chosen uniformly and independently from $\{x_1, \neg x_1, \dots, x_n, \neg x_n\}$.

A clause is satisfied if **not all literals** take the same truth value, in other words the clause is violated only when all literals are simultaneously true or simultaneously false. Again, one of the central questions is to determine the range of λ for which a random instance is satisfied with high probability when $n \rightarrow \infty$.

1.3.2 Random hypergraph 2-coloring

A hypergraph is a generalization of a graph in which edges, called **hyperedges**, may connect more than two vertices. Random k -uniform hyperedges can be constructed selecting $m \sim Po(\lambda n)$ hyperedges, each consisting of k vertices chosen uniformly at random from the set of n vertices.

The random hypergraph 2-coloring problem asks whether it is possible to assign one of two colors to each vertex of a random k -uniform hypergraph in such a way that no hyperedge is **monochromatic**, i.e. every hyperedge contains at least one vertex of each color.

As in the other CSPs, one is interested in the satisfiability threshold, the critical density λ beyond which a proper 2-coloring ceases to exist with high probability.

1.3.3 Random k -XORSAT

Let $x_1, \dots, x_n \in \{0, 1\}$. Construct $m \sim Po(\lambda n)$ clauses independently, each of the form

$$x_{i_1} \oplus x_{i_2} \oplus \dots \oplus x_{i_k} = b \quad (1.6)$$

where indices $\{i_1, \dots, i_k\}$ are chosen uniformly at random and $b \in \{0, 1\}$ is chosen uniformly and independently. A clause is satisfied if the **XOR** (exclusive or) of the variables equals b .

An instance is satisfiable if the resulting linear system over $\mathbb{F}_2$ admits a solution.

So by comparing these models, which share structural similarities but differ in their symmetries and constraint types, we aim to understand whether disorder chaos is a model-specific phenomenon or a more universal property of random constraint satisfaction problems.

Chapter 2

Disorder chaos in diluted models

Traditionally, in the context of dynamical systems, chaos refers to the extreme sensitivity of trajectories to their initial conditions. In the context of the statistical physics of disordered systems (such as spin glasses), chaos takes on a slightly different but related meaning: a system is said to exhibit chaos when its macroscopic state responds dramatically to arbitrarily small perturbations of external parameters, such as temperature or the underlying disorder.

The phenomenon of chaos in disorder and temperature has been first discussed by Bray and Moore in the article *"Chaotic Nature of the Spin-Glass Phase"*. Bray and Moore showed that the description of the spin glass phase in terms of a $T = 0$ fixed point implies a chaotic phase in which the relative orientations of spins are sensitive to small changes δJ in the disorder or in the temperature δT . In particular, small changes in the disorder (or temperature) may result in dramatic changes in the location of the ground state with energy $\max_{\sigma} H_n(\sigma)$ as well as the energy landscape.

As discussed in Chapter 1, a way to measure whether the system is chaotic is to study the overlap $R_{12} = R(\sigma^1, \sigma^2) = \frac{1}{n} \sum_{i=1}^n \sigma_i^1 \sigma_i^2$ between two equilibrium configurations of the system.

Formally, one considers two systems with correlated disorder and studies the distribution of their mutual overlap under the Gibbs measure. Disorder chaos is said to occur when this overlap concentrates near zero as the system size n diverges.

In this thesis work, we are mainly going to focus on **disorder chaos**. In the next sections we will introduce the mathematical framework for this definition, following the approach of Chen and Panchenko in *Disorder chaos in some diluted spin glass models*[3], and analyze its manifestation in random k -SAT model.

2.1 Disorder chaos in the SK model

Consider two correlated SK Hamiltonians:

$$H^1(\sigma) = \frac{1}{\sqrt{n}} \sum_{i,j=1}^n J_{ij}^1 \sigma_i \sigma_j \quad \text{and} \quad H^2(\sigma) = \frac{1}{\sqrt{n}} \sum_{i,j=1}^n J_{ij}^2 \sigma_i \sigma_j,$$

where as before, (J_{ij}^1, J_{ij}^2) , $1 \leq i, j \leq n$, are n^2 pairs of i.i.d. jointly Gaussian random variables. We assume that $\mathbb{E}[J_{ij}^1] = \mathbb{E}[J_{ij}^2] = 0$ and $\text{Var}(J_{ij}^1) = \text{Var}(J_{ij}^2) = 1$ for all $1 \leq i, j \leq n$. Moreover, the disorders of the Hamiltonians are correlated in such a way that for all $1 \leq i, j \leq n$,

$$\mathbb{E}[J_{ij}^1 J_{ij}^2] = e^{-t}, \tag{2.1}$$

where $t \geq 0$ is a **correlation parameter**. We can see that if $t \approx 0$, then $J_{ij}^1 \approx J_{ij}^2$ with high probability, while if $t \rightarrow \infty$, the disorders of the two system become fully decoupled or independent.

For each system, we can define the Gibbs measure, i.e. the probability distribution of the configurations σ at inverse temperature β , as:

$$G_s(\sigma) = \frac{\exp(\beta H^s(\sigma))}{Z^s}, \quad s \in \{1, 2\}, \quad (2.2)$$

where Z^s is the partition function of the system defined as:

$$Z^s = \sum_{\sigma} \exp(\beta H^s(\sigma)). \quad (2.3)$$

Then, it is possible to sample from the Gibbs measures two independent (given the disorder) configurations:

$$\sigma^1 \sim G_1 \quad \text{and} \quad \sigma^2 \sim G_2$$

and measure how these two coincide by evaluating their overlap (previously defined in Chapter 1)

$$R_{12}(t) = \frac{1}{n} \sum_{i=1}^n \sigma_i^1 \sigma_i^2.$$

A model shows **disorder chaos** if for a small positive correlation parameter t , the overlap $R_{12}(t)$ is close to zero. The mathematical proof of disorder chaos in the SK model at all temperatures β was given by S. Chatterjee in *Disorder chaos and multiple valleys in spin glasses* [4]. A concise statement of the following result can also be found in the book *Superconcentration and Related Topics* [5]:

Theorem (Disorder chaos in the SK model, [5, Theorem 1.11]). *For any integer $k \geq 1$,*

$$\mathbb{E} \left(R_{12}^{2k}(t) \right) \leq (C_1(\beta)k)^k n^{-C_2(\beta)k \min\{1, t\}},$$

where $C_1(\beta)$ and $C_2(\beta)$ are positive constants that depend only on β .

This theorem provides a quantitative formulation of disorder chaos.

It shows that the overlap between two systems with correlated disorder decreases polynomially fast with the system size.

More precisely, for any fixed correlation parameter $t > 0$ the expected even moments for the overlap $R_{12}(t)$ vanish as $n \rightarrow \infty$.

This implies that, although the two Hamiltonians share almost the same realization of the disorder, their equilibrium configurations become asymptotically uncorrelated.

In physics terms we could say that an infinitesimal perturbation of the disorder, corresponding to the resampling of a small fraction of the couplings, is sufficient to produce a macroscopic rearrangement of the Gibbs states.

In other words, even when the two disorders realization are highly correlated, configurations σ sampled from the Gibbs measure become **asymptotically orthogonal**.

This implies that arbitrarily **small perturbations** in the disorder lead to a complete reorganization of the system's equilibrium state.

Generally, the presence of disorder chaos is well understood for **fully connected models**, such as the SK model.

The previous theorem by Chatterjee can alternatively stated as:

Theorem (Chaos under independent flips, [5, Theorem 7.6]). *Consider the SK model at inverse temperature β . Take any integer n and $p \in [0, 1]$. Suppose that a randomly chosen fraction p of the couplings $(g_{ij})_{1 \leq i < j \leq n}$ are replaced by independent copies to give a perturbed Gibbs measure.*

Let σ^1 be chosen from the original Gibbs measure and σ^2 is chosen from the perturbed measure. Let R_{12} be the overlap between the two configurations. Then:

$$\mathbb{E}(R_{12}^2) \leq \frac{C(\beta)}{p \log n},$$

where $C(\beta)$ depends only on β .

The fully connected structure makes these models analytically tractable, and rigorous results establish that disorder chaos occurs as soon as the disorder realizations are not identical.

However, the main focus of this thesis lies in **diluted systems**, models in contrast to the mean field ones, diluted systems refer to a sparse version of a model. Each spin interacts with a finite number of randomly chosen others. The diluted setting is therefore closer to systems on sparse random graphs, such as random k -SAT, random k -NAESAT, random hypergraph 2-coloring and random k -XORSAT. To bridge the gap between these two settings, it is useful to introduce the **mixed p -spin** model as a fully connected approximation.

2.1.1 p -spin models

The p -spin model is an infinite-range interaction (or mean-field) model, it was invented in the course of the theoretical study of spin-glasses: here, the parameter p denotes the **order of the interaction**, so that the interactions are restricted to groups of exactly p -spins. Fixing $p \geq 1$, its Hamiltonian is thus defined as

$$H_p(\sigma) = -\frac{1}{n^{\frac{p-1}{2}}} \sum_{1 \leq i_1 < i_2 < \dots < i_p \leq n} J_{i_1, \dots, i_p} \sigma_{i_1} \cdots \sigma_{i_p}, \quad (2.4)$$

where the random variables $J_{i_1, \dots, i_p}$ are standard Gaussian and independent for all $(i_1, \dots, i_p)$. Hence, the Hamiltonian is a homogeneous function of the spins, it only contains terms of the form of p -spin interactions.

For instance for $p = 2$ we recover the SK model:

$$H_2(\sigma) = \frac{1}{\sqrt{n}} \sum_{1 \leq i < j \leq n} J_{ij} \sigma_i \sigma_j.$$

This Hamiltonian describes **pairwise** interactions between spins. Larger values of p describe higher-order interactions among spins. For example, for the $p = 3$ case one obtains interactions among triples of spins:

$$H_3(\sigma) = \frac{1}{n} \sum_{i, j, k=1}^n J_{ijk} \sigma_i \sigma_j \sigma_k.$$

More generally, if we consider a linear combination of the several **pure** interaction terms with deterministic weights $(\beta_p)_{p \geq 2}$, we define a **mixed p -spin** model, which

combines interactions of different orders

$$H(\sigma) = \sum_{p \geq 2} \beta_p H_p(\sigma), \quad (2.5)$$

where the case $p = 1$ is excluded since such a term would correspond to a random field acting independently on each site rather than to an interaction among spins. Including it would also break the global spin flip symmetry, which is characteristic of spin glass models.

To ensure that the Hamiltonian is well-defined when the Hamiltonian includes infinitely many term, we assume that the weights $(\beta_p)_{p \geq 2}$ decrease fast enough, for example by imposing the condition $\sum_p 2^p \beta_p^2 < \infty$. A key feature is that its covariance depends only on the overlap $R_{1,2} = \frac{1}{n} \sum_i \sigma_i^1 \sigma_i^2$:

$$\mathbb{E} H(\sigma^1) H(\sigma^2) = n \sum_{p \geq 2} \beta_p^2 (R_{1,2})^p.$$

So if we define a function:

$$\xi(x) = \sum_{p \geq 2} \beta_p^2 x^p,$$

the covariance can be written more compactly as:

$$\mathbb{E} H(\sigma^1) H(\sigma^2) = n \xi(R_{1,2}).$$

Thus, the function ξ completely encodes the structure of the mixed p -spin model.

2.2 Disorder Chaos in the Diluted Random k -SAT model

As discussed in the previous section, disorder chaos has been rigorously established in fully connected models. In this section, we turn to diluted systems, focusing on the work done by Chen and Panchenko [3] on the random k -SAT model, where they proved disorder chaos at zero temperature for large connectivity.

Chen and Panchenko show that resampling a small proportion of disorders (or signs), as well as clauses (or configurations), leads to configurations that **nearly maximize** the original Hamiltonian becoming **nearly orthogonal** to those that nearly maximize the perturbed Hamiltonian, with high probability.

In Chapter 1 we have briefly discussed the random k -SAT model, let us just recall what this model is:

Let $x_1, \dots, x_n$ be n Boolean variables which take values in $\{0, 1\}$ (where 0=False, 1=True). We can construct a k -SAT formula $\Phi = \Phi(k, n, m)$ by selecting $m \sim Po(\lambda n)$ clauses independently. Each clause of size k is obtained by choosing independently and uniformly at random each literal from $\{x_1, \neg x_1, \dots, x_n, \neg x_n\}$. The central question in random k -SAT is to determine, as a function of the density parameter λ , whether the formula admits a satisfying assignment, and to what extent such an assignment can be found efficiently by algorithms.

Importantly, random k -SAT model can be recast as a spin glass model.

We can define a **clause** as the following random function $\theta(\sigma_1, \dots, \sigma_k)$ on $\{-1; +1\}^k$:

$$\theta(\sigma_1, \dots, \sigma_k) := - \prod_{j \leq k} \frac{1 + J_j \sigma_j}{2}, \quad (2.6)$$

where $(J_j)_{j \geq 1}$ are i.i.d. Rademacher random variables and $(\sigma_j)_{j \geq 1} \in \{-1; +1\}^k$ represents an assignment of the variables in a given clause to specific truth values. Now, we can see that the definition of this random function θ is actually a mapping of the instances of the random k -SAT:

$$\theta(\sigma_1, \dots, \sigma_k) = \begin{cases} 0, & \text{if and only if there exists at least one } j \text{ such that } \sigma_j \neq J_j, \\ -1, & \text{if } \sigma_j = J_j \ \forall j \leq k. \end{cases}$$

So the random function θ takes value -1 when all literals in the clause are unsatisfied, and 0 otherwise. Then, denote by θ_c independent copies of the function θ for various indices c :

$$\theta_c(\sigma_1, \dots, \sigma_k) := - \prod_{j \leq k} \frac{1 + J_{j,c} \sigma_j}{2},$$

with i.i.d. copies $J_{j,c}$.

Then, introducing a **connectivity parameter** $\lambda > 0$, we can write the Hamiltonian of the model as:

$$H_\lambda(\sigma) = \sum_{c \leq \pi(\lambda N)} \theta_c(\sigma_{i_{1,c}}, \dots, \sigma_{i_{k,c}}), \quad (2.7)$$

where $\pi(\lambda n)$ is a Poisson random variable with mean λn , and the coordinate indices $i_{k,c}$ are independent for different pairs (k, c) and are chosen uniformly from $\{1, \dots, n\}$.

Thus, the Hamiltonian counts, up to a minus sign, the number of violated clauses in the random k -SAT instance. This makes the model comparable to the spin glass Hamiltonian, where energy contributions arise from random interactions among subsets of variables.

The analogy between these two systems becomes clearer if one interprets spins as the variable assignment within a clause θ and J_j 's signs as the random interaction coefficient in a spin glass system.

This analogy is key because it allows us to reinterpret the satisfiability problem as a disordered statistical mechanics system, and therefore to study phenomena such as disorder chaos in the same mathematical framework used for spin glasses.

In particular, we want to find the assignment of variables σ_i that maximizes the number of satisfied clauses, which for a given clause means that at least one $\sigma_j = -J_j$ for $1 \leq j \leq k$.

Within this framework, let us consider two copies H_λ^1 and H_λ^2 of the Hamiltonian $H_\lambda(\sigma) = \sum_{c \leq \pi(\lambda n)} \theta_c(\sigma_{i_{1,c}}, \dots, \sigma_{i_{k,c}})$. Similarly to Chapter 1, we will fix a correlation parameter t , but now we assume that $t \in [0, 1]$.

These copies can be defined either by **resampling clauses** or **resampling signs**, the latter will be the main approach in this thesis work. These two approaches will be described in the following subsections.

2.2.1 Resampling clauses

The two Hamiltonians H_λ^1 and H_λ^2 will have $\text{Po}(t\lambda n)$ **common** clauses and two independent $\text{Po}((1-t)\lambda n)$ **independently** generated clauses. So in this case one would resample both indices $(i_{k,j})$, and signs (J) .

So for $\ell = 1, 2$ the Hamiltonians will have the form:

$$H_\lambda^\ell = - \sum_{c \leq \pi(t\lambda n)} \prod_{j \leq k} \frac{1 + J_{j,c} \sigma_{j,c}}{2} - \sum_{c \leq \pi_\ell((1-t)\lambda n)} \prod_{j \leq k} \frac{1 + J_{j,c}^\ell \sigma_{j,c}}{2}. \quad (2.8)$$

Where $\pi(t\lambda n), \pi_1((1-t)\lambda n), \pi_2((1-t)\lambda n)$ are independent Poisson random variables with means $t\lambda n$ and $(1-t)\lambda n$.

Notice also the role that the correlation parameter t plays in this kind of resampling: if $t = 1$, only the first term of H_λ^ℓ exists, meaning that the two systems will share the same set of $\pi(\lambda n)$ clauses, they are fully correlated. Whereas, if $t = 0$ the two systems are going to be independent.

2.2.2 Resampling random signs

Here, the number of clauses $\pi(t\lambda n)$ will be the same as well as their indices, since we are resampling only random signs.

In order to resample random signs we replace the random variable J_j in each clause by two correlated copies J_j^1 and J_j^2 such that their correlation is $\mathbb{E}J_j^1 J_j^2 = t$. This procedure can be done in two ways:

1. Resample with probability $1 - t$ all random signs $J_{1,j}^1, \dots, J_{k,j}^1$ simultaneously to produce $J_{1,j}^2, \dots, J_{k,j}^2$ independently for each clause θ_j .
In this setting for each clause j , the entire vector $(J_{1,j}^1, \dots, J_{k,j}^1)$ is replaced by an independent clause with probability $1 - t$ and kept unchanged with probability t .
Consequently each pair $(J_{k,j}^1, J_{k,j}^2)$ satisfies $\mathbb{E}J_{k,j}^1 J_{k,j}^2 = t$.
2. Replace each $J_{k,j}^1$ with probability $1 - t$ to produce $J_{k,j}^2$, and keep it unchanged with probability t such that $\mathbb{E}J_{k,j}^1 J_{k,j}^2 = t$. Notice that in this way, each copy $(J_{k,j}^1, J_{k,j}^2)$ is independent for all different pairs (k, j) , while with the first setting the pairs $(J_{k,j}^1, J_{k,j}^2)$ within the same clause are correlated.

These equivalent procedures of resampling act like an independent spin flip, a perturbation. We do this because we want to compare the Hamiltonian with an independent copy of the latter, which acts like a perturbed Hamiltonian. In such a way, we have two independent systems.

For all these procedures described above and for large connectivity λ , with high probability all near maximizers are nearly orthogonal to each other.

This is stated in the **theorem** that we are going to prove in the next chapter, by Chen and Panchenko from [3, Theorem 1].

Theorem (Disorder Chaos in Diluted Random k -SAT [3, Theorem 1]). *For any $\varepsilon, t \in (0, 1)$ there exists small enough $\eta > 0$ such that for large enough λ the following holds for large enough n with probability at least $1 - Le^{-N\eta^2/L}$: for any configurations $\sigma^1, \sigma^2 \in V$*

that nearly maximize the corresponding Hamiltonian

$$\frac{1}{n} H_\lambda^\ell(\sigma^\ell) \geq \frac{1}{N} \max_{\sigma \in V} H_\lambda^\ell(\sigma) - \eta \sqrt{\lambda} \quad \text{for } \ell = 1, 2, \quad (2.9)$$

the overlap $R_{1,2} = \frac{1}{n} \sum_{i \leq n} \sigma_i^1 \sigma_i^2$ between them satisfies $|R_{1,2}| \leq \varepsilon$.

This theorem states that when both the system size n and the connectivity parameter λ are large, the configurations that nearly maximize the energy in the system are almost orthogonal to those that nearly maximize the energy of the perturbed system. In other words, even if the two systems differ only by a small perturbation in the disorder, their sets of near ground states have vanishing overlap.

This is precisely the manifestation of disorder chaos in diluted random k -SAT: arbitrarily small perturbations of the disorder lead to a complete reorganization of the near optimal solutions.

To conclude this section, let us observe that this implies the existence of exponentially many in n near maximizers of the Hamiltonian H_λ which are nearly orthogonal to each other.

2.3 Outline of the proof

To conclude this chapter, let us see a high-level description of the proof of Theorem 2.2.2. The proof will be fully discussed with more details in the next chapter.

The main idea is to interpolate between the diluted model and its fully connected approximation and study the free energy under the constraint $|R_{12}| > \varepsilon$. The derivative of the interpolated free energy will split into a Gaussian term (coming from the fully connected model) and a Poisson term (coming from the diluted model). If the overlap were to stay bounded away from 0, these two contributions would lead to incompatible estimates for the increment of the free energy. This contradiction shows that the overlap must concentrate near zero, which is the manifestation of disorder chaos.

To prove this theorem we need to introduce the fully connected approximation of the random k -SAT Hamiltonian. The latter is given by the mixed p -spin Hamiltonian

$$H(\sigma) = \sum_{p=1}^k \sqrt{\binom{k}{p} \frac{1}{n^{p-1}}} \sum_{1 \leq i_1, \dots, i_p \leq n} g_{i_1, \dots, i_p} \sigma_{i_1} \cdots \sigma_{i_p}, \quad (2.10)$$

where the coefficients $(g_{i_1, \dots, i_p})$ are $\mathcal{N}(0, 1)$ random variables indexed by $1 \leq i_1 < \dots < i_p \leq n$, $1 \leq p \leq k$, which are independent for all possible index choices.

Given this approximation it was also shown in [6, 7] that:

$$\frac{1}{n} \mathbb{E} \max_{\sigma \in V} H_\lambda(\sigma) = -\frac{\lambda}{2^k} + \frac{\sqrt{\lambda}}{2^k} \frac{1}{n} \mathbb{E} \max_{\sigma \in V} H(\sigma) + \mathcal{O}(\lambda^{1/3}) \quad (2.11)$$

as $\lambda \rightarrow \infty$, uniformly in n .

The theorem by Chen and Panchenko will be proved by contradiction.

We are going to suppose that exist two correlated system with the same near maximizers.

First we will introduce an **interpolating Hamiltonian**:

$$H(s, \sigma^1, \sigma^2) = \sum_{\ell=1}^2 (\delta H_{\lambda(1-s)}^\ell(\sigma^\ell) + \sqrt{s} \beta H^\ell(\sigma^\ell)), \quad (2.12)$$

where $s \in [0, 1]$ is the interpolation parameter and the correlated Hamiltonians $H_{\lambda(1-s)}^\ell$, for $\ell = 1, 2$, are defined in the same way as the diluted Hamiltonians (2.7) previously introduced in this chapter, and δ and β are two inverse temperature parameters, and will be chosen later in the proof. The Hamiltonian (2.12) considers configuration pairs (σ^1, σ^2) instead of single configurations σ because the interpolation is designed to couple two correlated spin system whose disorders share a prescribed correlation structure.

This allows one to track how the overlap $R_{12} = \frac{1}{n} \sum_i \sigma_i^1 \sigma_i^2$ between the two systems evolves as the correlation in the disorder changes, a key quantity in the analysis of disorder chaos.

Moreover $H(0, \sigma^1, \sigma^2)$ corresponds to:

$$H(0, \sigma^1, \sigma^2) = \sum_{\ell=1}^2 \delta H_{\lambda(1-s)}^\ell(\sigma^\ell),$$

two correlated diluted models at inverse temperature δ , and $H(1, \sigma^1, \sigma^2)$ is:

$$H(1, \sigma^1, \sigma^2) = \sum_{\ell=1}^2 \sqrt{s} \beta H^\ell(\sigma^\ell),$$

corresponding to two correlated fully connected models at inverse temperature β . In particular, for the fully connected model at $s = 1$, disorder chaos is already known to occur. After the interpolated Hamiltonian, it is useful to introduce the **interpolated free energy**, defined as:

$$\varphi(s) = \frac{1}{n} \mathbb{E} \sum_{|R_{12}| > \varepsilon} \exp H(s, \sigma^1, \sigma^2) \quad (2.13)$$

of the two correlated systems, coupled by the overlap constraint $|R_{12}| > \varepsilon$.

The central part of the proof will be about the extended computation of the derivative of this free energy density with respect to s .

The derivative naturally decomposes into two terms: a Gaussian contribution from the fully connected approximation and a Poissonian contribution from the diluted system.

The essence of the proof is to show that these two contributions cannot be simultaneously consistent if the overlap remains bounded away from zero, which yields the desired contradiction and establishes disorder chaos.

Chapter 3

Proof of Disorder Chaos in the random k -SAT model

In this chapter we will go through the detailed proof at the heart of the paper from Chen and Panchenko, about the existence of disorder chaos in the diluted random k -SAT model.

The aim of this chapter is to provide a rigorous derivation of the proof, previously described in Chapter 2, comparing the diluted model of the random k -SAT with its fully connected approximation, by means of an interpolated Hamiltonian.

We begin by recalling the definition of the diluted Hamiltonian and its fully connected approximation. We then introduce the interpolating Hamiltonian and the associated free energy, constrained by the overlap condition. A key step is the computation of the derivative of this interpolated free energy, which naturally splits into two contributions: a Gaussian term, coming from the fully connected approximation, and a Poissonian term, coming from the diluted system.

The central part of the chapter is devoted to analyzing these two contributions and showing that they cannot be consistent under the assumption that the overlap between near maximizers is bounded away from zero. This contradiction completes the proof, thereby establishing disorder chaos in the diluted random k -SAT model.

3.1 Approximation to the fully connected model

In this section, we argue that it is reasonable to approximate the random k -SAT model by a mixed p -spin model for large values of λ , and large but fixed number of clauses n .

Most importantly we want to show that the covariance of the fully connected model, depends only on the overlap between two configurations.

Claim. *The k -SAT model Hamiltonian*

$$H_\lambda(\sigma) = - \sum_{c \leq \pi(\lambda n)} \prod_{j \leq k} \frac{1 + J_{j,c} \sigma_{i_j,c}}{2} \quad (3.1)$$

can be approximated by the following mixed p -spin Hamiltonian:

$$H(\sigma) = \sum_{p=1}^k \sqrt{\binom{k}{p} \frac{1}{n^{p-1}}} \sum_{1 \leq i_1, \dots, i_p \leq n} g_{i_1, \dots, i_p} \sigma_{i_1} \cdots \sigma_{i_p}, \quad (3.2)$$

where the coefficients $(g_{i_1, \dots, i_p})$ are $\mathcal{N}(0, 1)$ random variables indexed by $1 \leq i_1, \dots, i_p \leq n$, $1 \leq p \leq k$, which are independent for all possible index choices.

Proof. Let us consider a single clause

$$\theta_c(\sigma_{i_1}, \dots, \sigma_{i_k}) = - \prod_{j=1}^k \frac{(1 + J_{j,c} \sigma_{i_j, c})}{2}.$$

Consider a fixed clause and let $k = 3$, then expand the product (for simplicity we will drop the index c in the example, since the clause is fixed):

$$- \prod_{j=1}^k \frac{(1 + J_j \sigma_{i_j})}{2} = - \frac{1}{2^3} (1 + J_1 \sigma_{i_1}) (1 + J_2 \sigma_{i_2}) (1 + J_3 \sigma_{i_3}).$$

By explicitly computing this product, we can group together the terms that have the same type of interactions: among the 8 terms of the product, we will have single interaction terms ($p = 1$), pairwise interactions ($p = 2$), interaction between three terms ($p = 3$) and a constant term ($p = 0$):

$$1 + J_1 \sigma_{i_1} + J_2 \sigma_{i_2} + J_3 \sigma_{i_3} + J_1 J_2 \sigma_{i_1} \sigma_{i_2} + J_1 J_3 \sigma_{i_1} \sigma_{i_3} + J_2 J_3 \sigma_{i_2} \sigma_{i_3} + J_1 J_2 J_3 \sigma_{i_1} \sigma_{i_2} \sigma_{i_3}.$$

Then, a single clause can be recasted in a more compact form, taking into account the order of interactions, as:

$$-2^{-k} \sum_{p=0}^k \sum_{\substack{S \subseteq [k] \\ |S|=p}} \prod_{j \in S} J_j \sigma_{i_j},$$

so each subset $S \subseteq \{1, \dots, k\}$ with cardinality $|S| = p$ gives a term of order p . Moreover, we can define the following form for a single clause:

$$X_c(\sigma) := \theta_c + 2^{-k} = -2^{-k} \sum_{p=1}^k \sum_{\substack{S \subseteq [k] \\ |S|=p}} \prod_{j \in S} J_j \sigma_{i_j},$$

in order to absorb the constant term. By doing so, since $J_{j,c}$'s are i.i.d. Rademacher and indices i_j are i.i.d uniformly distributed over $\{1, \dots, n\}$ we can say that:

$$\mathbb{E}[X_c] = 0.$$

The Hamiltonian becomes:

$$H_\lambda(\sigma) = - \frac{\pi(\lambda n)}{2^k} + \sum_{c \leq \pi(\lambda n)} X_c(\sigma).$$

We can firstly compute the covariance for a single clause. Start by calculating the expected value over the signs:

$$\mathbb{E}_J[X(\sigma^1)X(\sigma^2)] = 2^{-2k} \sum_{p=1}^k \sum_{|S|=p} \prod_{j \in S} \sigma_{i_j}^1 \sigma_{i_j}^2,$$

Finally, we can define a rescaling for the Hamiltonian in such a way to eliminate the prefactor $2^{-2k}\lambda n$, as:

$$\hat{H}_\lambda(\sigma) := \frac{2^k}{\sqrt{\lambda}} \sum_{c \leq M} X_c(\sigma).$$

In this way, the covariance reads:

$$\mathbb{E} [\hat{H}_\lambda(\sigma^1) \hat{H}_\lambda(\sigma^2)] = n \sum_{p=1}^k \binom{k}{p} R_{12}^p.$$

The dependence of the covariance only on the overlap $R_{1,2}$ plays a crucial role in the analysis of disorder chaos.

This structure allows the correlation between perturbed copies of the system to be entirely encoded in the overlap between configurations. As a result, the response of the system to small perturbations in the disorder can be effectively tracked by observing the behavior of the overlap as a function of the correlation parameter t . In particular, the sharp transition in the overlap indicates the existence of disorder chaos.

Fix $p \in \{1, \dots, k\}$ and an ordered p -uple $(i_1, \dots, i_p) \in \{1, \dots, n\}^p$.

For a given clause c , let:

$$Y_c^{(i_1, \dots, i_p)} := \sum_{\substack{S \subseteq [k] \\ |S|=p}} \mathbb{1}\{(i_{j,c})_{j \in S} = (i_1, \dots, i_p)\} \prod_{j \in S} J_{j,c},$$

which is the total contribution of a clause c to the term $\sigma_{i_1} \cdots \sigma_{i_p}$ obtained by summing over all p -subsets S of positions inside the clause, the product of corresponding signs whenever the clause chooses exactly the indices $(i_1, \dots, i_p)$ at those positions. Then define:

$$C_{i_1, \dots, i_p}^\lambda := -\frac{1}{\sqrt{\lambda}} \sum_{c \leq M} Y_c^{(i_1, \dots, i_p)},$$

and rewrite the Hamiltonian as:

$$\hat{H}_\lambda(\sigma) = \sum_{p=1}^k \sum_{1 \leq i_1, \dots, i_p \leq n} C_{i_1, \dots, i_p}^\lambda \sigma_{i_1} \cdots \sigma_{i_p}.$$

In this way we have created a sum over clauses, so i.i.d. blocks, over which we can use the Central Limit Theorem.

Notice that:

$$\mathbb{E} Y_c = 0$$

$$\begin{aligned} \text{Var}(Y_c) &= \mathbb{E}[Y_c^2] = \sum_{\substack{S_1, S_2 \subseteq [k] \\ |S_1|=|S_2|=p}} \mathbb{P}((i_{j,c})_{j \in S_1} = (i_{j,c})_{j \in S_2} = (i_1, \dots, i_p)) \mathbb{E} \left[\prod_{j \in S_1} J_{j,c} \prod_{j \in S_2} J_{j,c} \right] \\ &= \sum_{\substack{S \subseteq [k] \\ |S|=p}} \frac{1}{n^p} = \binom{k}{p} n^{-p}. \end{aligned}$$

To perform the CLT with a Poisson number of clauses we could condition over the number of clauses:

for $M = m : \{Y_c\}_{c=1}^m$ are i.i.d. with variance $\pi_{p,n}$.

For the CLT: $\frac{1}{\sqrt{m}} \sum_{c=1}^m Y_c = \mathcal{N}(0, \pi_{p,n})$.

It follows that: $C_{i_1, \dots, i_p}^\lambda | M = m \implies \mathcal{N}(0, (m/\lambda)\pi_{p,n})$.

Now, for $M \sim \text{Po}(\lambda n)$ it holds:

$$\mathbb{E}M = \lambda n, \quad \text{Var}(M) = \lambda n$$

when: $\lambda n \rightarrow \infty \implies \frac{M}{\lambda n} \rightarrow 1$

So:

$$C_{i_1, \dots, i_p}^\lambda \implies \mathcal{N}(0, n\pi_{p,n}) = \mathcal{N}\left(0, n^{1-p} \binom{k}{p}\right), \quad \lambda n \rightarrow \infty$$

Now, we can define:

$$g_{i_1, \dots, i_p} := \frac{n^{\frac{p-1}{2}}}{\sqrt{\binom{k}{p}}} C_{i_1, \dots, i_p}^\lambda$$

Then: $\text{Var}(g_{i_1, \dots, i_p}^\lambda) = 1 \implies g_{i_1, \dots, i_p}^\lambda \sim \mathcal{N}(0, 1)$. Substituting back we get:

$$\hat{H} = \sum_{p=1}^k \sqrt{\binom{k}{p}} n^{-\frac{p-1}{2}} \sum_{1 \leq i_1, \dots, i_p \leq n} g_{i_1, \dots, i_p}^\lambda \sigma_{i_1} \cdots \sigma_{i_p}$$

□

Remark (Heuristic nature of the Gaussian limit). *The passage from the diluted Hamiltonian to the fully connected Gaussian model is heuristic. Indeed, the clause contributions are neither independent nor identically distributed: they share indices and create non-trivial dependencies across products $\sigma_{i_1} \cdots \sigma_{i_p}$. The approximation above relies on the fact that, in the large-connectivity regime $\lambda \rightarrow \infty$ (with n large and $M \sim \text{Po}(\lambda n)$), each coefficient $C_{i_1, \dots, i_p}^\lambda$ is a sum of many weakly dependent, centered terms with variance of order $n^{1-p} \binom{k}{p}$. Under this high-connectivity scaling, a central-limit principle suggests that the vector of coefficients becomes approximately Gaussian after centering and normalization, and cross-covariances vanish at leading order. What is used in the sequel is precisely the covariance structure*

$$\mathbb{E} \hat{H}_\lambda(\sigma^1) \hat{H}_\lambda(\sigma^2) \approx n \sum_{p=1}^k \binom{k}{p} R_{12}^p,$$

rather than a fully rigorous CLT for all coefficients. Establishing a complete normal approximation with quantitative error bounds is delicate and beyond the scope of this thesis. Our use of the mixed p -spin surrogate should therefore be viewed as an effective mean field approximation: the Gaussian distribution is an effective description of the collective behavior, not a rigorous convergence result.

Our heuristic fully connected approximation can in principle be made rigorous via a Lindeberg type replacement argument, done by Chatterjee in "A generalization of the Lindeberg

principle" [8], provided one establishes appropriate bounds on influences and dependencies of clause contributions. We do not attempt this verification here.

3.2 Guerra-Toninelli interpolation method

For $s \in [0, 1]$, let us consider the following interpolating Hamiltonian

$$H(s, \sigma^1, \sigma^2) = \sum_{l=1}^2 (\delta H_{\lambda(1-s)}^l(\sigma^l) + \sqrt{s} \beta H^l(\sigma^l))$$

Then let us recall the interpolated free-energy of correlated systems coupled by the overlap constraint $|R_{1,2}| > \varepsilon$:

$$\varphi(s) = \frac{1}{N} \mathbb{E} \log \sum_{|R_{1,2}| > \varepsilon} \exp(H(s, \sigma^1, \sigma^2))$$

We are interested in computing the derivative of the free energy with respect to s : in order to do that we need to decompose the Hamiltonian as the sum of two terms introducing a change of variables.

$$H(s, \sigma^1, \sigma^2) = u(s, \sigma^1, \sigma^2) + v(s, \sigma^1, \sigma^2)$$

So now the free energy will be defined as:

$$\varphi(s) = \Phi(u(s), v(s))$$

Now, the computation of the derivative of the free energy takes the following form:

$$\frac{d}{ds} \varphi(s) = \frac{\partial \Phi}{\partial u} \frac{du}{ds} + \frac{\partial \Phi}{\partial v} \frac{dv}{ds}$$

This derivative gives rise to two separate terms, this allows us to rewrite the derivative of the free energy, as the sum of two terms:

$$\frac{d}{ds} \varphi = I + II$$

Respectively:

$$I = \left\langle \frac{\partial}{\partial s} Gauss \right\rangle_s \text{ and } II = \left\langle \frac{\partial}{\partial s} Poiss \right\rangle_s$$

3.3 Derivative of the Gaussian term

In this section, we will see the explicit derivation of the first term in the derivative of the free energy.

The latter being the derivative with respect to s of the Gaussian part of the interpolated Hamiltonian, corresponding to the fully connected approximation of the

model.

$$\begin{aligned}
I &= \frac{\beta}{2\sqrt{sn}} \frac{\partial}{\partial u} \mathbb{E} \log \sum_{|R_{12}| > \varepsilon} \exp \left(H(s, \sigma^1, \sigma^2) \right) \\
&= \frac{\beta}{2\sqrt{sn}} \mathbb{E} \frac{\sum_{\ell} H^{\ell}(\sigma^{\ell}) \exp(H(s, \sigma^1, \sigma^2))}{\sum_{|R_{12}| > \varepsilon} \exp(H(s, \sigma^1, \sigma^2))} \\
&= \frac{\beta}{2\sqrt{sn}} \mathbb{E} \left\langle \sum_{\ell} H^{\ell}(\sigma^{\ell}) \right\rangle_s
\end{aligned}$$

Where we used:

$$G_s(\sigma^1, \sigma^2) = \frac{\exp H(s, \sigma^1, \sigma^2)}{\sum_{|R_{12}| > \varepsilon} \exp H(s, \sigma^1, \sigma^2)} \quad (3.3)$$

which is the average with respect to the Gibbs measure on $\{(\sigma^1, \sigma^2) \in V^2 : |R_{1,2}| > \varepsilon\}$ corresponding to the Hamiltonian $H(s, \sigma^1, \sigma^2)$.

Gaussian integration by parts

To fully get the explicit expression in terms of the overlaps for this first contribution, let us focus on how to compute the expected value of the average with respect to the Gibbs measure. The whole expression can be written as follows:

$$\mathbb{E} \left\langle H^1(\sigma^1) + H^2(\sigma^2) \right\rangle_s = \sum_{\sigma^1, \sigma^2} \mathbb{E} \left[(H^1(\sigma^1) + H^2(\sigma^2)) G_s(\sigma^1, \sigma^2) \right]$$

Then, defining: $H^1(\sigma^1) + H^2(\sigma^2) := x(\sigma)$, we have:

$$\mathbb{E} \langle x(\sigma) \rangle_s = \sum_{\sigma} \mathbb{E} [x(\sigma) G_s(\sigma)]$$

Lemma (Chatterjee[5], Talagrand[9]). *Suppose g is a Gaussian random variable, and $f : \mathbb{R} \rightarrow \mathbb{R}$ is an absolutely continuous function. Under the assumption that $\mathbb{E}|f'(x)| < \infty$, an application of integration by parts gives the identity:*

$$\mathbb{E} g f(g) = \mathbb{E} f'(g)$$

This identity can be generalized to n dimensions. Suppose that $g = (g_1, \dots, g_n)$ is a centered Gaussian vector. If $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is an absolutely continuous function such that $|\nabla f(g)|$ has finite expectation, then for any i ,

$$\mathbb{E}(g_i f(g)) = \sum_{j=1}^n \mathbb{E}(g_i g_j) \mathbb{E}(\partial_i f(g))$$

Using the above Lemma, we can explicitly compute the expected value over the Hamiltonian as:

$$\sum_{\sigma} \mathbb{E} [x(\sigma) G_s(\sigma)] = \sum_{\sigma} \sum_{\tau} \mathbb{E} [x(\sigma) x(\tau)] \mathbb{E} \left[\frac{\partial G_s(\sigma)}{\partial x(\tau)} \right]$$

Now let us evaluate the derivative:

$$\begin{aligned}
& \frac{\partial}{\partial x(\tau)} \frac{\exp(\sqrt{s}\beta x(\sigma))}{\sum_{\rho:|R|>\varepsilon} \exp(\sqrt{s}\beta x(\rho))} \\
&= \frac{(\sqrt{s}\beta \exp(\sqrt{s}\beta x(\sigma))\delta_{\sigma=\tau}) \left(\sum_{\rho} \exp(\sqrt{s}\beta x(\rho)) \right) - (\exp(\sqrt{s}\beta x(\sigma))) (\sqrt{s}\beta \sum_{\tau} \exp(\sqrt{s}\beta x(\tau)))}{\left(\sum_{\rho} \exp(\sqrt{s}\beta x(\rho)) \right)^2} \\
&= \frac{\sqrt{s}\beta \exp(\sqrt{s}\beta x(\sigma)) \left(\delta_{\sigma=\tau} \sum_{\rho} \exp(\sqrt{s}\beta x(\rho)) - \sum_{\tau} \exp(\sqrt{s}\beta x(\tau)) \right)}{\left(\sum_{\rho} \exp(\sqrt{s}\beta x(\rho)) \right) \left(\sum_{\rho} \exp(\sqrt{s}\beta x(\rho)) \right)} \\
&= \sqrt{s}\beta \frac{\exp(\sqrt{s}\beta x(\sigma))}{\sum_{\rho} \exp(\sqrt{s}\beta x(\rho))} \frac{\left(\delta_{\sigma=\tau} \sum_{\rho} \exp(\sqrt{s}\beta x(\rho)) - \sum_{\tau} \exp(\sqrt{s}\beta x(\tau)) \right)}{\sum_{\rho} \exp(\sqrt{s}\beta x(\rho))} \\
&= \sqrt{s}\beta G_s(\sigma) \left(\delta_{\sigma=\tau} - \sum_{\tau} G_s(\tau) \right)
\end{aligned}$$

Now, we can substitute the derivative back into the formula:

$$\begin{aligned}
& \sqrt{s}\beta \sum_{\sigma} \sum_{\tau} \mathbb{E} \left[x(\sigma)x(\tau) \left(G_s(\sigma)\delta_{\sigma=\tau} - G_s(\sigma) \sum_{\tau} G_s(\tau) \right) \right] \\
&= \sqrt{s}\beta \mathbb{E} \left[\sum_{\sigma} x(\sigma)x(\sigma)G_s(\sigma) - \sum_{\sigma} \sum_{\tau} x(\sigma)x(\tau)G_s(\sigma)G_s(\tau) \right] \\
&= \sqrt{s}\beta \mathbb{E} \left[\sum_{\sigma} \langle x(\sigma)^2 \rangle_s - \sum_{\sigma} \sum_{\tau} \langle x(\sigma)x(\tau) \rangle_s \right]
\end{aligned}$$

Then:

$$\sqrt{s}\beta \mathbb{E} \left[\sum_{\sigma^1 \sigma^2} \left\langle \left(H^1(\sigma^1) + H^2(\sigma^2) \right)^2 \right\rangle_s - \sum_{\sigma^1 \sigma^2} \sum_{\tau^1 \tau^2} \left\langle \left(H^1(\sigma^1) + H^2(\sigma^2) \right) \left(H^1(\tau^1) + H^2(\tau^2) \right) \right\rangle_s \right]$$

Now we can deal with these two terms independently:

$$\begin{aligned}
& \sum_{\sigma^1 \sigma^2} \mathbb{E} \left\langle \left(H^1(\sigma^1) + H^2(\sigma^2) \right)^2 \right\rangle_s \\
&= \sum_{\sigma^1 \sigma^2} \mathbb{E} \left\langle H^1(\sigma^1)^2 \right\rangle_s + \sum_{\sigma^1 \sigma^2} \mathbb{E} \left\langle H^2(\sigma^2)^2 \right\rangle_s + 2 \sum_{\sigma^1 \sigma^2} \mathbb{E} \left\langle H^1(\sigma^1) H^2(\sigma^2) \right\rangle_s
\end{aligned}$$

Now we know that the covariance of the Gaussian Hamiltonians is given by:

$$\mathbb{E}H^1(\sigma^1)H^2(\sigma^2) = n\zeta(R_{12})$$

We can also denote the overlap between the i.i.d. replicas coming from the Gibbs measure as:

$$R_{\ell,\ell'}^{j,j'} = \frac{1}{n} \sum_{i=1}^n \sigma_i^{\ell,j} \sigma_i^{\ell',j'} \quad (3.4)$$

So:

$$\begin{aligned} \sum_{\sigma^1} \mathbb{E} \left\langle H^1(\sigma^1) H^1(\sigma^1) \right\rangle_s &= n\zeta(1) \\ \sum_{\sigma^2} \mathbb{E} \left\langle H^2(\sigma^2) H^2(\sigma^2) \right\rangle_s &= n\zeta(1) \\ 2 \sum_{\sigma^1 \sigma^2} \mathbb{E} \left\langle H^1(\sigma^1) H^2(\sigma^2) \right\rangle_s &= 2n\mathbb{E} \langle \zeta(R_{11}^{12}) \rangle_s \end{aligned}$$

As well we can write in terms of the overlaps the second term as follows:

$$\begin{aligned} &\sum_{\sigma^1 \sigma^2} \sum_{\tau^1 \tau^2} \mathbb{E} \left\langle \left(H^1(\sigma^1) + H^2(\sigma^2) \right) \left(H^1(\tau^1) + H^2(\tau^2) \right) \right\rangle_s \\ &= \sum_{\sigma^1 \sigma^2} \sum_{\tau^1 \tau^2} \mathbb{E} \left(\langle H^1(\sigma^1) H^1(\tau^1) \rangle_s + \langle H^1(\sigma^1) H^2(\tau^2) \rangle_s + \langle H^2(\sigma^2) H^1(\tau^1) \rangle_s + \langle H^2(\sigma^2) H^2(\tau^2) \rangle_s \right) \end{aligned}$$

so:

$$\begin{aligned} \sum_{\sigma^1 \sigma^2} \sum_{\tau^1 \tau^2} \mathbb{E} \langle H^1(\sigma^1) H^1(\tau^1) \rangle_s &= n\mathbb{E} \langle \zeta(R_{12}^{11}) \rangle_s \\ \sum_{\sigma^1 \sigma^2} \sum_{\tau^1 \tau^2} \mathbb{E} \langle H^1(\sigma^1) H^2(\tau^2) \rangle_s &= n\mathbb{E} \langle \zeta(tR_{12}^{12}) \rangle_s \\ \sum_{\sigma^1 \sigma^2} \sum_{\tau^1 \tau^2} \mathbb{E} \langle H^2(\sigma^2) H^1(\tau^1) \rangle_s &= n\mathbb{E} \langle \zeta(tR_{12}^{21}) \rangle_s \\ \sum_{\sigma^1 \sigma^2} \sum_{\tau^1 \tau^2} \mathbb{E} \langle H^2(\sigma^2) H^2(\tau^2) \rangle_s &= n\mathbb{E} \langle \zeta(R_{12}^{22}) \rangle_s \end{aligned}$$

Where the term $t \in (0, 1)$ is the correlation parameter between two different Hamiltonians.

Putting everything together, we finally get that:

$$I = \frac{\beta^2}{2} \left(2n\zeta(1) + 2\mathbb{E} \langle \zeta(R_{11}^{12}) \rangle_s - \mathbb{E} \left\langle \zeta(R_{12}^{11}) + \zeta(tR_{12}^{12}) + \zeta(tR_{12}^{21}) + \zeta(R_{12}^{22}) \right\rangle_s \right) \quad (3.5)$$

3.4 Derivative of the Poissonian term

The second term corresponds to the diluted Hamiltonian and it has to be computed by taking the Poisson derivative with respect to s . In order to do that we can use the following lemma:

Lemma. Let $X \sim \text{Po}((1-s)\lambda N)$ and $f : \mathbb{R} \rightarrow \mathbb{R}$ be a function s.t: $\mathbb{E}[|f(x)|] < \infty$, then:

$$\frac{\partial}{\partial s} \mathbb{E}[f(X)] = \mathbb{E}[f(X+1)] - \mathbb{E}[f(X)]$$

Proof. The proof follows from direct calculations

$$\text{Given the Poisson distribution: } P(X = k) = \frac{[\lambda(1-s)N]^k}{k!} e^{-\lambda(1-s)N}$$

$$\text{Then: } \mathbb{E}[f(X)] = \sum_{k=0}^{\infty} f(k) \frac{[\lambda(1-s)N]^k}{k!} e^{-\lambda(1-s)N}$$

Now we can compute the derivative with respect to s of the expectation value:

$$\begin{aligned} \frac{\partial}{\partial s} \mathbb{E}[f(X)] &= \sum_{k=0}^{\infty} \frac{f(k)}{k!} \frac{\partial}{\partial s} \left([\lambda(1-s)N]^k e^{-\lambda(1-s)N} \right) \\ &= \sum_{k=0}^{\infty} \frac{f(k)}{k!} \left(-k[\lambda(1-s)N]^{k-1} (\lambda N) e^{-\lambda(1-s)N} - [\lambda(1-s)N]^k (\lambda N) e^{-\lambda(1-s)N} \right) \\ &= - \sum_{k=1}^{\infty} \frac{f(k)}{(k-1)!} [\lambda(1-s)N]^{k-1} e^{-\lambda(1-s)N} \lambda N - \sum_{k=0}^{\infty} \frac{f(k)}{k!} [\lambda(1-s)N]^k e^{-\lambda(1-s)N} \lambda N \\ &= -\lambda N \left[\sum_{k=1}^{\infty} \frac{f(k)}{(k-1)!} [\lambda(1-s)N]^{k-1} e^{-\lambda(1-s)N} + \sum_{k=0}^{\infty} \frac{f(k)}{k!} [\lambda(1-s)N]^k e^{-\lambda(1-s)N} \right] \end{aligned}$$

then we make the following change of index for the first term: $\ell = k - 1$

$$\begin{aligned} &= -\lambda N \left[\sum_{\ell=0}^{\infty} \frac{f(\ell+1)}{\ell!} [\lambda(1-s)N]^{\ell} e^{-\lambda(1-s)N} + \sum_{k=0}^{\infty} \frac{f(k)}{k!} [\lambda(1-s)N]^k e^{-\lambda(1-s)N} \right] \\ &= -\lambda N (\mathbb{E}[f(X+1)] + \mathbb{E}[f(X)]) \end{aligned}$$

□

Applying the above lemma, we get:

$$\begin{aligned} II &= \frac{\partial \Phi}{\partial v} \frac{dv}{ds} \\ &= \frac{\partial}{\partial v(s)} \frac{1}{n} \mathbb{E} \log \sum_{|R_{12}| > \varepsilon} \exp H(s, \sigma^1, \sigma^2) \\ &= -\lambda \left(\mathbb{E} \log \sum_{|R_{12}| > \varepsilon} \exp H^+(s, \sigma^1, \sigma^2) + \mathbb{E} \log \sum_{|R_{12}| > \varepsilon} \exp H(s, \sigma^1, \sigma^2) \right) \end{aligned}$$

As one can see from the proof of the lemma, computing the derivative of a Poisson random variable with respect to its parameter, corresponds to adding a clause. The

direct application of this lemma to our term results in this full calculation:

$$\begin{aligned}
II &= \frac{1}{n} \frac{\partial}{\partial s} \mathbb{E} \log \sum_{|R_{12}| > \varepsilon} \exp \left(\sum_{\ell=1}^2 \delta H_{\lambda(1-s)}^{\ell}(\sigma^{\ell}) \right) \\
&= \frac{1}{n} \frac{\partial}{\partial s} \mathbb{E} \log \sum_{|R_{12}| > \varepsilon} \exp \left(\delta \left(H_{\lambda(1-s)}^1(\sigma^1) + H_{\lambda(1-s)}^2(\sigma^2) \right) \right) \\
&= \frac{1}{n} \frac{\partial}{\partial s} \mathbb{E} \log \sum_{|R_{12}| > \varepsilon} \exp \left(\delta \sum_{j=1}^{\lambda \pi(1-s)n} \left(\theta_j^1(\sigma^1) + \theta_j^2(\sigma^2) \right) \right) \\
&= -\lambda \mathbb{E} \log \sum_{|R_{12}| > \varepsilon} \exp \left(\delta \sum_{j=1}^{(\lambda \pi(1-s)n)+1} \left(\theta_j^1(\sigma^1) + \theta_j^2(\sigma^2) \right) \right) + \\
&\quad + \lambda \mathbb{E} \log \sum_{|R_{12}| > \varepsilon} \exp \left(\delta \sum_{j=1}^{\lambda \pi(1-s)n} \left(\theta_j^1(\sigma^1) + \theta_j^2(\sigma^2) \right) \right) \\
&= -\lambda \left(\mathbb{E} \log \sum_{|R_{12}| > \varepsilon} \exp \left(\delta \sum_{j=1}^{(\lambda \pi(1-s)n)+1} \left(\theta_j^1(\sigma^1) + \theta_j^2(\sigma^2) \right) \right) \right. \\
&\quad \left. - \mathbb{E} \log \sum_{|R_{12}| > \varepsilon} \exp \left(\delta \sum_{j=1}^{\lambda \pi(1-s)n} \left(\theta_j^1(\sigma^1) + \theta_j^2(\sigma^2) \right) \right) \right) \\
&= -\lambda \left(\mathbb{E} \log \sum_{|R_{12}| > \varepsilon} \exp H^+(s, \sigma^1, \sigma^2) - \mathbb{E} \log \sum_{|R_{12}| > \varepsilon} \exp H(s, \sigma^1, \sigma^2) \right)
\end{aligned}$$

Where

$$H^+(s, \sigma^1, \sigma^2) = H(s, \sigma^1, \sigma^2) + \delta \theta^1(\sigma_{i_1}^1 \cdots \sigma_{i_k}^1) + \delta \theta^2(\sigma_{i_1}^2 \cdots \sigma_{i_k}^2), \quad (3.6)$$

and these clauses are independent of $H(s, \sigma^1, \sigma^2)$ and are given by:

$$\theta^{\ell}(\sigma_{i_1}^{\ell}, \dots, \sigma_{i_k}^{\ell}) = - \prod_{j \leq k} \frac{1 + J_j^{\ell} \sigma_j^{\ell}}{2},$$

where the random signs J_j^{ℓ} are correlated in the sense that they are resampled with probability $1 - t$ either independently or simultaneously within this one clause. Now, notice that the term II can be rewritten as follows:

$$\begin{aligned}
&= -\lambda \left(\mathbb{E} \log \sum_{|R_{12}| > \varepsilon} \exp H(s, \sigma^1, \sigma^2) \exp(\delta \theta^1) \exp(\delta \theta^2) - \mathbb{E} \log Z_s \right) \\
&= -\lambda \mathbb{E} \log \frac{\sum_{|R_{12}| > \varepsilon} H(s, \sigma^1, \sigma^2) \exp \delta \theta^1 \exp \delta \theta^2}{Z_s} \\
&= -\lambda \mathbb{E} \log \left\langle \exp \delta \theta^1(\sigma_{i_1}^1, \dots, \sigma_{i_k}^1) \exp \delta \theta^2(\sigma_{i_1}^2, \dots, \sigma_{i_k}^2) \right\rangle_s
\end{aligned}$$

Now, since each clause $\theta \in \{-1, 0\}$ we can rewrite each exponential term as:

$$\exp \delta \theta = 1 + (1 - e^{-\delta}) \theta$$

then:

$$\begin{aligned}
\exp \delta \theta^1 \exp \delta \theta^2 &= \left(1 + (1 - e^\delta) \theta^1\right) \left(1 + (1 - e^\delta) \theta^2\right) \\
&= 1 + (1 - e^\delta) \theta^2 + (1 - e^\delta) \theta^1 + (1 - e^\delta)^2 \theta^1 \theta^2 \\
&= 1 - (1 - e^\delta) \left(\prod_{j \leq k} \frac{1 + J_j^2 \sigma_{i_j}^2}{2} \right) - (1 - e^\delta) \left(\prod_{j \leq k} \frac{1 + J_j^1 \sigma_{i_j}^1}{2} \right) + (1 - e^\delta)^2 \prod_{j \leq k} \frac{1 + J_j^1 \sigma_{i_j}^1}{2} \frac{1 + J_j^2 \sigma_{i_j}^2}{2} \\
&= 1 - (1 - e^\delta) \left(\prod_{j \leq k} \frac{1 + J_j^1 \sigma_{i_j}^1}{2} + \prod_{j \leq k} \frac{1 + J_j^2 \sigma_{i_j}^2}{2} - (1 - e^\delta) \prod_{j \leq k} \frac{1 + J_j^1 \sigma_{i_j}^1}{2} \frac{1 + J_j^2 \sigma_{i_j}^2}{2} \right) \\
&= 1 - (1 - e^\delta) \Delta(\sigma^1, \sigma^2)
\end{aligned}$$

Notice that:

$$0 \leq \Delta(\sigma^1, \sigma^2) \leq 1 + e^{-\delta}$$

$$\text{that also means: } \Delta(\sigma^1, \sigma^2)(1 - e^{-\delta}) \leq 1 - e^{-2\delta}$$

$$\text{leading to the fact that: } \Delta(\sigma^1, \sigma^2) < 1$$

Knowing these bounds, we can safely expand the logarithm using the Taylor series as:

$$II = \lambda \sum_{n \geq 1} \frac{(1 - e^{-\delta})^n}{n} \mathbb{E} \langle \Delta(\sigma^1, \sigma^2) \rangle_s^n + III \quad (3.7)$$

where we define the remainder

$$III := \lambda \sum_{n \geq 3} \frac{(1 - e^{-\delta})^n}{n} \mathbb{E} \langle \Delta(\sigma^1, \sigma^2) \rangle_s^n.$$

Since $0 \leq \Delta(\sigma^1, \sigma^2) < 1$, we have $|\mathbb{E} \langle \Delta \rangle_s^n| \leq 1$ for all $n \geq 1$, hence

$$|III| \leq \lambda \sum_{n \geq 3} \frac{(1 - e^{-\delta})^n}{n} \leq \lambda \sum_{n \geq 3} (1 - e^{-\delta})^n = \lambda \frac{(1 - e^{-\delta})^3}{e^{-\delta}} = \mathcal{O}(\lambda \delta^3),$$

uniformly in s . Using replicas we can represent the last term as follows:

$$\mathbb{E} \langle \Delta(\sigma^1, \sigma^2) \rangle_s^n = \mathbb{E} \left\langle \prod_{\ell \leq n} \Delta(\sigma^{\ell,1}, \sigma^{\ell,2}) \right\rangle_s = \mathbb{E} \left\langle \mathbb{E}' \prod_{\ell \leq n} \Delta(\sigma^{\ell,1}, \sigma^{\ell,2}) \right\rangle_s$$

Where $\mathbb{E}'$ is the expectation with respect to the randomness J_k^1, J_k^2 and i_k of the clauses θ^1 and θ^2 , which is independent of the randomness in $\langle \cdot \rangle_s$. So we can compute the expectation with respect to the randomness knowing that the correlations between disorders J_j^1, J_j^2 is given as: $\mathbb{E} J_j^1, J_j^2 = t$, so:

$$\mathbb{E}_J \prod_{j \leq k} \frac{1 + J_j^1 \sigma_{i_j}^1}{2} \frac{1 + J_j^2 \sigma_{i_j}^2}{2} = \prod_{j \leq k} \frac{1 + \mathbb{E}_J J_j^2 \sigma_{i_j}^2 + \mathbb{E}_J J_j^1 \sigma_{i_j}^1 + \mathbb{E}_J J_j^1 J_j^2 \sigma_{i_j}^1 \sigma_{i_j}^2}{4} = \prod_{j \leq k} \frac{1 + t \sigma_{i_j}^1 \sigma_{i_j}^2}{4}$$

while taking the expectation with respect to the random indices i_k , we get:

$$\mathbb{E}' \prod_{j \leq k} \frac{1 + J_j^1 \sigma_{i_j}^1}{2} \frac{1 + J_j^2 \sigma_{i_j}^2}{2} = \prod_{j \leq k} \frac{1 + J_j^2 \mathbb{E}' \sigma_{i_j}^2 + J_j^1 \mathbb{E}' \sigma_{i_j}^1 + J_j^1 J_j^2 \mathbb{E}' \sigma_{i_j}^1 \sigma_{i_j}^2}{4} = \frac{1 + \xi(tR_{1,2})}{4^k}$$

Now we can compute $\mathbb{E} \langle \Delta(\sigma^1, \sigma^2) \rangle_s^n$ for $n = 1, 2$. First, for $n = 1$:

$$\begin{aligned} \mathbb{E} \langle \Delta(\sigma^1, \sigma^2) \rangle_s &= \mathbb{E} \left\langle \prod_{j \leq k} \frac{1 + J_j^1 \sigma_{i_j}^1}{2} + \prod_{j \leq k} \frac{1 + J_j^2 \sigma_{i_j}^2}{2} - (1 - e^{-\delta}) \prod_{j \leq k} \frac{1 + J_j^1 \sigma_{i_j}^1}{2} \frac{1 + J_j^2 \sigma_{i_j}^2}{2} \right\rangle_s \\ &= \frac{2}{2^k} - \frac{1 - e^{-\delta}}{4^k} (1 + \mathbb{E} \langle (tR_{11}^{12}) \rangle_s) \end{aligned}$$

while, for $n = 2$ (we will separate the terms that contain the factor $(1 - e^{-\delta})$):

$$\mathbb{E} \langle \Delta(\sigma^1, \sigma^2) \rangle_s^2 = \mathbb{E} \left[\left(\sum_{\sigma^1 \sigma^2} G_s(\sigma^1, \sigma^2) \Delta(\sigma^1, \sigma^2) \right)^2 \right]$$

here we need to introduce two separate copies $(\sigma^1 \sigma^2), (\tau^1 \tau^2)$ such that:

$$G_s^\otimes \left((\sigma^1 \sigma^2) (\tau^1 \tau^2) \right) = G_s(\sigma^1 \sigma^2) G_s(\tau^1 \tau^2)$$

then:

$$\begin{aligned} &= \sum_{\sigma^1 \sigma^2} G_s(\sigma^1 \sigma^2) \Delta(\sigma^1 \sigma^2) \sum_{\tau^1 \tau^2} G_s(\tau^1 \tau^2) \Delta(\tau^1 \tau^2) \\ &= \langle \Delta(\sigma^1 \sigma^2) \Delta(\tau^1 \tau^2) \rangle_s \end{aligned}$$

For the sake of clarity and simplicity let us change notation for this computation. Recall that:

$$\Delta(\sigma^1 \sigma^2) = \prod_{j \leq k} \frac{1 + J_j^1 \sigma_{i_j}^1}{2} + \prod_{j \leq k} \frac{1 + J_j^2 \sigma_{i_j}^2}{2} - (1 - e^{-\delta}) \prod_{j \leq k} \frac{1 + J_j^1 \sigma_{i_j}^1}{2} \frac{1 + J_j^2 \sigma_{i_j}^2}{2}$$

Then let us define:

$$\Delta(\sigma^1 \sigma^2) := A(\sigma^1) + B(\sigma^2) + cA(\sigma^1)B(\sigma^2)$$

$$\Delta(\tau^1 \tau^2) := A(\tau^1) + B(\tau^2) + cA(\tau^1)B(\tau^2)$$

with: $c = 1 - e^{-\delta}$

so:

$$\begin{aligned} &\mathbb{E} \langle \Delta(\sigma^1 \sigma^2) \Delta(\tau^1 \tau^2) \rangle_s \\ &= \mathbb{E} \left\langle \left(A(\sigma^1) + B(\sigma^2) + cA(\sigma^1)B(\sigma^2) \right) \cdot \left(A(\tau^1) + B(\tau^2) + cA(\tau^1)B(\tau^2) \right) \right\rangle_s \\ &= \mathbb{E} \langle A(\sigma^1) A(\tau^1) \rangle_s + \mathbb{E} \langle A(\sigma^1) B(\tau^2) \rangle_s + \mathbb{E} \langle B(\sigma^2) A(\tau^1) \rangle_s + \mathbb{E} \langle B(\sigma^2) B(\tau^2) \rangle_s + III \\ &= \mathbb{E} \left\langle \frac{1 + \xi(R_{12}^{11})}{4^k} + \frac{1 + \xi(tR_{12}^{12})}{4^k} + \frac{1 + \xi(tR_{12}^{21})}{4^k} + \frac{1 + \xi(R_{12}^{22})}{4^k} \right\rangle_s + III \end{aligned}$$

Now we can insert everything inside II :

$$II = \frac{2\lambda(1 - e^{-\delta})}{2^k} - \frac{\lambda(1 - e^{-\delta})^2}{4^k} (1 + \mathbb{E}\langle \xi(tR_{11}^{12}) \rangle_s) \\ + \frac{4\lambda(1 - e^{-\delta})^2}{2 \cdot 4^k} + \frac{\lambda(1 - e^{-\delta})^2}{2 \cdot 4^k} \mathbb{E}\langle \xi(R_{12}^{11}) + \xi(tR_{12}^{12}) + \xi(tR_{12}^{21}) + \xi(R_{12}^{22}) \rangle_s + III$$

Then, we make the following choice of β :

$$\beta = \frac{\sqrt{\lambda}(1 - e^{-\delta})}{2^k}$$

With this choice of β we finally get that the derivative of the free energy is:

$$\varphi'(s) = \frac{2\lambda(1 - e^{-\delta})}{2^k} + \frac{2\lambda(1 - e^{-\delta})^2}{2 \cdot 2^k} + III$$

And by performing a Taylor expansion around $\delta \rightarrow 0$:

$$\varphi'(s) = \frac{2\lambda\delta}{2^k} + \mathcal{O}(\lambda\delta^3)$$

Then, integrating between 0 and 1, with respect to s :

$$\int_0^1 \varphi'(s) ds = \varphi(1) - \varphi(0) = \frac{2\lambda\delta}{2^k} + \mathcal{O}(\lambda\delta^3)$$

So:

$$\left| \varphi(0) + \frac{2\lambda\delta}{2^k} - \varphi(1) \right| = \mathcal{O}(\lambda\delta^3) \\ \left| \frac{1}{\delta} \varphi(0) + \frac{2\lambda}{2^k} - \frac{1}{\delta} \varphi(1) \right| = \mathcal{O}(\lambda\delta^2)$$

3.5 Bounds for the energy

Now we want to estimate the upper-bound and the lower-bound for the expression of the free energy.

To do so, let us first take the exponential term, and notice that we can write the

following estimate to get the lower-bound for the free energy:

$$\begin{aligned}
\max_{\sigma^1 \sigma^2} \exp H(s, \sigma^1, \sigma^2) &= \exp H(s, \sigma^{1*}, \sigma^{2*}) \leq \sum_{|R_{12}| > \varepsilon} \exp H(s, \sigma^1, \sigma^2) \\
\log \exp H(s, \sigma^{1*}, \sigma^{2*}) &\leq \log \sum_{|R_{12}| > \varepsilon} \exp H(s, \sigma^1, \sigma^2) \\
\max_{\sigma^1 \sigma^2} H(s, \sigma^1, \sigma^2) &\leq \log \sum_{|R_{12}| > \varepsilon} \exp H(s, \sigma^1, \sigma^2) \\
\frac{1}{n} \mathbb{E} \max_{\sigma^1 \sigma^2} H(s, \sigma^1, \sigma^2) &\leq \frac{1}{n} \mathbb{E} \log \sum_{|R_{12}| > \varepsilon} \exp H(s, \sigma^1, \sigma^2) \\
\frac{1}{n} \mathbb{E} \max_{\sigma^1 \sigma^2} H(s, \sigma^1, \sigma^2) &\leq \varphi(s)
\end{aligned}$$

Moreover:

$$\begin{aligned}
\exp H(s, \sigma^1, \sigma^2) &\leq \max_{\sigma^1 \sigma^2} \exp H(s, \sigma^1, \sigma^2) \\
\sum_{|R_{12}| > \varepsilon} \exp H(s, \sigma^1, \sigma^2) &\leq \sum_{|R_{12}| > \varepsilon: \{(\sigma^1, \sigma^2) \in \max\}} \exp H(s, \sigma^1, \sigma^2) \\
\sum_{|R_{12}| > \varepsilon} \exp H(s, \sigma^1, \sigma^2) &\leq 2^{2n} \max_{\sigma^1 \sigma^2} \exp H(s, \sigma^1, \sigma^2) \\
\log \sum_{|R_{12}| > \varepsilon} \exp H(s, \sigma^1, \sigma^2) &\leq 2n \log 2 + \max_{\sigma^1 \sigma^2} H(s, \sigma^1, \sigma^2) \\
\varphi(s) &\leq \frac{1}{n} \mathbb{E} \max_{\sigma^1 \sigma^2} H(s, \sigma^1, \sigma^2) + 2 \log 2
\end{aligned}$$

Finally we see that the free energy is bounded as:

$$\begin{aligned}
\frac{1}{n} \mathbb{E} \max_{\sigma^1 \sigma^2} H(s, \sigma^1, \sigma^2) &\leq \varphi(s) \leq \frac{1}{n} \mathbb{E} \max_{\sigma^1 \sigma^2} H(s, \sigma^1, \sigma^2) + 2 \log 2 \\
\left| \frac{1}{n} \mathbb{E} \max_{|R_{12}| > \varepsilon} H(s, \sigma^1, \sigma^2) - \varphi(s) \right| &\leq 2 \log 2
\end{aligned}$$

Then we can evaluate this bound for $s = 0$ and $s = 1$:

$$\begin{aligned}
s = 0 \rightarrow H(0, \sigma^1, \sigma^2) &= \delta \left(H_\lambda^1(\sigma^1) + H_\lambda^2(\sigma^2) \right) \\
\left| \frac{1}{n} \mathbb{E} \max_{|R_{12}| > \varepsilon} \left(H_\lambda^1(\sigma^1) + H_\lambda^2(\sigma^2) \right) - \frac{1}{\delta} \varphi(0) \right| &\leq \frac{2 \log 2}{\delta}
\end{aligned}$$

$$\begin{aligned}
s = 1 \rightarrow H(1, \sigma^1, \sigma^2) &= \beta \left(H^1(\sigma^1) + H^2(\sigma^2) \right) \\
\left| \frac{\beta}{\delta} \frac{1}{n} \mathbb{E} \max_{|R_{12}| > \varepsilon} \left(H^1(\sigma^1) + H^2(\sigma^2) \right) - \frac{1}{\delta} \varphi(1) \right| &\leq \frac{2 \log 2}{\delta}
\end{aligned}$$

Then from these two bounds we find that:

$$\frac{1}{\delta}\varphi(0) = \frac{1}{n}\mathbb{E} \max_{|R_{12}|>\varepsilon} \left(H_{\lambda}^1(\sigma^1) + H_{\lambda}^2(\sigma^2) \right) + \mathcal{O}\left(\frac{1}{\delta}\right)$$

and

$$\frac{1}{\delta}\varphi(1) = \frac{\beta}{\delta} \frac{1}{n}\mathbb{E} \max_{|R_{12}|>\varepsilon} \left(H^1(\sigma^1) + H^2(\sigma^2) \right) + \mathcal{O}\left(\frac{1}{\delta}\right)$$

Now, from the result of the previous integration, we know that:

$$\frac{1}{\delta}\varphi(0) = \frac{1}{\delta}\varphi(1) - \frac{2\lambda}{2^k} + \mathcal{O}(\lambda\delta^2) \quad (3.8)$$

So:

$$\frac{1}{n}\mathbb{E} \max_{|R_{12}|>\varepsilon} \left(H_{\lambda}^1(\sigma^1) + H_{\lambda}^2(\sigma^2) \right) = \frac{\beta}{\delta} \frac{1}{n}\mathbb{E} \max_{|R_{12}|>\varepsilon} \left(H^1(\sigma^1) + H^2(\sigma^2) \right) - \frac{2\lambda}{2^k} + \mathcal{O}\left(\lambda\delta^2 + \frac{1}{\delta}\right) \quad (3.9)$$

Now we can perform a Taylor expansion on our choice of β to get the following:

$$\beta = \frac{\sqrt{\lambda}}{2^k} + \mathcal{O}\left(\sqrt{\lambda\delta^2}\right) \quad (3.10)$$

And by substituting it into the previous expression we get:

$$\frac{1}{n}\mathbb{E} \max_{|R_{12}|>\varepsilon} \left(H_{\lambda}^1(\sigma^1) + H_{\lambda}^2(\sigma^2) \right) = \frac{\sqrt{\lambda}}{2^k} \frac{1}{n}\mathbb{E} \max_{|R_{12}|>\varepsilon} \left(H^1(\sigma^1) + H^2(\sigma^2) \right) - \frac{2\lambda}{2^k} + \mathcal{O}\left(\lambda\delta^2 + \frac{1}{\delta} + \sqrt{\lambda\delta^2}\right) \quad (3.11)$$

With the choice of $\delta = \lambda^{-\frac{1}{3}}$, the error term is $\mathcal{O}(\lambda^{\frac{1}{3}})$.

Now using the following:

Theorem. [3, Theorem 2] for any $\varepsilon, t \in (0, 1)$, there exists $\eta > 0$ such that, for large enough n :

$$\frac{1}{n}\mathbb{E} \max_{|R_{12}|>\varepsilon} \left(H^1(\sigma^1) + H^2(\sigma^2) \right) \leq \frac{2}{n}\mathbb{E} \max_{\sigma} H(\sigma) - \eta \quad (3.12)$$

We can write the following inequality:

$$\frac{1}{n}\mathbb{E} \max_{|R_{12}|>\varepsilon} \left(H_{\lambda}^1(\sigma^1) + H_{\lambda}^2(\sigma^2) \right) \leq -\frac{2\lambda}{2^k} + \frac{2\sqrt{\lambda}}{2^k}\mathbb{E} \max_{\sigma} H(\sigma) - \frac{2\sqrt{\lambda}\eta}{2^k} + L\lambda^{\frac{1}{3}} \quad (3.13)$$

Moreover, we can write the latter expression as a function of H_{λ} by using the approximation:

$$\frac{1}{n}\mathbb{E} \max_{\sigma} H_{\lambda}(\sigma) = -\frac{\lambda}{2^k} + \frac{\sqrt{\lambda}}{2^k} \frac{1}{n}\mathbb{E} \max_{\sigma \in V} H(\sigma) + \mathcal{O}(\lambda^{\frac{1}{3}}), \text{ as } \lambda \rightarrow \infty$$

And we get:

$$\frac{1}{n} \mathbb{E} \max_{|R_{12}| > \varepsilon} \left(H_\lambda^1(\sigma^1) + H_\lambda^2(\sigma^2) \right) \leq \frac{2}{n} \mathbb{E} \max_{\sigma} H_\lambda(\sigma) - \frac{2\sqrt{\lambda}\eta}{2^k} + L\lambda^{\frac{1}{3}} \quad (3.14)$$

Now, for $\lambda \geq L\eta^{-6}$ for large enough constant $L = L(k)$:

$$\frac{1}{n} \mathbb{E} \max_{|R_{12}| > \varepsilon} \left(H_\lambda^1(\sigma^1) + H_\lambda^2(\sigma^2) \right) \leq \frac{2}{n} \mathbb{E} \max_{\sigma} H_\lambda(\sigma) - \frac{\sqrt{\lambda}\eta}{L} \quad (3.15)$$

Finally, by Azuma's inequality:

$$\frac{1}{n} \max_{|R_{12}| > \varepsilon} \left(H_\lambda^1(\sigma^1) + H_\lambda^2(\sigma^2) \right) \leq \frac{1}{n} \left(\max_{\sigma} H_\lambda^1(\sigma^1) + \max_{\sigma} H_\lambda^2(\sigma^2) \right) - \frac{\sqrt{\lambda}\eta}{L}$$

with probability at least: $1 - Le^{-\frac{n\eta^2}{L}}$

On this event, the existence of σ^1, σ^2 such that:

$$\frac{1}{n} H_\lambda^\ell(\sigma^\ell) \geq \frac{1}{n} \max_{\sigma} H_\lambda^\ell(\sigma^\ell) - \frac{\sqrt{\lambda}\eta}{3L} \quad (3.16)$$

and such that $|R_{12}| > \varepsilon$ would lead to contradiction.

Chapter 4

Other Models

In this chapter we extend the discussion of disorder chaos from diluted random k -SAT to three closely related CSPs: random k -NAESAT, random hypergraph 2-coloring problem and random k -XORSAT.

For each model, we will follow the same strategy adopted to show the evidence of disorder chaos in diluted random k -SAT. So each model will be reformulated as a spin system, by expressing the instance as a Hamiltonian on $\{-1; +1\}^n$. The contribution of each clause, or hyperedge, will be expanded to identify the effective order of interactions. Then, the high-connectivity regime will be discussed, first letting $n \rightarrow \infty$ at fixed connectivity λ , then $\lambda \rightarrow \infty$, in which the diluted Hamiltonian can be approximated by a fully connected mean-field model.

Finally, we show that, in this regime, the resulting Gaussian Hamiltonian has a covariance that depends only on the overlap $R(\sigma^1, \sigma^2)$, which is the key structural property underlying the analysis of mixed p -spin models and the rigorous theory of disorder chaos.

As well as in Chapter 3, also in this chapter it is crucial to stress that the fully connected approximation that will be discussed for each model, is heuristic. The convergence of the diluted Hamiltonian to a Gaussian process cannot be justified by a direct application of the canonical central limit theorem. In the diluted models, each spin variable appears in a random number of clauses, and the corresponding interaction terms are neither identically distributed nor independent, they share indices and exhibit complex correlations, as a result, the standard CLT assumptions: independence and uniform variance scaling, are violated. In this thesis, we rely on the heuristic observation that in the high-connectivity regime the clause contributions become weakly dependent, and their cumulative effect is approximately Gaussian. This motivates the use of a fully connected mixed p -spin model as an effective description of the diluted systems, allowing us to discuss, at a qualitative level, the emergence of disorder chaos across these different random CSPs.

4.1 Disorder Chaos in random k -NAESAT

As defined in Section 1.3.1, each clause in the random k -NAESAT model forbids all k literals from taking the same value, in contrast to the k -SAT where clauses are penalized only when none of the variables are satisfied. We will see that this symmetry under global spin flip leads to the cancellation of all odd-interaction terms in the expansion of the Hamiltonian.

In this section we reformulate the model as a spin system and develop its fully connected approximation, which will later allow us to study its covariance structure and the emergence of disorder chaos.

4.1.1 The k -NAESAT Hamiltonian

To study the problem using tools from statistical physics we reformulate the problem in terms of a spin system.

Each boolean variable x_i is mapped to a spin variable σ_i .

$$x_i = 1 \longleftrightarrow \sigma_i = +1 \quad \text{and} \quad x_i = 0 \longleftrightarrow \sigma_i = -1$$

As briefly described above, each clause in the k -NAESAT problem imposes a constraint: not all literals in the clause should have the same truth value. In spin variables this means that the clause is violated if all σ_i s are equal, i.e., either all $+1$ or -1 . We define an energy penalty for such violating configurations, the contribution of each clause can be defined as:

$$\theta(\sigma_1, \dots, \sigma_k) := -\prod_{j=1}^k \frac{1 + J_j \sigma_j}{2} - \prod_{j=1}^k \frac{1 - J_j \sigma_j}{2}, \quad (\sigma_1, \dots, \sigma_k) \in \{-1, 1\}^k.$$

This expression equals -1 if all literals are simultaneously true or false, and 0 otherwise, therefore the full Hamiltonian is the sum over all such clause contributions:

$$\begin{aligned} H_\lambda(\sigma) &= \sum_{c \leq \pi(\lambda n)} \left[\left(-\prod_{j \leq k} \frac{1 + J_{j,c} \sigma_{i_{j,c}}}{2} \right) - \left(\prod_{j \leq k} \frac{1 - J_{j,c} \sigma_{i_{j,c}}}{2} \right) \right] \\ &= - \sum_{c \leq \pi(\lambda n)} \left[\left(-\prod_{j \leq k} \frac{1 + J_{j,c} \sigma_{i_{j,c}}}{2} \right) + \left(\prod_{j \leq k} \frac{1 - J_{j,c} \sigma_{i_{j,c}}}{2} \right) \right]. \end{aligned}$$

This defines a spin-glass-type energy landscape for the k -NAESAT model.

4.1.2 The k -NAESAT fully connected approximation

The Hamiltonian of the k -NAESAT model can be approximated, in a very similar way to the k -SAT Hamiltonian, by fully connected model in the large connectivity regime.

Proof. Let us consider the contribution of a single clause:

$$\theta(\sigma_1, \dots, \sigma_k) = -\prod_{j=1}^k \frac{1 + J_j \sigma_j}{2} - \prod_{j=1}^k \frac{1 - J_j \sigma_j}{2}.$$

We can try to expand the single factors in one clause as:

$$\begin{aligned} \prod_{j=1}^k (1 + J_j \sigma_j) &= \sum_{S \subseteq [k]} \prod_{j \in S} J_j \sigma_j, \\ \prod_{j=1}^k (1 - J_j \sigma_j) &= \sum_{S \subseteq [k]} (-1)^{|S|} \prod_{j \in S} J_j \sigma_j. \end{aligned}$$

Recalling that $|S|$ is the cardinality of the number of interactions between spins, this means that in each clause, only spins with an even number of interactions will contribute, and we obtain:

$$\theta = -2 \sum_{\substack{S \subseteq k \\ |S| \text{ even}}} \prod_{j \in S} J_j \sigma_j$$

So the Hamiltonian takes the form:

$$H_\lambda = -\frac{2}{2^k} \sum_{c \leq \pi(\lambda n)} \sum_{\substack{S \subseteq k \\ |S| \text{ even}}} \prod_{j \in S} J_{j,c} \sigma_{i_{j,c}},$$

grouping terms by their interaction order $p = |S|$, the Hamiltonian becomes

$$H_\lambda = -\frac{2}{2^k} \sum_{\substack{p=2 \\ |S|=p \\ |S| \text{ even}}}^k \sum_{\substack{S \subseteq k \\ |S|=p \\ |S| \text{ even}}} \left(\sum_{c \leq \pi(\lambda n)} \prod_{j \in S} J_{j,c} \sigma_{i_{j,c}} \right). \quad (4.1)$$

This expression has the same structure as the expansion of the diluted k -SAT model, except for the fact that in this case only terms that enter the Hamiltonian with an even number of interactions, contribute to the energy. In the large connectivity limit ($\lambda \rightarrow \infty$), each coefficient in front of the product $\sigma_{i_1} \dots \sigma_{i_p}$ is the sum of many **weakly** dependent random contributions. After centering and normalizing so that $\text{Var}(H) \sim n$, these coefficients become approximately Gaussian by the central limit theorem.

It is important to remark again that, in the high connectivity regime, the clause coefficients can be viewed as approximately Gaussian by a central limit-type argument. The rigorous justification of such a limit typically relies on the Lindeberg interpolation method [8] which we do not reproduce here. Hence, the k -NAESAT Hamiltonian can be approximated by the following fully connected Gaussian model:

$$H(\sigma) = \sum_{\substack{p=2 \\ p \text{ even}}}^k \sqrt{\binom{k}{p} \frac{1}{n^{p-1}}} \sum_{1 \leq i_1, \dots, i_p \leq n} g_{i_1, \dots, i_p} \sigma_{i_1} \dots \sigma_{i_p},$$

where coefficients $g_{i_1, \dots, i_p}$ are independent standard random variables. □

4.1.3 Covariance computation

Let us verify whether in this case as well the covariance depends only on the overlap

$$\begin{aligned}
\mathbb{E}H(\sigma^1)H(\sigma^2) &= \sum_{\substack{p=2 \\ p \text{ even}}}^k \binom{k}{p} \frac{1}{n^{p-1}} \sum_{i_1, \dots, i_p} \sigma_{i_1}^1 \cdots \sigma_{i_p}^1 \sigma_{i_1}^2 \cdots \sigma_{i_p}^2 \\
&= \sum_{\substack{p=2 \\ p \text{ even}}}^k \binom{k}{p} \frac{1}{n^{p-1}} n^p R_{1,2}^p \\
&\text{or} \\
&= \sum_{p=1}^{k/2} \binom{k}{2p} \frac{1}{n^{2p-1}} n^{2p} R_{1,2}^{2p}
\end{aligned}$$

We can see that also in this case the covariance depends only on the overlap. This allows us to extend the previous proof on the existence of disorder chaos also to the k -NAESAT model.

In conclusion, the Hamiltonian of the k -NAESAT model in the fully connected limit retains the essential Gaussian properties, with a covariance that depends only on the overlap between configurations.

In the fully connected limit, the k -NAESAT Hamiltonian belongs to the class of **even** p -spin models, characterized by a covariance that is an even polynomial of the overlap. Consequently, small independent perturbations of the disorder are expected to cause rapid decorrelation between equilibrium configurations: an explicit manifestation of disorder chaos analogous to that observed in even p -spin systems. This establishes that, in the high connectivity regime, the k -NAESAT model exhibits the same qualitative mechanism of chaos as its fully connected counterpart.

4.2 Disorder Chaos in hypergraph 2-coloring

Let us now turn to the Hypergraph 2-coloring model which, as introduced in Section 1.3.2, is defined on a k -uniform hypergraph where each hyperedge connects k vertices chosen uniformly at random. The goal is to assign one of two colors (or equivalently, spin values $\sigma_i \in -1, +1$) to each vertex so that no hyperedge is monochromatic, meaning that not all vertices within it share the same color.

Importantly, this model is closely related to the random k -NAESAT model: both forbid configurations in which all variables in a clause, or hyperedge, take the same value, or color. However, in contrast to k -NAESAT, the hypergraph 2-coloring problem contains no random signs J_j , and hence its disorder arises only from the random choice of hyperedges. In what follows, we reformulate the model as a spin glass Hamiltonian and derive its fully connected approximation, which will again exhibit a covariance depending only on the overlap.

4.3 The Hypergraph 2-col Hamiltonian

Each hyperedge $e_j \subseteq 1, \dots, n$ contributes an energy penalty if all spins within it are equal. The contribution of a single hyperedge can therefore be written as

$$\theta_{e_j}(\sigma) = - \left(\prod_{i \in e_j} \frac{1 + \sigma_i}{2} + \prod_{i \in e_j} \frac{1 - \sigma_i}{2} \right)$$

where $e_j \subset [n]$ is a hyperedge of cardinality k .

This function evaluates to 1 if all spins in the clause are equal (i.e., the hyperedge is monochromatic), and to 0 otherwise. Given a k -uniform random hypergraph with a Poisson-distributed number of edges $\pi(\lambda n)$, the total Hamiltonian is then:

$$H_\lambda(\sigma) = - \sum_{j=1}^{\pi(\lambda n)} \left(\prod_{i \in e_j} \frac{1 + \sigma_i}{2} + \prod_{i \in e_j} \frac{1 - \sigma_i}{2} \right),$$

where each hyperedge e_j is chosen uniformly at random from the set of all k -tuples of vertices. This Hamiltonian counts the total number of violated constraints and therefore defines the energy landscape of the 2-coloring problem.

4.3.1 The Hypergraph 2-coloring fully connected approximation

Recalling that a single clause is:

$$\theta_{e_j}(\sigma) = - \left(\prod_{i \in e_j} \frac{1 + \sigma_i}{2} + \prod_{i \in e_j} \frac{1 - \sigma_i}{2} \right)$$

And also remarking that each clause is a hyper-edge $e_j = \{1, \dots, i_k\} \subseteq n$ chosen uniformly at random. So we have e_j hyper-edges, formed by choosing k vertices from n .

We can expand each term of a clause as:

$$\begin{aligned} \prod_{i \in e_j} \frac{1 + \sigma_i}{2} &= \frac{1}{2^k} \sum_{S \subseteq e_j} \prod_{i \in S} \sigma_i, \\ \prod_{i \in e_j} \frac{1 - \sigma_i}{2} &= \frac{1}{2^k} \sum_{S \subseteq e_j} (-1)^{|S|} \prod_{i \in S} \sigma_i, \end{aligned}$$

adding these expressions cancels all terms of odd order, so only even-order interactions contribute. The Hamiltonian can thus be written as:

$$H_\lambda(\sigma) = \frac{1}{2^{k-1}} \sum_{j=1}^{\pi(\lambda n)} \sum_{\substack{S \subseteq e_j \\ |S| \text{ even}}} \prod_{i \in S} \sigma_i.$$

Now, by fixing a subset $S \subseteq n, |S| = p \leq k, p$ even, and grouping terms with the same interaction order, the Hamiltonian can be redefined as follows:

$$H_\lambda(\sigma) = \sum_{\substack{S \subseteq n \\ |S| \text{ even} \leq k}} c_S \prod_{i \in S} \sigma_i,$$

where

$$c_S := \sum_j^{\pi(\lambda n)} \frac{1}{2^{k-1}} \mathbb{I}(S \subseteq e_j). \quad (4.2)$$

Now we have $\pi(\lambda n)$ clauses, and each e_j clause is independent. We can take the $\lambda \rightarrow \infty$ limit and use the CLT over c_S coefficients to get the fully connected approximation of the Hamiltonian as:

$$H(\sigma) = \sum_{\substack{p=2 \\ p \text{ even}}}^k \sum_{i_1, \dots, i_p} A_p g_{i_1, \dots, i_p} \sigma_{i_1} \cdots \sigma_{i_p},$$

where $g \sim \mathcal{N}(0, 1)$. Although the original Hamiltonian contains only positive contributions, the randomness in the hyperedge sampling induces fluctuations in the coefficients c_S . After centering around their mean, these coefficients behave approximately as independent Gaussian variables in the large-connectivity limit, by the Poisson-normal approximation. Hence, the fully connected model can still be represented as a Gaussian process with the same covariance structure.

To determine the correct normalization, we compute the variance of the Hamiltonian and impose that it remains extensive, that is $\mathbb{E}[H^2] \sim n$. This requirement fixes the prefactor A_p in front of each p -spin term ensuring that the total variance of $H(\sigma)$ grows proportionally to the system size:

$$\mathbb{E}[H^2] = \binom{k}{p} \frac{1}{n^p} \xrightarrow{\text{imposing}} \binom{k}{p} \frac{1}{n^p} = A_p^2 n$$

We get the prefactor leading, also in this case, to the form of the fully connected Hamiltonian:

$$H(\sigma) = \sum_{\substack{p=2 \\ p \text{ even}}}^k \sqrt{\binom{k}{p} \frac{1}{n^{p-1}}} \sum_{1 \leq i_1, \dots, i_p \leq n} g_{i_1, \dots, i_p} \sigma_{i_1} \cdots \sigma_{i_p}$$

Consequently, we obtain the fully connected approximation in the standard form of a mixed even p -spin model, which captures the same large-scale statistical properties as the original diluted system.

4.3.2 Covariance computation

Having established the fully connected form of the Hamiltonian, we now verify that its covariance depends only on the overlap between configurations. For two configurations σ^1 and σ^2 , we compute the expected value of the product of the Hamiltonians, which reads:

$$\begin{aligned} \mathbb{E}H(\sigma^1)H(\sigma^2) &= \sum_{\substack{p=2 \\ p \text{ even}}}^k \binom{k}{p} \frac{1}{n^{p-1}} \sum_{i_1, \dots, i_p} \sigma_{i_1}^1 \cdots \sigma_{i_p}^1 \sigma_{i_1}^2 \cdots \sigma_{i_p}^2 \\ &= \sum_{\substack{p=2 \\ p \text{ even}}}^k \binom{k}{p} \frac{1}{n^{p-1}} n^p R_{1,2}^p \\ &\text{or} \\ &= \sum_{p=1}^{k/2} \binom{k}{2p} \frac{1}{n^{2p-1}} n^{2p} R_{1,2}^{2p}. \end{aligned}$$

In conclusion, the hypergraph 2-coloring model shares the same effective mean-field structure as the k -NAESAT model, leading to a covariance that depends solely on the overlap between configurations. This structural analogy implies that, in the large-connectivity regime, the model is also expected to exhibit disorder chaos: small perturbations in the underlying disorder or in the graph structure induce a macroscopic reorganization of equilibrium configurations.

4.4 Disorder Chaos in random k -XORSAT

We now turn to another closely related system: the random k -XORSAT model, defined in Section 1.3.3, which despite its algebraic solvability, also reveals a rich spin glass-like structure and provides further insight into the universality of disorder chaos across different random constraint satisfaction problems. In contrast with other CSP models like the k -SAT or the k -NAESAT, the k -XORSAT has a linear algebraic structure: each clause represents an equation over the field $\mathbb{F}_2$, and the whole instance can be solved in polynomial time through gaussian elimination.

Despite this polynomial-time solvability, random k -XORSAT formulas exhibit rich structural properties, such as phase transitions and clustering phenomena, making this model very interesting for studying complex behaviors in disordered systems. In particular, also the k -XORSAT can be mapped onto a spin system, allowing us to use methods from statistical physics to better understand it.

4.4.1 The k -XORSAT Hamiltonian

To study the k -XORSAT problem using tools from statistical physics, we reformulate it in terms of spin variables, each Boolean variable $x_i \in \{0, 1\}$ is mapped to a spin variable $\sigma_i \in \{-1, +1\}$ via:

$$x_i = 1 \longleftrightarrow \sigma_i = +1 \quad \text{and} \quad x_i = 0 \longleftrightarrow \sigma_i = -1$$

Within this mapping, the XOR constraint

$$x_{i_1} \oplus x_{i_2} \oplus \cdots \oplus x_{i_k} = b$$

is satisfied if and only if the product of the corresponding spin equals $(-1)^b$, i.e.:

$$\prod_{j=1}^K \sigma_{i_j} = (-1)^b.$$

This allows us to define an energy penalty function for a single violated clause:

$$\theta = -\frac{1 - (-1)^b \prod_{k=1}^K \sigma_{i_k}}{2}$$

This convention ensures that satisfied clauses contribute zero energy, while violated ones contribute positively. Therefore, given an instance of m clauses indexed by $\alpha = 1, \dots, m$, the total Hamiltonian is:

$$H_\lambda(\sigma) = -\sum_{\alpha=1}^m \frac{1 - (-1)^{b_\alpha} \prod_{j=1}^k \sigma_{i_j}^\alpha}{2}.$$

This defines a spin glass system with k -body interactions, where the disorder is encoded in the right-hand side values b_α and in the random choice of variables in each clause, orrnally similar to other diluted models introduced before.

4.4.2 The k -XORSAT fully connected approximation

We now construct a fully connected approximation to the k -XORSAT model Hamiltonian, replacing the sparse interaction structure with a dense mean-field one. Neglecting additive constant, the Hamiltonian can be written as:

$$H_\lambda \approx -\frac{1}{2} \sum_{\alpha=1}^M J_\alpha \prod_{k=1}^K \sigma_{i_k}^\alpha, \quad (4.3)$$

where $(-1)^{b_\alpha} := J_\alpha$ are i.i.d. random signs taking values $+1, -1$ with equal probability. This shows that the random k -XORSAT model can be viewed as a random k -spin spin glass system with disorder given by the J_α 's coefficients. In the large-connectivity limit, we approximate the sum over the m random clauses by a sum over all possible k -tuples of spins, each weighted by an independent Gaussian coefficient. This leads to the fully connected Hamiltonian

$$H(\sigma) = \frac{1}{n^{(k-1)/2}} \sum_{1 \leq i_1, \dots, i_p \leq n} g_{i_1, \dots, i_k} \sigma_{i_1} \cdots \sigma_{i_p}$$

Which is the Hamiltonian for a k -spin spin model. The prefactor comes from computing the covariance of the Hamiltonian, since we want an energy which grows with n , in analogy with the pure p -spin model, where the coefficients $g_{i_1, \dots, i_k}$ are independent standard Gaussian random variables. The normalization is chosen so that $\text{Var}(H)$ scales linearly with n , ensuring that the energy is extensive, as in the pure p -spin model. Again, it is important to emphasize that this construction is

heuristic: the clause contributions in the diluted model are not independent, but in the high-connectivity regime their cumulative fluctuations behave approximately Gaussian. Such mean-field approximations and Gaussian interpolation methods are the subject of dedicated research, e.g. *Optimization on sparse random hypergraphs and spin glasses* [10], and are not trivial.

4.4.3 Covariance computation

The covariance can be computed explicitly. For two configurations σ^1 and σ^2 , the expectation of their product is:

$$\mathbb{E}[H(\sigma^1)H(\sigma^2)] = \frac{1}{n^{k-1}} \sum_{i_1, \dots, i_k} (\sigma_{i_1}^1 \sigma_{i_1}^2) \cdots (\sigma_{i_k}^1 \sigma_{i_k}^2),$$

using the overlap $R(\sigma^1, \sigma^2) = \frac{1}{N} \sum_{i=1}^N \sigma_i^1 \sigma_i^2$, we get:

$$\begin{aligned} \mathbb{E}[H(\sigma^1)H(\sigma^2)] &= \frac{1}{n^{k-1}} \left(\sum_{i=1}^n \sigma_i^1 \sigma_i^2 \right)^k \\ &= \frac{1}{n^{k-1}} n^k R^k = n R^k \end{aligned}$$

Thus, the covariance depends only on the overlap, as in the pure p -spin model. This confirms that, in the fully connected limit, random k -XORSAT belongs to the class of k -spin models whose covariance structure is determined solely by the overlap between configurations.

Notice that the prefactor appearing in the computation for the covariance follows again from the direct calculation to get the energy to be extensive:

$$\mathbb{E}[H(\sigma^1)H(\sigma^2)] = \frac{1}{n^{2\alpha}} \sum_{1 \leq i_1, \dots, i_p \leq n} \mathbb{E}[g_{i_1, \dots, i_p}]^2 \prod_{\ell=1}^k \sigma_{i_\ell}^1 \sigma_{i_\ell}^2 = \frac{n R^k}{n^{2\alpha}}$$

So

$$\frac{n^k}{n^{2\alpha}} \sim n \implies \alpha = \frac{k-1}{2}.$$

The fully connected approximation provides an effective Gaussian representation of the k -XORSAT Hamiltonian, characterized by a covariance proportional to R_{12}^k . This places the model within the same universality class as the p -spin spin glass, and thus the analytical results and heuristic arguments on disorder chaos developed for mean-field models apply here as well. Despite the algebraic solvability of XORSAT, its spin glass formulation captures the same qualitative phenomenon: small perturbations of the disorder lead to a rapid decorrelation between the equilibrium configurations, confirming the robustness of this phenomenon across different random CSPs.

Chapter 5

Simulations

The rigorous results obtained in the previous chapter establish the occurrence of **disorder chaos** in diluted random constraint satisfaction problems with large connectivity via an interpolation argument. However, the proof is by necessity asymptotic and does not provide quantitative information about the finite-size systems.

In order to complement the theoretical analysis we now turn to simulations. For each model, we investigate the behavior of the overlap as a function of the correlation parameter t : our goal is to observe how the average overlap $R(t)$ changes as the correlation decreases. In particular we aim to determine whether $R(t)$ exhibits a rapid decay when the correlation slightly decreases from $t = 1$. For each case, we generate pairs of correlated instances with correlation parameter t , compute ground states and measure their mutual overlap. This numerical exploration provides finite-size a perspective on the phenomenon of disorder chaos and complements theoretical results discussed in Chapter 3 and Chapter 4.

5.1 Random k -SAT

We performed numerical experiments on the random k -SAT problem with $n = 100$ variables, clause arity $k = 3$ and $M = 400$ number of clauses.

For each correlation parameter $t \in [0, 1]$, two formulas were generated with correlated disorder: given a set of random signs J^1 , the corresponding correlated set J^2 was obtained by keeping each literal identical with probability t and resampling it otherwise.

This interpolation produces two instances whose disorder is coupled with probability t , so that:

$$\mathbb{E}[J^1 J^2] = t.$$

In this way the parameter t directly controls the correlation of the disorder between the two formulas.

Each formula was optimized using a greedy local search algorithm with multiple restarts: starting from a random assignment, variables are flipped at random and the move is accepted only if it does not decrease the number of satisfied clauses. The best solution across the restart is retained.

To quantify the similarity between the solutions of two correlated formulas, we measure the overlap between the corresponding assignments σ^1 and σ^2 :

$$R(\sigma^1, \sigma^2) = \frac{1}{n} \left| \sum_{i=1}^n \sigma_i^1 \sigma_i^2 \right|,$$

where we take the absolute value because we are interested in the magnitude of the correlation between the two configurations, not in its orientation, this ensures that configurations which are globally or partially inverted are still recognized as strongly correlated. For each value of the correlation parameter t , we compute the overlap $R(\sigma^1, \sigma^2)$ between the two solutions and then average it several independent realizations of the disorder.

In this way we obtain the function $t \mapsto \mathbb{E}[R(t)]$, which describes how the similarity between ground states decays as the correlation of the disorders is reduced.

The result is showed in Fig.5.1. As expected, for $t = 1$ the two formulas coincide,

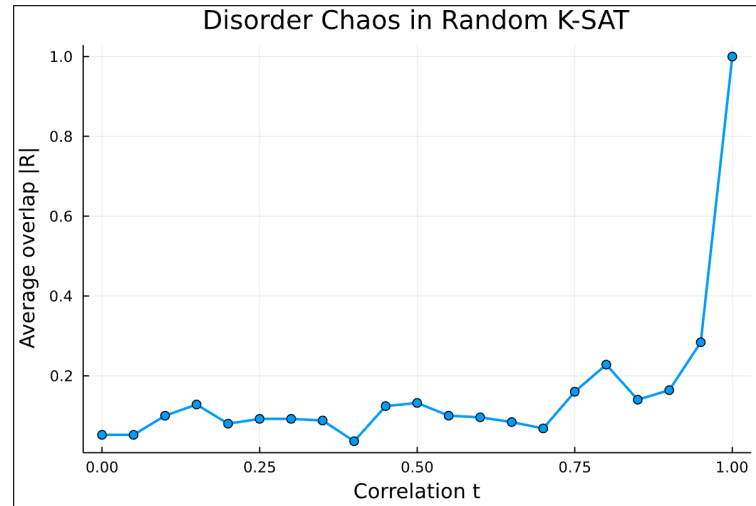


FIGURE 5.1: Average overlap R between solutions of correlated 3-SAT instances as a function of the correlation parameter t . The sharp decrease of $R(t)$ away from $t = 1$ illustrates the occurrence of disorder chaos in random k -SAT.

hence the same solution is recovered and the overlap is equal to 1. For $t = 0$, the two instances are completely independent and the overlap approaches to zero.

Most importantly, the decay of $R(t)$ as $t < 1$ is very sharp: even a small perturbation of the disorder produces ground state configurations that are essentially uncorrelated.

This can be seen in Fig.5.2, where it is shown that for $t = 0.95$ the overlap $R(t)$ the overlap rapidly decreases from 1 to 0.284.

t	Overlap
0.0	0.052
0.05	0.052
0.1	0.1
0.15	0.128
0.2	0.08
0.25	0.092
0.3	0.092
0.35	0.088
0.4	0.036
0.45	0.124
0.5	0.132
0.55	0.1
0.6	0.096
0.65	0.084
0.7	0.068
0.75	0.16
0.8	0.228
0.85	0.14
0.9	0.164
0.95	0.284
1.0	1.0

FIGURE 5.2: Numerical values of the overlap $R(t)$ for values of the correlation parameter t . Even at $t = 0.95$ the overlap has already dropped significantly. So the data confirms a sharp decay of the overlap away from $t = 1$.

5.1.1 Pseudo-code for the random k -SAT simulation

Algorithm 1 Simulation of Disorder Chaos in Random k -SAT

```

1: Input:  $n, k, M, ts, n_{\text{trials}}, max\_iter, n_{\text{restarts}}$ 
2: Output: List of average overlaps for each value of  $t$ 
3: function GENERATECLAUSE( $n, k$ )
4:   return  $k$  variables randomly selected from  $\{1, \dots, n\}$ 
5: end function
6: function CORRELATEDSIGNS( $t, k$ )
7:    $J_1 \leftarrow$  list of  $k$  random signs in  $\{-1, +1\}$ 
8:   for  $i = 1$  to  $k$  do
9:      $J_2[i] \leftarrow J_1[i]$  with probability  $t$ , otherwise a new random sign
10:  end for
11:  return  $(J_1, J_2)$ 
12: end function
13: function CLAUSESATISFIED( $clause, signs, \sigma$ )
14:  return  $\bigvee_i (\text{sign}_i \cdot \sigma_{\text{clause}_i} = 1)$ 
15: end function
16: function EVALUATEFORMULA( $clauses, signs, \sigma$ )
17:  return number of clauses satisfied by  $\sigma$ 
18: end function
19: function GREEDYSEARCH( $clauses, signs$ )
20:  for  $r = 1$  to  $n_{\text{restarts}}$  do
21:     $\sigma \leftarrow$  random assignment
22:    for  $k = 1$  to  $max\_iter$  do
23:      randomly choose  $i$  in  $1, \dots, n$ 
24:      flip  $\sigma[i]$ 
25:      if formula score improves then
26:        accept the change
27:      else
28:        revert the flip
29:      end if
30:    end for
31:    store the best  $\sigma$  found
32:  end for
33:  return best  $\sigma$ 
34: end function
35: Initialize  $results \leftarrow []$ 
36: for each  $t \in ts$  do
37:    $overlaps \leftarrow []$ 
38:   for  $trial = 1$  to  $n_{\text{trials}}$  do
39:     generate  $M$  random clauses
40:     generate  $(J_1, J_2)$  for each clause with correlation  $t$ 
41:     if  $t == 1$  then
42:        $\sigma \leftarrow$  GREEDYSEARCH( $clauses, J_1$ )
43:        $\sigma_1 \leftarrow \sigma_2 \leftarrow \sigma$ 
44:     else
45:        $\sigma_1 \leftarrow$  GREEDYSEARCH( $clauses, J_1$ )
46:        $\sigma_2 \leftarrow$  GREEDYSEARCH( $clauses, J_2$ )
47:     end if
48:      $R \leftarrow \frac{1}{n} |\sum_i \sigma_1[i] \cdot \sigma_2[i]|$ 
49:     append  $R$  to  $overlaps$ 
50:   end for
51:   append mean( $overlaps$ ) to  $results$ 
52: end for
53: return  $results$ 

```

5.2 Random k -NAESAT

We now turn to the random k -NAESAT problem. As seen in Chapter 1, this model is defined in the same setting as the k -SAT, but a clause is satisfied if and only if not all literals take the same value. We performed the same simulation procedure as before, generating pairs of correlated formulas with parameter t and optimizing them via greedy local search.

The result of the simulation shows the presence of disorder chaos also in this case.

We can see from the plot in Fig.5.3 that when the value of the correlation parameter t becomes slightly smaller than 1, a sharp transition of the overlap occurs.

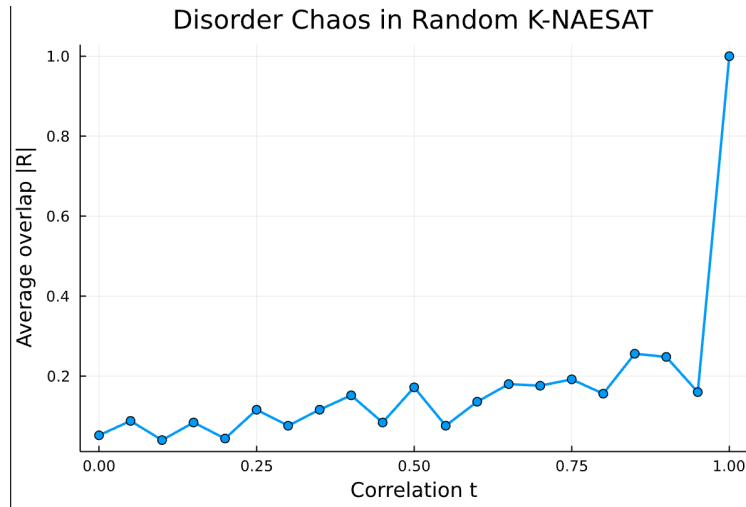


FIGURE 5.3: Average overlap R between solutions of correlated 3-NAESAT instances as a function of the correlation parameter t . The sharp decrease of $R(t)$ away from $t = 1$ illustrates also in this case the occurrence of disorder chaos in the random k -NAESAT. Notice also that the transition to chaos in this case is sharper than the one occurring in random k -SAT.

What is interesting to notice here, is the difference in the sharpness of the transition to disorder chaos between the k -SAT model and the k -NAESAT.

This could be due to symmetry reasons. In the k -NAESAT a clause is satisfied if and only if not all literals take the same value. This introduces a global symmetry in the model: if σ is a solution, then $-\sigma$ is a solution as well.

Heuristically, this would mean that optimal solutions "compete in pairs". A small change in the disorder can flip the preference between a configuration σ and its negation, so even tiny perturbations can cause the Gibbs measure to jump from one "cluster" to its opposite and the overlap collapses.

But in the k -SAT, this global symmetry does not exist. Therefore small perturbations tend to move the solutions "gradually".

This can be seen in the table shown in Fig.5.4, where we observe that for values of $t < 1$, like $t = 0.95$, the overlap rapidly decreases.

5.3 Random hypergraph 2-coloring

We next consider the random hypergraph 2-coloring problem. As seen in Chapter 1, in this model we generate a random k -uniform hypergraph with $n = 100$ vertices

t	Overlap
0.0	0.052
0.05	0.088
0.1	0.04
0.15	0.084
0.2	0.044
0.25	0.116
0.3	0.076
0.35	0.116
0.4	0.152
0.45	0.084
0.5	0.172
0.55	0.076
0.6	0.136
0.65	0.18
0.7	0.176
0.75	0.192
0.8	0.156
0.85	0.256
0.9	0.248
0.95	0.16
1.0	1.0

FIGURE 5.4: Numerical values of the overlap $R(t)$ for values of the correlation parameter t . We can clearly see that disorder chaos manifests for small deviation from $t = 1$. Moreover, the transition is sharper than the one occurring in k -SAT.

and $M = 400$ hyperedges.

Each hyperedge is a subset of k vertices, and the constraint requires that its vertices are not all assigned the same color.

To investigate the effect of disorder chaos, we generate two correlated hypergraphs with correlation parameter $t \in [0, 1]$.

Let us recall the hypergraph 2-coloring problem Hamiltonian. The function θ is given as:

$$\theta_{e_j}(\sigma) = \left(\prod_{i \in e_j} \frac{1 + \sigma_i}{2} + \prod_{i \in e_j} \frac{1 - \sigma_i}{2} \right),$$

where $e_j \subset [n]$ is a hyperedge of cardinality k .

Then, given a k -random uniform hypergraph with a Poisson-distributed number of clauses $\pi(\lambda n)$, the Hamiltonian takes the form:

$$H_\lambda(\sigma) = \sum_{j=1}^{\pi(\lambda n)} \left(\prod_{i \in e_j} \frac{1 + \sigma_i}{2} + \prod_{i \in e_j} \frac{1 - \sigma_i}{2} \right),$$

where each hyperedge e_j is chosen uniformly at random from the set of all k -tuples of vertices. Hence, the Hamiltonian counts the number of monochromatic hyperedges. Notice that, in this case there are no literal signs (J 's), the only source of randomness is the set of hyperedges itself.

To introduce a correlation $t \in [0, 1]$ between two instances we proceed as follows: for each position $m = 1, \dots, M$ we keep the m -hyperedge identical in the two copies with probability t , and we resample it otherwise.

In other words, if $e_m^{(1)}$ and $e_m^{(2)}$ denote the m -th hyperedge in the two instances, then:

$$\mathbb{P}(e_m^{(1)} = e_m^{(2)}) = t.$$

In this way, t directly controls the structural correlation of the disorder between the two hypergraphs. Then, as before, to quantify the similarity between two colorings σ^1 and σ^2 we can study the overlap $R(\sigma^1, \sigma^2) = \frac{1}{n} |\sum_{i=1}^n \sigma^1_i \sigma^2_i|$ between them. Then, following the same procedure as previously, the overlap is averaged for each value of t over several independent realizations of the disorder, leading to the function $t \mapsto \mathbb{E}[R(t)]$.

The results shown in Fig.5.5, confirm the presence of disorder chaos in random hypergraph 2-coloring. For $t = 1$, the two instances coincide and the overlap is equal to 1. As $t < 1$, the overlap decreases, indicating that even a small perturbation of the disorder leads to macroscopic rearrangements of the ground state coloring.

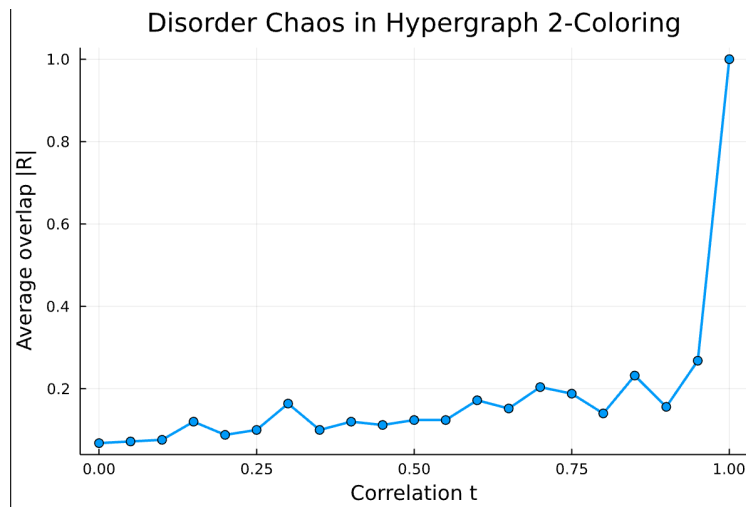


FIGURE 5.5: Average overlap R between optimized colorings of two correlated k -uniform hypergraphs ($n = 100$, $k = 3$, $M = 400$) as a function of the correlation t . The curve decays as soon as $t < 1$, evidencing disorder chaos.

Notice that at the level of local constraints, the hypergraph 2-coloring problem is very similar to the random k -NAESAT.

A k -NAESAT clause is violated if and only if all its k literals are equal and a k uniform hyperedge is violated if and only if all its k vertices have the same coloring.

In other words, a NAESAT clause is violated when:

$$J_{i_1} \sigma_{i_1} = J_{i_2} \sigma_{i_2} = \dots = J_{i_k} \sigma_{i_k},$$

while a 2-coloring hyperedge is violated when:

$$\sigma_{i_1} = \sigma_{i_2} = \dots = \sigma_{i_k}.$$

Now, to highlight this structural similarity of the two models, it could be possible to perform a change of variable that absorbs the J signs and turns each NAESAT clause into a **"not-all-equal-hyperedges"** constraint.

Moreover, both models share the same global $\mathbb{Z}_2$ symmetry: if σ is a solution, then $-\sigma$ is also a solution.

In this sense the cost functions θ and the Hamiltonians are nearly identical.

Despite the near identity of the local constraints, the mechanism of perturbation is actually harsher in k -NAESAT.

In NAESAT, random signs J produce spin-glass-like mean-zero interactions. A small decrease of t flips many clause preference and this triggers abrupt jumps of the Gibbs measure between opposite clusters.

In 2-coloring, by contrast, interactions are uniform, i.e.: there are no random signs J and the disorder is purely which hyperedges exist.

This could make the optimal coloring rearrange more gradually, so the drop in $\mathbb{E}[R(t)]$ near $t = 1$ is less steep, as shown in Fig.5.6.

t	Overlap
0.0	0.068
0.05	0.072
0.1	0.076
0.15	0.12
0.2	0.088
0.25	0.1
0.3	0.164
0.35	0.1
0.4	0.12
0.45	0.112
0.5	0.124
0.55	0.124
0.6	0.172
0.65	0.152
0.7	0.204
0.75	0.188
0.8	0.14
0.85	0.232
0.9	0.156
0.95	0.268
1.0	1.0

FIGURE 5.6: Numerical values of $\mathbb{E}[R(t)]$ for selected t . The data confirm a smoother decay near $t = 1$ than in k -NAESAT.

5.3.1 Pseudo-code for the hypergraph 2-col simulation

In this section the pseudo-code for the hypergraph 2-coloring is shown. Notice that the algorithm is very similar to the k -SAT simulation one, hence, only the specific operations for the hypergraph 2-col are shown. Greedy local search with restarts, the overlap computation, the averaging over trials, and the $t=1$ shortcut $\sigma^{(1)}=\sigma^{(2)}$ are identical to the k -SAT case and are omitted here.

Algorithm 2 Only the components specific to hypergraph 2-coloring

```

1: function CORRELATEDHYPEREDGES( $t, n, M, k$ )
2:    $C1 \leftarrow [], C2 \leftarrow []$ 
3:   for  $m = 1$  to  $M$  do
4:      $e \leftarrow k$ -subset of  $\{1, \dots, n\}$  sampled u.a.r.
5:     if Bernoulli( $t$ ) = 1 then                                 $\triangleright$  keep the same edge w.p.  $t$ 
6:       append  $e$  to  $C1$  and to  $C2$ 
7:     else
8:       append  $e$  to  $C1$  and a fresh  $k$ -subset to  $C2$ 
9:     end if
10:  end for
11:  return ( $C1, C2$ )                                            $\mathbb{E}[1\{e_m^{(1)} = e_m^{(2)}\}] = t$ 
12: end function
13: function CLAUSESATISFIED-2COL( $e, \sigma$ )                       $\triangleright$  NAE: not all equal
14:  return  $\neg(\text{all } \sigma_i = +1 \text{ or all } \sigma_i = -1) \text{ for } i \in e$ 
15: end function
16: function EVALUATE( $C, \sigma$ )
17:  return # of  $e \in C$  with CLAUSESATISFIED-2COL( $e, \sigma$ ) true
18: end function

```

5.4 Random k -XORSAT

Finally, we will go through the random k -XORSAT simulation.

We consider instances with $n = 100$ variables, clause arity $k = 3$ and $M = 250$ clauses.

Notice that here we chose $M = 250$ with $n = 100$ so that, at this scale, we obtain a non-negligible fraction of both SAT instances (solvable exactly via Gaussian elimination over $\mathbb{F}_2$) and UNSAT instances, where we measure the minimum number of violated equations. With $M = 400$ the inconsistency probability at these sizes increases substantially, and the analysis would focus almost exclusively on the UNSAT regime. Using $M = 250$ therefore allows us to observe both regimes within the same numerical setting.

Let us briefly recall that each clause impose a constraint as:

$$x_{i_1} \oplus x_{i_2} \oplus \dots \oplus x_{i_k} = b, \quad b \in \{0; 1\}.$$

In spin variables $\sigma_i \in \{-1; +1\}$ this is satisfied if and only if $\prod_{j=1}^k \sigma_{i_j} = (-1)^b$.

As in previous models, we measure energy as the number of violated equations.

To probe disorder chaos, we generate for each $t \in [0, 1]$, two copies on the same clause hypergraph, but with right hand side vectors $b^{(1)}$ and $b^{(2)}$ coupled independently clause by clause: with probability t , $b_\alpha^{(2)} = b_\alpha^{(1)}$; otherwise $b_\alpha^{(2)}$ is resampled in $\{0, 1\}$. Equivalently for the signs $(-1)^b$ we have $\mathbb{E}[(-1)^{b^{(1)}}(-1)^{b^{(2)}}] = 2t - 1$. Thus $t = 1$ yields to identical instances and $t = 0$ yields independent right hand sides.

Then, if an instance is SAT, we compute an exact solution by **Gaussian elimination**

over $\mathbb{F}_2$. This gives us a not approximated solution.

While, if the system is in the UNSAT regime, we need to find the assignment with the minimum number of violated equations. This can be done by **simulated annealing**, which is more convenient than a greedy algorithm: in this model the optimization landscape in the UNSAT regime is extremely rugged meaning many nearly optimal configurations are separated by high energy or entropic barriers. Standard greedy local search algorithms, which only accept improving moves, tend to get trapped very quickly in suboptimal minima. For this reason, we adopted simulated annealing to approximate the ground states. This algorithm allows for temporary uphill moves controlled by a decreasing temperature schedule, making it more capable of escaping shallow minima and exploring different valleys of the energy landscape. The relevance of such barriers has been analyzed in detail by Bellitti, Ricci-Tersenghi, and Scardicchio in *Entropic barriers as a reason for hardness in both classical and quantum algorithms* [11], who studied the dynamics of classical and quantum algorithms on the 3-XORSAT problem and identified entropic barriers as a main source of algorithmic hardness. Their results show that even simulated annealing tends to converge to positive-energy states due to the complex structure of the configuration space, confirming that XORSAT exhibits glassy behavior and slow relaxation. In our context, we use simulated annealing not as an exact solver, but as a more robust heuristic for exploring this rugged landscape compared to purely local greedy dynamics.

Then, we can measure the similarity of the solutions by computing the overlap between them as before: $R(\sigma^1, \sigma^2) = \frac{1}{n} |\sum_{i=1}^n \sigma^1_i \sigma^2_i|$. Then, following the same procedure as previously, the overlap is averaged for each value of t over several independent realizations of the disorder, leading to the function $t \mapsto \mathbb{E}[R(t)]$.

As expected, when $t=1$ the two instances coincide and the overlap is equal to one. On the other hand, when $t=0$ the right-hand sides are completely independent and the overlap approaches zero. $t \mapsto \mathbb{E}[R(t)]$ shows a very sharp decay as soon as $t < 1$. Even a small perturbation of the right-hand sides produces solutions that are essentially uncorrelated.

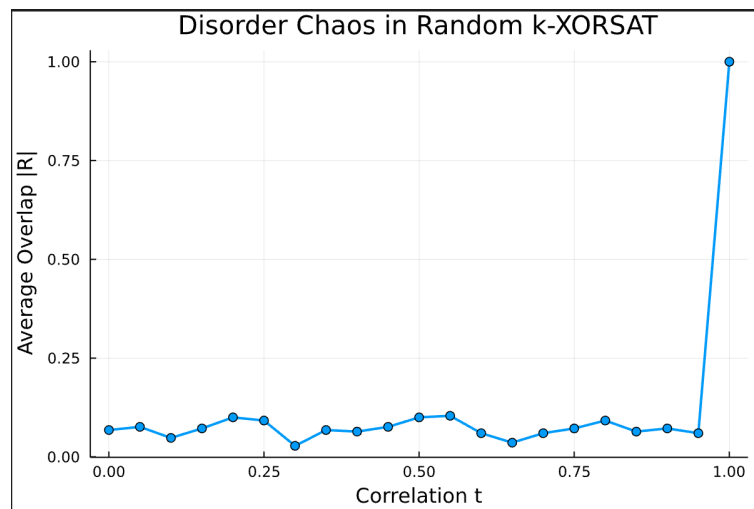


FIGURE 5.7: Average overlap R between solutions of correlated 3-XORSAT instances as a function of the correlation parameter t . The curve drops sharply as soon as $t < 1$, showing that even tiny perturbations of the right-hand side are sufficient to decorrelate the ground states. This provides numerical evidence of disorder chaos in random k -XORSAT.

t	Overlap
0.0	0.068
0.05	0.076
0.1	0.048
0.15	0.072
0.2	0.1
0.25	0.092
0.3	0.028
0.35	0.068
0.4	0.064
0.45	0.076
0.5	0.1
0.55	0.104
0.6	0.06
0.65	0.036
0.7	0.06
0.75	0.072
0.8	0.092
0.85	0.064
0.9	0.072
0.95	0.06
1.0	1.0

FIGURE 5.8: Numerical values of the average overlap $R(t)$ for selected values of the correlation parameter t . Already for $t = 0.95$ the overlap has decreased significantly, confirming the sharp sensitivity of 3-XORSAT solutions to disorder perturbations.

5.4.1 Pseudo-code for the simulated annealing (XORSAT)

Algorithm 3 Simulated Annealing for minimizing violated equations in k -XORSAT

```

1: function SA-MINIMIZE(clauses, rhs, var2cls, max_sweeps, restarts,  $T_0$ ,  $T_{min}$ , cooling)
2:    $best\sigma \leftarrow \emptyset$ ,  $bestE \leftarrow +\infty$ 
3:   for  $r = 1$  to restarts do
4:      $\sigma \leftarrow$  random assignment in  $\{0, 1\}^n$ 
5:     status  $\leftarrow$  CLAUSESTATUSES(clauses, rhs,  $\sigma$ )
6:      $E \leftarrow$  # of violated clauses in status
7:      $T \leftarrow T_0$ 
8:     steps  $\leftarrow$  max_sweeps  $\times n$ 
9:     for step = 1 to steps do
10:      choose variable  $i \in \{1, \dots, n\}$  uniformly
11:       $\Delta E \leftarrow$  DELTAENERGYFORFLIP( $i$ , var2cls, status)
12:      if  $\Delta E \leq 0$  or  $\text{rand}() < e^{-\Delta E/T}$  then
13:        flip  $\sigma[i]$ 
14:         $E \leftarrow E + \Delta E$ 
15:        UPDATESTATUSES AFTER FLIP( $i$ , var2cls, status)
16:        if  $E < bestE$  then
17:           $bestE \leftarrow E$ ,  $best\sigma \leftarrow \sigma$ 
18:          if  $bestE = 0$  then return  $best\sigma$ 
19:          end if
20:        end if
21:      end if
22:       $T \leftarrow \max(T \cdot \text{cooling}, T_{min})$ 
23:    end for
24:  end for
25:  return  $best\sigma$ 
26: end function

```

Chapter 6

Conclusions and Open Problems

This work suggests several directions for future research: one of them could be the study of Temperature Chaos for the studied models. Temperature Chaos is the statement that at fixed disorder two Gibbs measures at two different temperatures $\beta_1 \neq \beta_2$ yield typical configurations whose overlap, $R_{1,2} = \frac{1}{N} \sum_{i=1}^N \sigma_i^1 \sigma_i^2$ where $\sigma^1 \sim G_N(\beta_1)$ and $\sigma^2 \sim G_N(\beta_2)$, concentrates on a deterministic value as $N \rightarrow \infty$, where this value is typically zero.

In words: even a small change of temperature reshuffles the Gibbs weights so that configurations sampled at β_1 and β_2 become asymptotically decorrelated. The phenomenon of temperature chaos has been investigated in the physics literature for mean-field spin glasses, notably by Parisi and collaborators, but it remains largely open from a mathematical perspective. In particular, the paper *Chaos in Temperature in Diluted Mean-Field Spin Glasses* [12] by Parisi and Rizzo suggests that temperature chaos should be stronger in diluted systems than in fully connected ones. From a theoretical standpoint, proving temperature chaos even for the SK model remains an open and challenging problem, as current techniques based on interpolation and superconcentration, are not yet sufficient to capture this form of instability.

6.1 k -XORSAT: Temperature Chaos

One of the models which could be interesting to be studied from the point of view of Temperature Chaos, is the random k -XORSAT model. In the context of random k -XORSAT, introducing a positive temperature $T = \frac{1}{\beta}$ corresponds to relaxing the hard constraints of the system: we don't look for exact solutions for the XOR equation system, but we want to study the probability of having partially satisfying assignments in a thermodynamic context, modeled by an energy function and a Boltzmann distribution $P \propto \exp(-\beta H(x))$, where $H(x)$ is the Hamiltonian of the model. At zero temperature ($\beta \rightarrow \infty$), the measure concentrates on satisfying assignments, recovering the combinatorial optimization problem. At finite temperature, one expects configurations that minimize the Hamiltonian only approximately to contribute significantly. Understanding whether and how these Gibbs measures at nearby temperatures decorrelate, i.e., whether temperature chaos occurs in diluted systems such as random k -XORSAT, remains an open question. Based on heuristic arguments and numerical evidence (see again Parisi and Rizzo, [12]), one expects this effect to be even more pronounced than in fully connected models, due to the stronger influence of local fluctuations.

6.2 Conclusions

In this thesis we have investigated the phenomenon of disorder chaos, putting particular emphasis on diluted spin glass models and constraint satisfaction problems. Starting from the rigorous results established by Chen and Panchenko for the diluted random k -SAT model, we have provided a detailed reconstruction of the main ideas behind their proof and discussed how the phenomenon extends beyond fully connected mean-field systems.

We began by introducing the Sherrington-Kirkpatrick model and its generalization to mixed p -spin Hamiltonian, which serve as the fully connected limit of diluted models.

Within this framework, we reviewed the notion of overlap and showed that in the Sherrington-Kirkpatrick model, small perturbations in the disorder lead to a vanishing correlation between equilibrium configurations, an effect known as disorder chaos.

We then reinterpreted Random Constraint Satisfaction Problems: Random k -SAT, Random k -NAESAT, Hypergraph 2-coloring and Random k -XORSAT, as diluted spin systems, where variables interact through randomly selected constraints.

By studying perturbations in these models, either by resampling clauses (flipping fractions of spins) or by resampling signs (flipping independent spins), we demonstrated the close analogy between the two frameworks: even minor modifications in the disorder chaos cause optimal or near-optimal solutions to become almost orthogonal, indicating the same type of chaotic behavior observed in spin glasses.

Using an interpolation method between diluted and fully connected models, we followed the Guerra-Toninelli approach to analyze the derivative of the interpolated free-energy.

The incompatibility of the Gaussian and Poissonian distributions under the assumption of non-vanishing overlap leads to a contradiction, thereby confirming the emergence of disorder chaos in the diluted setting.

The rigorous argument was further supported by numerical simulations, which showed a sharp decay of the overlap as the correlation parameter t decreases from 1, across all the studied models.

Beyond the analytical results, this work suggests a broader message: the tools of statistical physics originally developed to study glassy systems, offer a unifying perspective on the complexity and instability of solution space in combinatorial optimization.

Understanding how solutions reorganize under disorder perturbations gives insight into the geometry of the energy landscape and may inform the design of more robust algorithms for random optimization problems.

Finally, several directions remain open. Extending these results to finite temperature, i.e. the study of temperature chaos in diluted mean-field models, hopefully is a next step, as well as disorder chaos in random k -SAT model for smaller connectivity values λ , which would require a proof that does not interpolate to the fully connected SK model.

Bibliography

- [1] David Sherrington and Scott Kirkpatrick. Solvable model of a spin-glass. *Physical review letters*, 35(26):1792, 1975.
- [2] Giorgio Parisi. Infinite number of order parameters for spin-glasses. *Physical Review Letters*, 43(23):1754, 1979.
- [3] Wei-Kuo Chen and Dmitry Panchenko. Disorder chaos in some diluted spin glass models. 2018.
- [4] Sourav Chatterjee. Disorder chaos and multiple valleys in spin glasses. *arXiv preprint arXiv:0907.3381*, 2009.
- [5] Sourav Chatterjee. *Superconcentration and related topics*, volume 15. Springer, 2014.
- [6] Dmitry Panchenko. On the k-sat model with large number of clauses. *Random Structures & Algorithms*, 52(3):536–542, 2018.
- [7] Luca Leuzzi and Giorgio Parisi. The k-sat problem in a simple limit. *Journal of Statistical Physics*, 103(5):679–695, 2001.
- [8] Sourav Chatterjee. A generalization of the lindeberg principle. 2006.
- [9] Michel Talagrand. *Spin glasses: a challenge for mathematicians: cavity and mean field models*, volume 46. Springer Science & Business Media, 2003.
- [10] Subhabrata Sen. Optimization on sparse random hypergraphs and spin glasses. *Random Structures & Algorithms*, 53(3):504–536, 2018.
- [11] Matteo Bellitti, Federico Ricci-Tersenghi, and Antonello Scardicchio. Entropic barriers as a reason for hardness in both classical and quantum algorithms. *Physical Review Research*, 3(4):043015, 2021.
- [12] Giorgio Parisi and Tommaso Rizzo. Chaos in temperature in diluted mean-field spin-glass. *Journal of Physics A: Mathematical and Theoretical*, 43(23):235003, 2010.