

Replica Symmetry Breaking without Replicas I. Pure States and Ultrametricity

Simone Franchini ^{*†}

February 1, 2018

Abstract

This is the first of a series of papers devoted to the interpretation of the Replica Symmetry Breaking (RSB) of Parisi et Al. [2]. Let Ω^N be the space of configurations of an N -spins system, each spin having a finite set Ω of inner states, and let $\mu : \Omega^N \rightarrow [0, 1]$ be some probability measure. Here we give an argument to encode any probability measure μ into a kernel function $M : [0, 1]^2 \rightarrow \Omega$, and use this notion to reinterpret the assumptions of the RSB ansatz, without using replicas. The present work is mainly intended to give the necessary mathematical background for further developments. ¹

1 Introduction

Originally introduced by Parisi in order to interpret its exact solution of the Sherrington-Kirkpatrick model (SK) [1, 2], the Replica Symmetry Breaking (RSB) ansatz proved to be an extremely valuable tool in explaining properties of disordered systems. Despite many technical advances, worth to cite the proof of the free energy functional by Guerra and Talagrand [3, 4], some of its fundamental features remain quite mysterious after more than thirty years.

A central role is played by the elusive concept of *pure state*. Despite a precise definition is still lacking, it is widely acknowledged that they must satisfy some properties. For example, it is expected that the connected correlation functions associated to these states vanish in the thermodynamic limit [2, 5]. This implies that in some sense the measure conditioned to those states can be described by a mean field model of some kind (see Part III of [5], updated 2014 version, for a non-rigorous but detailed discussion of the *finite volume pure states*).

Perhaps, the most striking and unconventional property is that the pure states are predicted to have a hierarchical structure such that the support of the

^{*}Goethe Mathematics Institute, 10 Robert Mayer Str, 60325 Frankfurt, Germany

[†]Sapienza Università di Roma, Piazzale Aldo Moro 1, 00185 Roma, Italy

¹MSC2000: 82D30, 60F10, Keywords: Spin Glasses, Replica Symmetry Breaking, Pure States, Parisi Formula

overlaps is ultrametric [2]. A considerable amount of work has been produced on this subject, culminating in a proof of ultrametricity for the SK model by Panchenko [6]. Anyway, whether ultrametricity and other properties of the pure states hold in some general framework, including their representation as well defined mathematical objects, proved to be an extremely hard task and remains an open question.

Inspired by a remarkable series of papers by Coja-Oghlan et al. which introduce tools from Graph Theory to study Belief Propagation algorithms [9, 10, 11], we propose that the set of pure states can be represented using a kernel, a two variables function

$$\hat{M} : [0, 1]^2 \rightarrow \Omega \tag{1.1}$$

that encodes the states of magnetization (hereafter we assume $\Omega = \{-1, 1\}$). As we shall see, this will allow for critical simplifications in reproducing the results of the RSB scheme, which we interpret as a technique to approximate some probability measure through a *mean field* filtered algebra² (in a sense that will become clear along the path). Since our theory need quite much space to be properly explored, we decided to split its presentation into a series of papers, the first being dedicated to the introduction and discussion of the general mathematical structures that subtend it.

The paper is organized as follows. In the next section we introduce a kernel representation for probability distributions. Then in Sections 2 and 3 we introduce an approximation argument to use matrices instead of kernels, which will greatly simplify our notation. We introduce our notion of Pure States in Section 4, and use it to construct a martingale representation. Finally, in Section 5 we use this representation to deal with the ultrametricity assumption of the Parisi ansatz, and identify the associated trial probability distribution (which is the main result of the present paper).

We do not apply our results to the Sherrington-Kirkpatrick model, being the applicative aspects the main subject of the second paper of this series. Although we will briefly confront our findings with the classical methods along the exposition (and when this is possible) we won't give an exhaustive comparative analysis since this would require substantial additional work to be carried on in a dedicated work.

2 Kernel representation

Let Ω^N be the product space of N spins with finite set Ω of inner states, let $\mathcal{P}(\Omega^N)$ be the ensemble of all probability measures on Ω^N and let $\mu \in \mathcal{P}(\Omega^N)$ some probability measure. We denote by μ_K the marginal distribution of μ over a subset $K \subset \{1 \dots N\}$ of $|K|$ distinct spin coordinates (hereafter the modulus $|\cdot|$ when applied to a set returns its size). Then, if S_α , $\alpha \geq 1$ is a partition of Ω^N into a number of disjoint subsets we call μ^α the measure conditioned

²In a *filtered algebra* there are distinguished subspaces S_α , indexed by the elements of a totally ordered group, such that $S_\alpha \subseteq S_\beta$ for $\alpha < \beta$, and $S_\alpha \otimes S_\beta \subseteq S_{\alpha+\beta}$.

to S_α and by μ_K^α the marginal distribution of μ^α over K . For this paper we indicate random variables with bold characters (for example any bold variable \mathbf{X} is assumed to be random, while X is not).

The connection between kernels and the pure states of the RSB ansatz has been first noticed in [9], where a kernel encoding of μ is introduced in order to prove the following

Lemma. (*Bapst, Coja-Oghlan, 2016*) *For any $\mu \in \mathcal{P}(\Omega^N)$ it is possible to take some arbitrary small $\epsilon > 0$ and a partition of Ω^N into a finite number $n \geq n(\epsilon, |K|)$, not dependent from N , of disjoint subsets S_α , $0 \leq \alpha \leq n$ such that $\mu(S_0) \leq \epsilon$ and*

$$\sum_{K \in \{1, \dots, N\}^{|K|}} \left\| \mu_K^\alpha - \bigotimes_{i \in K} \mu_i^\alpha \right\|_{TV} \leq \epsilon N^{|K|}, \quad \forall \alpha, |K| \geq 1 \quad (2.1)$$

if N is chosen large enough (we denoted by $\|\cdot\|_{TV}$ the total variation³). For example, in the case $|K| = 2$ we can write

$$\sum_{\{i,j\} \in \{1, \dots, N\}^2} \left\| \mu_{\{i,j\}}^\alpha - \mu_i^\alpha \otimes \mu_j^\alpha \right\|_{TV} \leq \epsilon N^2, \quad \forall \alpha. \quad (2.2)$$

Proof. It is essentially a measure theoretic version of the Szemerédi Regularity Lemma, see Chapter 9.2 and 9.3 of [14] for a detailed review. A proof of Eq.s (2.1) and (2.2) can be found in the first part of [9], after the statements of Theorem 2.2 and Corollaries 2.3-2.5. \square

The above result tells us that for any measure μ describing a system of variables with finite set Ω of inner states we can decompose our sample space Ω^N into a finite number $n(\epsilon, |K|)$ of regular disjoint subsets S_α , $1 \leq \alpha \leq n(\epsilon, |K|)$ plus one irregular S_0 with $\mu(S_0) \leq \epsilon$ such that for any regular subset S_α the marginals of μ^α over a randomly chosen set K can be approximated by a product measure in the sense of Eq. (2.1). Surprisingly, the number $n(\epsilon, |K|)$ of such regular subsets only depends on k , $|\Omega|$ and the level of precision ϵ we want to achieve for our approximation, and it does not depend on the size N of the system.

This and many other results can be obtained by noticing that both probability measures and graphs can be exactly encoded into kernel functions. For example, in [9, 10] a new distance on $\mathcal{P}(\Omega^N)$ based on Graph Theory is introduced to characterize Gibbs Measures directly in the thermodynamic limit (Cut Distance^{4,5}, see Chapter 8.2 of [14], or [10] for the measure theoretic approach).

³Given two measures $\mu, \nu : S \rightarrow [0, 1]$ and some $A \subseteq S$ the *total variation distance* between μ and ν is given by the formula $\|\mu - \nu\|_{TV} = 2 \sup_A |\mu(A) - \nu(A)|$.

⁴Let M, W be two kernels and let $\theta = (\theta_1, \theta_2)$ be a pair of measure preserving maps. We call *cut norm* the positive quantity $\|M\|_\square = \sup_{A,B} |\int_{A,B} M(x,y) dx dy|$ and *cut distance* $D_\square(W, U) = \inf_\theta \|W - U^\theta\|_\square$, where U^θ stands for $U(\theta_1(x), \theta_2(y))$ [14].

⁵In the context of probability theory the cut distance between $\mu, \nu : S \rightarrow [0, 1]$ is the cut distance $D_\square(M_\mu, M_\nu)$ between the associated kernels M_μ, M_ν of Eq.(2.6) below. It can be

The above Lemma is itself a probabilistic version of the celebrated Szemerédi Regularity Lemma (Chapter 9.2 and 9.3 of [14]), whose typical mark can be recognized in the independence of the partition size from the size of the partitioned set.

Unfortunately the technical complexity of these arguments would require quite much space to be properly explained. Since the original arguments presented in this work do not necessarily require the use Szemerédi Partitions (at the cost of working with finite systems only) we won't systematically deepen this argument here, but we stress that these are useful mathematical concepts and we warmly advice the reader to look at [14] for further reading on this important subject. In general, here we try to avoid mathematical concepts unfamiliar to the physics literature as much as we can (indeed, we try to avoid inessential technicalities in general).

Before entering in the core of the discussion some mathematical introduction is mandatory in order to justify our later arguments. In the following we show how to encode any probability distribution $\mu \in \mathcal{P}(\Omega^N)$ into a spin array with a countably infinite number of entries. We start by introducing an exact Kernel Representation of μ

Definition 1. (*Magnetization Kernel*) *Let label the sample space as follows*

$$\Omega^N = \{\sigma^a : 1 \leq a \leq 2^N\} \quad (2.3)$$

where $\sigma^a \in \Omega^N$ are the sample vectors

$$\Omega^N \ni \sigma^a = \{\sigma_k^a \in \Omega : 1 \leq k \leq N\}. \quad (2.4)$$

Then, the Magnetization Kernel

$$\hat{M}_\mu : [0, 1]^2 \rightarrow \Omega \quad (2.5)$$

associated to $\mu \in \mathcal{P}(\Omega^N)$ is the step function

$$\hat{M}_\mu(x, y) = \sum_{a=1}^{2^N} \sum_{k=1}^N \sigma_k^a \mathbb{I}_{\{x \in \hat{V}_k, y \in \hat{S}_a\}} \quad (2.6)$$

where \hat{V}_k and \hat{S}_a are the intervals

$$\hat{V}_k = [x_{k-1}, x_k), \quad \hat{S}_a = [y_{a-1}, y_a) \quad (2.7)$$

with boundaries x_k and y_a given by

$$x_k = k/N, \quad y_a = \sum_{a'=1}^a \mu(\sigma^{a'}). \quad (2.8)$$

An example is given in Figure 2.1 .

shown that the kernel space is compact cut distance, and that convergence in cut distance is stronger than weak* convergence when dealing with intensive quantities, such as the free energy density associated to a Gibbs measure (see Chapter 8 of [14] and therein, or the first part of [9, 10]).

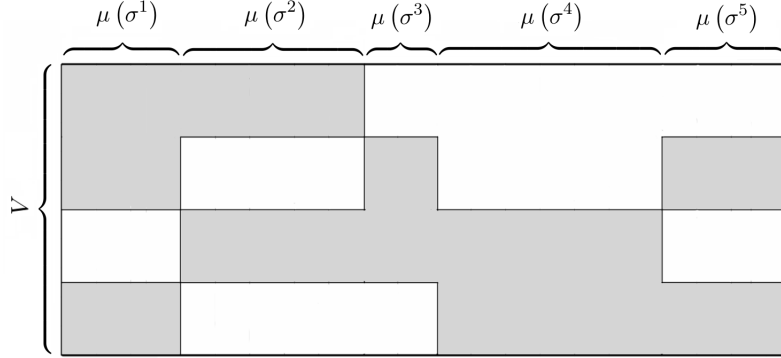


Figure 2.1: Kernel representation $\hat{M}_\mu(x, y)$ of Eq. (2.6) for a system of $N = 4$ spins described by an atomic probability distribution with $\mu(\sigma^a) = 0$ for $6 \leq a \leq 16$ and $\mu(\sigma^a) > 0$ for $1 \leq a \leq 5$, with $\sigma^1 = (+, +, -, +)$, $\sigma^2 = (+, -, +, -)$, $\sigma^3 = (-, +, +, -)$, $\sigma^4 = (-, -, +, +)$, $\sigma^5 = (-, +, -, +)$. The column index y has been rescaled for a better visualization.

It is intuitively clear that for finite spin systems the magnetization kernel \hat{M}_μ exactly encodes the associated probability measure μ from which is computed. In fact, we can reconstruct μ from \hat{M}_μ using the following lemma

Lemma 2. (Reconstruction Lemma) Given the kernel $\hat{M}_\mu : [0, 1]^2 \rightarrow \Omega$ associated to the probability density function $\mu \in \mathcal{P}(\Omega^N)$, then for any sequence

$$\hat{x} = \{ \hat{x}_k \in \hat{V}_k : 1 \leq k \leq N \} \quad (2.9)$$

and any measure preserving map $\theta : [0, 1] \rightarrow [0, 1]$ the probability density function μ is exactly reconstructed using the following relation

$$\mu(\sigma) = \int_{y \in [0, 1]} dy \prod_{k=1}^N \left(\frac{1 + \hat{M}_\mu(\hat{x}_k, \theta(y)) \sigma_k}{2} \right). \quad (2.10)$$

Proof. We start from the chain of identities

$$\mu(\sigma) = \langle \mathbb{I}_{\{\sigma = \sigma\}} \rangle_\mu = \left\langle \prod_{k=1}^N \left(\frac{1 + \sigma_k \sigma_k}{2} \right) \right\rangle_\mu = \sum_{a=1}^{2^N} \mu(\sigma^a) \prod_{k=1}^N \left(\frac{1 + \sigma_k^a \sigma_k}{2} \right) \quad (2.11)$$

then we notice that by definition

$$\mu(\sigma^a) = \int_{y \in \hat{S}_a} dy \quad (2.12)$$

and that $\hat{M}_\mu(\hat{x}_k, y_a) = \sigma_k^a$ for any $y \in \hat{S}_a$. Putting together we can write

$$\begin{aligned} \mu(\sigma) &= \sum_{a=1}^{2^N} \int_{y \in \hat{S}_a} dy \prod_{k=1}^N \left(\frac{1 + \hat{M}_\mu(\hat{x}_k, y) \sigma_k}{2} \right) = \\ &= \int_{y \in \bigcup_a \hat{S}_a} dy \prod_{k=1}^N \left(\frac{1 + \hat{M}_\mu(\hat{x}_k, y) \sigma_k}{2} \right) = \\ &= \int_{y \in [0,1]} dy \prod_{k=1}^N \left(\frac{1 + \hat{M}_\mu(\hat{x}_k, y) \sigma_k}{2} \right). \end{aligned} \quad (2.13)$$

The invariance under $y \rightarrow \theta(y)$ is granted by the fact that the integral operator is independent from the order of the infinitesimal steps dy . \square

An intriguing fact is that \hat{M}_μ is a *graphon*⁶, but is not clear at first look what kind of physical meaning could have its associated graph, whose edges “connect” in some sense the spin space with the states of the magnetization.

Even if array representations in the context of Spin Glasses have been considered since the very beginning (ie, overlap matrix of [2]), their use to represent probability distributions is quite recent. Before [9, 10], for example, the Aldous-Hoover theorem has been invoked in [7] in order to encode the replicated distribution of the SK model into a function $W : [0, 1]^4 \rightarrow \Omega$, but the use of a four dimensional order parameter is redundant in view of the Kernel representation above, that only requires a two dimensional function. The formula of Lemma 2 should be rather confronted with the De Finetti representation for exchangeable sequences [12]. Notice that for finite systems the kernel of μ is not unique because there is an infinite number of possible choices {for \hat{x} and for the map θ that points to the same μ , and we can also change the integral on y with a sum on $\hat{y}_a \in \hat{S}_a$. Also notice that \hat{M}_μ is defined up to an arbitrary index relabeling. Formally, if we take two bijections

$$\theta_{\hat{V}} : \{1 \dots N\} \rightarrow \{1 \dots N\}, \quad \theta_{\hat{S}} : [0, 1] \rightarrow [0, 1] \quad (2.14)$$

such that $k' = \theta_{\hat{V}}(k)$, $y' = \theta_{\hat{S}}(y)$, then clearly holds that

$$\mu(\sigma) = \int_{y \in [0,1]} dy \prod_{k=1}^N \left(\frac{1 + \hat{M}_\mu(\hat{x}_k, y) \sigma_k}{2} \right) = \int_{y' \in [0,1]} dy' \prod_{k'=1}^N \left(\frac{1 + \hat{M}_\mu(\hat{x}_{k'}, y') \sigma_{k'}}{2} \right). \quad (2.15)$$

By the way, these are the only sources of ambiguity when encoding measures into kernels, and it will prove a useful property later.

The following simple lemma express one of the main technical advantages of the kernel representation.

⁶Contraction for graph(function), first introduced by Lovasz. See Chapter 7.1 of [14] and therein for a detailed review of kernel functions in the context of Graph Theory.

Lemma 3. (*Correlations and Overlaps*) Let $1 \leq k, k' \leq N$ and let $\hat{x}_k \in \hat{V}_k$, $\hat{x}_{k'} \in \hat{V}_{k'}$, then the scalar product between the two rows \hat{x}_k and $\hat{x}_{k'}$ of \hat{M}_μ is the two point correlation function

$$\langle \sigma_k \sigma_{k'} \rangle_\mu = \int_{y \in [0,1]} dy \hat{M}_\mu(\hat{x}_k, y) \hat{M}_\mu(\hat{x}_{k'}, y). \quad (2.16)$$

Moreover, let σ^a and $\sigma^{a'}$ be two states of magnetization, and let $y_a^* \in \hat{S}_a$ and $y_{a'}^* \in \hat{S}_{a'}$, then the scalar product between the columns \hat{y}_a and $\hat{y}_{a'}$ of the kernel \hat{M}_μ is the magnetization overlap between these states

$$q(\sigma^a, \sigma^{a'}) = \frac{1}{N} \sum_{i=1}^N \sigma_k^a \sigma_k^{a'} = \int_{x \in [0,1]} dx \hat{M}_\mu(x, \hat{y}_a) \hat{M}_\mu(x, \hat{y}_{a'}). \quad (2.17)$$

Proof. The proof trivially follows by substituting the definition of \hat{M}_μ into the above formulas. \square

As we can see the kernel function provides a powerful visual encoding of both correlations and overlaps (and the event algebra in general). Clearly we can write higher order correlation functions and overlaps using the same proceeding, which suggests a first definition for the Pure States of the magnetization.

Definition 4. (*Microscopic Pure States*) We call *Microscopic Pure States* of μ the stepfunctions corresponding to the columns $\{\hat{M}_\mu(x, y) : x \in [0, 1]\}$ of \hat{M}_μ , with $y \in [0, 1]$.

Although they are still not the Pure States as intended in the Parisi Ansatz, the latter will be obtained starting from the above objects.

3 Matrix approximation

Another important fact about the kernel representation is that in Eq. (2.10) the weighted average $\langle \cdot \rangle_\mu$ is replaced by a uniform average on $y \in [0, 1]$. Clearly, if we could encode most of our distribution μ into some matrix M rather than a real valued kernel \hat{M}_μ then we would still work with a uniform average on the pure states while recovering all the simplification given by working with atomic measures.

Definition 5. (*Magnetization Matrix*) Let S and V be two integer vectors

$$V = \{i \in \mathbb{N} : 1 \leq i \leq |V|\}, \quad S = \{\alpha \in \mathbb{N} : 1 \leq \alpha \leq |S|\}, \quad (3.1)$$

then we call *Magnetization Matrix* $M : S \otimes V \rightarrow \Omega$ any binary matrix of the kind

$$M = \{m_i^\alpha \in \Omega : 1 \leq i \leq |V|, 1 \leq \alpha \leq |S|\} \quad (3.2)$$

and define the atomic probability distribution $\eta_M : \Omega^V \rightarrow [0, 1]$ that is obtained from the Magnetization Matrix as follows

$$\eta_M(\sigma) = \frac{1}{|S|} \sum_{\alpha \in S} \prod_{i \in V} \left(\frac{1+m_i^\alpha \sigma_i}{2} \right). \quad (3.3)$$

We call *Microscopic Pure States* of η_M the columns of M , ie the magnetization vectors corresponding to the atoms of η_M . Hereafter we will assume $\Omega = \{-1, 1\}$, $|V| = N$ and $|S| = n > N$.

In [9, 10] the concept of ϵ -regularity is introduced to approximate μ with an atomic probability distribution with finite number of atoms independent from the dimension N of σ , this at the cost of an error ϵ when computing the averages of intensive quantities. This can be done by using Szemerédi Regularity Lemma to approximate in cut distance the kernel \hat{M}_μ with a step function with finite numbers of steps, and then use the fact that cut distance convergence is stronger than weak convergence (Chapter 8, Lemma 8.22 of [14]).

This approach allows to work directly in the thermodynamic limit $N \rightarrow \infty$, and we point out that some form of Szemerédi partitioning seems unavoidable in order to achieve this remarkable feature. Fortunately, if we drop this requirement we can approximate μ with η_M using almost trivial arguments.

Theorem 6. (*Matrix approximation of μ*) Let $\epsilon > 0$. For any probability density function $\mu \in \mathcal{P}(\Omega^N)$ there is an integer

$$n(\epsilon, N) \leq 2^{N+1}/\epsilon, \quad (3.4)$$

in general dependent from N , and a Magnetization Matrix

$$M : S \otimes V \rightarrow \Omega \quad (3.5)$$

having $|V| = N$ rows and $|S| = n \geq n(\epsilon, N)$ columns such that for any bounded function of the kind $f : \Omega^N \rightarrow \mathbb{R}$, with $\|f\| < \infty$ holds

$$|\langle f(\sigma) \rangle_\mu - \langle f(\sigma) \rangle_{\eta_M}| \leq \epsilon \|f\|, \quad (3.6)$$

where η_M is the distribution obtained from the matrix M using Eq. (3.3). Hereafter we write $\eta_M \approx \mu$, or equivalently $M \approx \hat{M}_\mu$, to indicate that the magnetization matrix M approximates \hat{M}_μ in the sense of Eq.(3.6) above.

Proof. The first step is to notice that if $\mu(\sigma^\alpha) \in \mathbb{Q}$ for any α then μ can be encoded into a Magnetization matrix exactly. By confronting the kernel \hat{M}_μ defined in Eq. (2.6) with the magnetization kernel associated to the regularized measure η_M

$$\hat{M}_{\eta_M}(x, y) = \sum_{\alpha \in S} \sum_{i \in V} m_i^\alpha \mathbb{I}_{\{x \in [(i-1)/N, i/N), y \in [(\alpha-1)/|S|, \alpha/|S|)\}}. \quad (3.7)$$

we see that the partition parameters of the rows $[0, 1) \ni x$ are already the same if we take $k = i$, and by definition of rational number some $n \in \mathbb{N}$ must exist

such that $n\mu(\sigma^a) \in \mathbb{N}$ for all a . We can split $[0, 1) \ni y$ into $|S| = n$ equally sized subsets being sure that

$$\{y_a : 1 \leq a \leq 2^N\} \subseteq \{\alpha/n : 1 \leq \alpha \leq n\}, \quad (3.8)$$

with y_a given in Eq. (2.8). Then, if $y_a - y_{a-1} = \mu(\sigma^a) \in \mathbb{Q}$ holds for all a it suffices to refine the column indexes $[0, 1) \ni y$ of \hat{M}_μ up to subsets of size $1/n$ to obtain \hat{M}_{η_M} .

The extension to general case $\mu(\sigma^a) \in \mathbb{R}$ can be obtained via the following surgery on $[0, 1) \ni y$. Starting from the definition of Eq. (2.6) we can rewrite the parameters y_a using the recursive formula

$$y_a = y_{a-1} + \mu(\sigma^a), \quad (3.9)$$

then we introduce the modified variables

$$y_{a,n} = y_{a-1} + \lfloor n\mu(\sigma^a) \rfloor / n \quad (3.10)$$

where y_{a-1} is unchanged and $\lfloor \cdot \rfloor$ indicates the lower integer part. Using this new set of variables we can split each interval \hat{S}_a into two components

$$\hat{S}_a = [y_{a-1}, y_a) = [y_{a-1}, y_{a,n}) \cup [y_{a,n}, y_a) = \bar{S}_a \cup [y_{a,n}, y_a) \quad (3.11)$$

such that $|\bar{S}_a| \in \mathbb{Q}$ and $|[y_{a,n}, y_a)| \in \mathbb{R}$. Let group these intervals into two sets, the first is

$$\bar{S} = \bigcup_{1 \leq a \leq 2^N} [y_{a-1}, y_{a,n}) = \bigcup_{1 \leq a \leq 2^N} \bar{S}_a \quad (3.12)$$

and is composed by the union of intervals of rational size. The second set

$$\bar{S}_0 = \bigcup_{1 \leq a \leq 2^N} [y_{a,n}, y_a) \quad (3.13)$$

is the union the reminders. Clearly $[0, 1) = \bar{S} \cup \bar{S}_0$, then $1 - |\bar{S}_0| = |\bar{S}|$ with

$$|\bar{S}_0| = \sum_{a=1}^{2^N} (\mu(\sigma^a) - \lfloor n\mu(\sigma^a) \rfloor / n). \quad (3.14)$$

The above procedure is illustrated in Figure 3.1.

At this point we are ready to compute the bound for the average $\langle f(\sigma) \rangle_\mu$ by separating the contributions of \bar{S} and \bar{S}_0 . By Lemma 2 we can split the two contributions into

$$\begin{aligned} \langle f(\sigma) \rangle_\mu &= \int_{y \in [0,1]} dy \sum_{\sigma \in \Omega^N} f(\sigma) \prod_{k=1}^N \left(\frac{1 + \hat{M}_\mu(\hat{x}_k, y) \sigma_k}{2} \right) = \\ &= \int_{y \in \bar{S}} dy \sum_{\sigma \in \Omega^N} f(\sigma) \prod_{k=1}^N \left(\frac{1 + \hat{M}_\mu(\hat{x}_k, y) \sigma_k}{2} \right) + \\ &\quad + \int_{y \in \bar{S}_0} dy \sum_{\sigma \in \Omega^N} f(\sigma) \prod_{k=1}^N \left(\frac{1 + \hat{M}_\mu(\hat{x}_k, y) \sigma_k}{2} \right), \quad (3.15) \end{aligned}$$

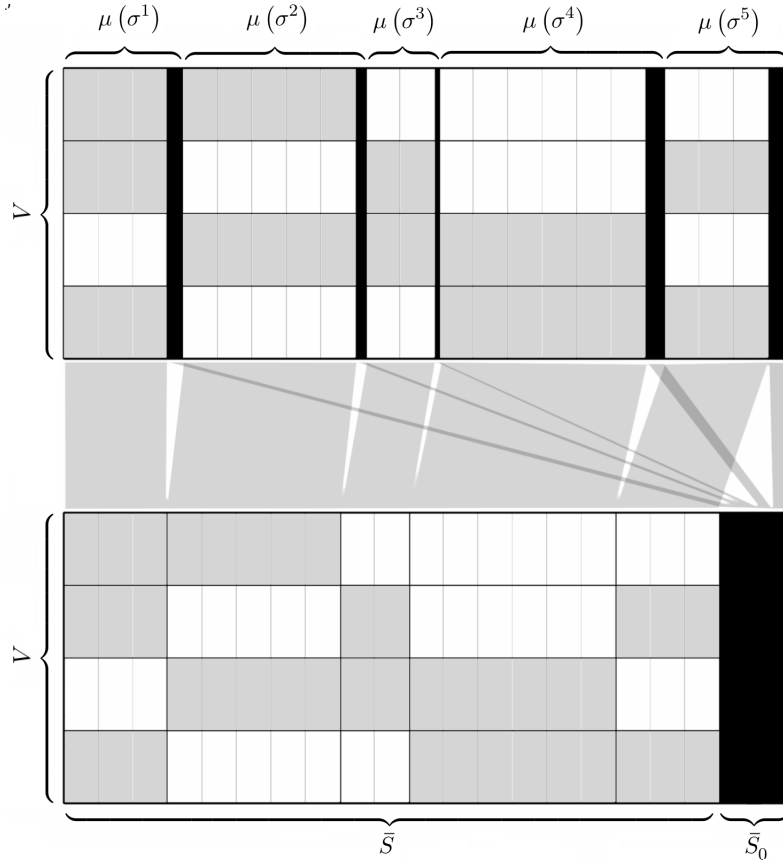


Figure 3.1: Surgery on \hat{M}_μ with $n = 19$ for the same μ of Figure 2.1. The black areas are the sub-kernel associated to the reminders $[y_{a,n}, y_a)$. A map is also applied (shown in gray color) that collects the reminders at the bottom of the kernel to highlight the Magnetization Matrix. The column index y has been rescaled as in Figure 2.1.

then, applying the definitions we find

$$\int_{y \in \bar{S}} dy \sum_{\sigma \in \Omega^N} f(\sigma) \prod_{k=1}^N \left(\frac{1 + \hat{M}_\mu(\hat{x}_k, y) \sigma_k}{2} \right) = |\bar{S}| \langle f(\sigma) \rangle_{\eta_M} = (1 - |\bar{S}_0|) \langle f(\sigma) \rangle_{\eta_M} \quad (3.16)$$

for the first term and

$$\int_{y \in \bar{S}_0} dy \sum_{\sigma \in \Omega^N} f(\sigma) \prod_{k=1}^N \left(\frac{1 + \hat{M}_\mu(\hat{x}_k, y) \sigma_k}{2} \right) = |\bar{S}_0| \langle f(\sigma) \rangle_{\mu^0} \quad (3.17)$$

for the correction. Then, we just bound with

$$|\langle f(\sigma) \rangle_\mu - \langle f(\sigma) \rangle_{\eta_M}| = |\bar{S}_0| |\langle f(\sigma) \rangle_{\mu^0} - \langle f(\sigma) \rangle_{\eta_M}| \leq 2 |\bar{S}_0| \|f\|. \quad (3.18)$$

It remains to compute the upper bound of the reminder $|\bar{S}_0|$. This can be done by arguing that $\mu(\sigma^a) - \lfloor n \mu(\sigma^a) \rfloor / n \leq 1/n$ holds for all $n \in \mathbb{N}$, then we must have $|\bar{S}_0| \leq 2^N/n < \epsilon/2$, from which $n > 2^{N+1}/\epsilon$. Notice that this bound is far from being optimal. Although sufficient for our aims, we remark that it could be greatly improved even in this very general setting (for example iterating the above construction to approximate the irregular subset \bar{S}_0 and so on, or using Szemerédi Regularity Lemma as in [9, 10]). \square

Obviously for $N < \infty$ we can always choose n big enough to make the error bound $2^{N+1}/\epsilon$ arbitrarily small, then for the rest of this paper we will mostly work with the Magnetization Matrix M , and automatically assume $M \approx \hat{M}_\mu$ with ϵ infinitesimally small (unless specified otherwise).

It worth notice that M , like \hat{M}_μ , is defined up to a pair of bijections

$$\theta_V : V \rightarrow V, \theta_S : S \rightarrow S \quad (3.19)$$

in fact also in this case if $i' = \theta_V(i)$, $\alpha' = \theta_S(\alpha)$ then obviously

$$\eta_M(\sigma) = \frac{1}{|S|} \sum_{\alpha \in S} \prod_{i \in V} \left(\frac{1 + m_i^\alpha \sigma_i}{2} \right) = \frac{1}{|S|} \sum_{\alpha' \in S} \prod_{i' \in V} \left(\frac{1 + m_{i'}^{\alpha'} \sigma_{i'}}{2} \right). \quad (3.20)$$

Working with a magnetization matrix largely simplifies the notation, and allows for an easier visualization of the kernel operations, as we can appreciate from the following simple lemmas that will be useful in the next sections.

Lemma 7. (*Correlation functions of η_M*) Given a magnetization matrix M having N rows, $n > N$ columns with $n/N \in \mathbb{N}$, and a generalized N -point function of the kind

$$G(\sigma) = \prod_{i \in V} g(\sigma_i) \quad (3.21)$$

with $g : \Omega \rightarrow \mathbb{R}$, $\|g\| < \infty$, then the average is given by

$$\langle G(\sigma) \rangle_{\eta_M} = \langle \prod_{i \in V} g(\sigma_i) \rangle_{\eta_M} = \frac{1}{|S|} \sum_{\alpha \in S} \prod_{i \in V} g(m_i^\alpha). \quad (3.22)$$

Proof. Follows directly from the definition of η_M

$$\begin{aligned} \langle \prod_{i \in V} g(\sigma_i) \rangle_{\eta_M} &= \sum_{\sigma \in \Omega^N} \eta_M(\sigma) \prod_{i \in V} g(\sigma_i) = \\ &= \frac{1}{|S|} \sum_{\alpha \in S} \prod_{i \in V} \sum_{\sigma_i \in \Omega} g(\sigma_i) \left(\frac{1+m_i^\alpha \sigma_i}{2} \right) = \frac{1}{|S|} \sum_{\alpha \in S} \prod_{i \in V} g(m_i^\alpha). \end{aligned}$$

□

Lemma 8. (*Overlaps distribution of η_M*) Given a magnetization matrix M having N rows, $n > N$ columns with $n/N \in \mathbb{N}$, let η_M be the associated measure and let

$$P_M(q) = \langle \mathbb{I}_{\{q-q(\sigma, \sigma')\}} \rangle_{\eta_M \otimes \eta_M} \quad (3.23)$$

be the overlap distribution between two independent samples σ and σ' distributed according to η_M . Then

$$P_M(q) = \frac{1}{n^2} \sum_{\alpha=1}^n \sum_{\alpha'=1}^n \delta(q - q(m^\alpha, m^{\alpha'})), \quad (3.24)$$

supported by the discrete set

$$Q = \{q(m^\alpha, m^{\alpha'}) : \alpha, \alpha' \in S\}, \quad (3.25)$$

of normalized scalar products between the columns of M ,

$$q(m^\alpha, m^{\alpha'}) = \frac{1}{N} \sum_{i=1}^N m_i^\alpha m_i^{\alpha'}. \quad (3.26)$$

Proof. Obvious from definition of η_M . The support of η_M is the set of columns of M by construction, all equally weighted with probability $1/n$. □

Let us end this section with a result that should help in understanding the connection of the kernel representation with the Replica Theory. In particular, we show how to encode the kernel into a square matrix.

Lemma 9. Given some M having N rows, $n > N$ columns with $n/N \in \mathbb{N}$, and the generalized N -point function of Eq. (3.21), there is a squared magnetization matrix M_{sq} with $|V| = |S| = n$, defined as follows

$$M_{sq} = \{\tau_j^\alpha \in \Omega : 1 \leq j \leq n, 1 \leq \alpha \leq n\}, \quad (3.27)$$

with associated atomic distribution

$$\eta_{M_{sq}}(\tau) = \frac{1}{n} \sum_{\alpha \in S} \prod_{j=1}^n \left(\frac{1+\tau_j^\alpha \tau_j}{2} \right), \quad (3.28)$$

such that the following bounds hold

$$\langle \prod_{j=1}^n g(\tau_j)^{1/n} \rangle_{\eta_{M_{sq}}} \leq \langle \prod_{i=1}^N g(\sigma_i) \rangle_{\eta_M}^{1/N} \leq \langle \prod_{j=1}^n g(\tau_j) \rangle_{\eta_{M_{sq}}}^{1/n}. \quad (3.29)$$

Proof. Let partition the V set into a number N of subsets V_i of equal sizes $|V_i| = n/N \in \mathbb{N}$. Then, let M_{sq} be as follows

$$\tau_j^\alpha = m_i^\alpha, \quad \forall j \in V_i. \quad (3.30)$$

The probability distribution associated to M_{sq} is

$$\eta_{M_{sq}}(\tau) = \frac{1}{n} \sum_{\alpha \in S} \prod_{j=1}^n \left(\frac{1+\tau_j^\alpha}{2} \tau_j \right) = \frac{1}{n} \sum_{\alpha \in S} \prod_{i \in V} \prod_{j \in V_i} \left(\frac{1+m_i^\alpha}{2} \tau_j \right). \quad (3.31)$$

Now let $\boldsymbol{\tau}$ be a random spin vector of n spins

$$\Omega^n \ni \boldsymbol{\tau} \sim \eta_{M_{sq}} \quad (3.32)$$

distributed according to $\eta_{M_{sq}}$. It is easy to verify that

$$\left\langle \prod_{i=1}^N g(\boldsymbol{\sigma}_i) \right\rangle_{\eta_M} = \left\langle \prod_{i=1}^N \prod_{j \in V_i} g(\boldsymbol{\sigma}_i)^{N/n} \right\rangle_{\eta_M} = \left\langle \prod_{j=1}^n g(\boldsymbol{\tau}_j)^{N/n} \right\rangle_{\eta_{M_{sq}}}, \quad (3.33)$$

and since $N/n < 1$ by Jensen inequality we find

$$\left\langle \prod_{j=1}^n g(\boldsymbol{\tau}_j)^{1/n} \right\rangle_{\eta_{M_{sq}}} \leq \left\langle \prod_{j=1}^n g(\boldsymbol{\tau}_j)^{N/n} \right\rangle_{\eta_{M_{sq}}}^{1/N} \leq \left\langle \prod_{j=1}^n g(\boldsymbol{\tau}_j) \right\rangle_{\eta_{M_{sq}}}^{1/n} \quad (3.34)$$

which proves our claim. \square

After this the meaning of our Microscopic Pure States should be a little bit more clear. In fact $\eta_{M_{sq}}$ can be interpreted as the probability distribution of n/N correlated replicas of the N -spin system described by η_M , and the Microscopic Pure States of η_M are then the same of $\eta_{M_{sq}}$ of the replicated system.

We do not further deepen the connection with the Replica Theory as many simplifications will come from working with the rectangular matrix M of the actual system. We hope to properly explore this argument in a dedicated paper.

4 Random distributions

At this point we should open a parenthesis on how to deal with random probability measures, since this is one of the main technical issues that comes with the approximation of Theorem 6. Let first spot that if we work with kernels the representation is almost trivial.

Let $\boldsymbol{\mu} \sim \phi$ be a random probability measure distributed according to

$$\phi : \mathcal{P}(\Omega^N) \rightarrow [0, 1] \quad (4.1)$$

Since there is a correspondence between measures and kernels it suffices to take $\hat{M} = \hat{M}_\mu$ so that we can define the associated kernel distribution

$$\psi(\hat{M}) = \int_{\mu \in \mathcal{P}(\Omega^N)} \phi(d\mu) \delta(\hat{M} - \hat{M}_\mu) \quad (4.2)$$

and denote by \mathcal{M} the kernel space. Then by construction holds that

$$\hat{M} \sim \psi : \mathcal{M} \rightarrow [0, 1]. \quad (4.3)$$

is an exact encoding of the random distribution μ . For example, let us consider some bounded function $f : \Omega^N \rightarrow [0, 1]$, and suppose that we are interested in computing the average of $\log \langle f(\sigma) \rangle_\mu$ respect to the distribution ϕ , then from Lemma 3,

$$\mu(\sigma) = \int_{y \in [0,1]} dy \prod_{k=1}^N \left(\frac{1 + \hat{M}_\mu(\hat{x}_k, y) \sigma_k}{2} \right), \quad (4.4)$$

plus a trivial change of measure one easily finds the formula

$$\mathbb{E}_\phi \log \langle f(\sigma) \rangle_\mu = \mathbb{E}_\psi \log \int_{y \in [0,1]} dy \sum_{\sigma \in \Omega^N} f(\sigma) \prod_{k=1}^N \left(\frac{1 + \hat{M}(\hat{x}_k, y) \sigma_k}{2} \right). \quad (4.5)$$

Unfortunately, if we try to obtain the same relation using matrices instead of kernels we encounter some technical problems.

Although the approximation argument of Theorem 6 ensures that at least the annealed average $\bar{\mu}(\sigma) = \mathbb{E}_\phi \mu(\sigma)$ can be encoded into a deterministic matrix by choosing $|S| = n$ larger than some threshold depending on to the number of atoms (essentially because

$$\bar{\mu}(\sigma) = \sum_{a=1}^{2^N} \mathbb{E}_\phi \mu(\sigma^a) \prod_{k=1}^N \left(\frac{1 + \sigma_k^a \sigma_k}{2} \right) = \sum_{a=1}^{2^N} \bar{\mu}(\sigma^a) \prod_{k=1}^N \left(\frac{1 + \sigma_k^a \sigma_k}{2} \right) \quad (4.6)$$

is still an atomic distribution) to deal with quenched averages we would rather like to approximate the samples μ with a collection of matrices $M(\mu)$ having the same number of rows and columns for all μ and such that

$$M(\mu) \approx \hat{M}_\mu, \forall \mu \in \mathcal{P}(\Omega^N). \quad (4.7)$$

It is clear that to extend our approximation in this sense we would need to find a common surgery over \hat{S} that allows to simultaneously approximate all the kernels \hat{M}_μ . But our crude bounds for the sizes of the irregular sets $|\bar{S}_0|$ grow with the number of atoms, that may be infinite even for finite N if ϕ is continuously supported.

This his is exactly the kind of technical problems for which we would need the Szemerédi Lemma. By the way, we can still approximate the empirical average over any subset of measurements

$$\mathcal{P}(\Omega^N) \supset \mathcal{X} = \{\mu_t \in \mathcal{P}(\Omega^N) : 1 \leq t \leq T\} \quad (4.8)$$

if we keep T large but finite. Let define the empirical distribution

$$\phi_{\mathcal{X}}(\mu) = \frac{1}{\Gamma_{\mathcal{X}}(\phi)} \phi(\mu) \mathbb{I}_{\{\mu \in \mathcal{X}\}} \quad (4.9)$$

where the normalization $\Gamma_{\mathcal{X}}(\phi)$ is the partition function

$$\Gamma_{\mathcal{X}}(\phi) = \int_{\mu \in \mathcal{P}(\Omega^N)} \phi(d\mu) \mathbb{I}_{\{\mu \in \mathcal{X}\}} = \sum_{t=1}^T \phi(\mu_t). \quad (4.10)$$

This distribution has T atoms by definition, each one being in itself a distribution of at most 2^N atoms. This gives a total of $2^N T$ atoms that clearly allows for a common surgery if n is taken large enough.

Then, if we accept to deal with $\phi_{\mathcal{X}}$ instead of ϕ (which is realized in virtually any experimental setting) we can always find a sequence of matrices of N rows and n columns

$$\mathcal{M}(\mathcal{X}) = \{M(\mu) : \mu \in \mathcal{X}\} \quad (4.11)$$

such that each $M(\mu)$ approximates the corresponding \hat{M}_{μ} in the sense of Lemma 6 simultaneously for all $\mu \in \mathcal{X}$, provided that the total number of columns is $n > 2^{N+1}T/\epsilon$. Then we can define the atomic distribution

$$\psi_{\mathcal{X}}(M) = \frac{1}{T} \sum_{\mu \in \mathcal{X}} \phi_{\mathcal{X}}(\mu) \mathbb{I}_{\{M=M(\mu)\}}, \quad (4.12)$$

where $M(\mu) \approx \hat{M}_{\mu}$ for each $\mu \in \mathcal{X}$, and introduce the notation

$$\mathbf{M} \overset{\mathcal{X}}{\approx} \hat{M}_{\mu} \quad (4.13)$$

to indicate that \mathbf{M} approximates \hat{M}_{μ} in the sense that $\mathbf{M} \sim \psi_{\mathcal{X}}$. We remark once again that this technical issue can be completely overcome by following the approximation argument of [9, 10], which uses a probabilistic version of the Szemerédi Lemma.

5 Pure States

It is now time to introduce the main mathematical objects we will deal with. Hereafter we work only with η_M , automatically assuming that $M \approx \hat{M}_{\mu}$. We do this because matrices greatly simplify the presentation of these ideas, but we also remark from now that everything we are going to define and prove in this section has its natural counterpart in a continuous kernel formulation.

Essentially, we will introduce a sequence of refinements for the sets S and then use this construction to define the Pure States, and a Martingale representation for η_M . The first definition that we need is a tree indexing to control the refinements.

Definition 10. (*Tree Indexing for S*) Let S be as in Definition 5, let L be a finite integer and let

$$s = \{s_{\ell} \in \mathbb{N} : 0 \leq \ell \leq L, \prod_{0 \leq \ell \leq L} s_{\ell} = n\}. \quad (5.1)$$

be a collection of $L + 1$ integer parameters such that their product is n . Then we define the following tree index

$$\alpha_1 \alpha_2 \dots \alpha_{L+1} \in \bigotimes_{\ell'=0}^{\ell} \{1, 2, \dots, s_{\ell'}\}, \quad (5.2)$$

with $0 \leq \ell \leq L$ and each subindex α_{ℓ} running from 1 to $s_{\ell-1}$. The level $\ell = 0$ is the first generation of nodes originating from the root. The root has s_0 children, indexed by the variables $1 \leq \alpha_1 \leq s_0$. All the other indexes α_{ℓ} , $\ell \geq 1$ up to $\ell = L$ label the layers of internal nodes $1 \leq \alpha_{\ell} \leq s_{\ell-1}$. The last level $\ell = L + 1$ represents the external nodes α_{L+1} (also called leaves)..

We can now use the above indexing to describe the main definition of this section, ie a partition structure for the sets S . We anticipate that the following construction will be central to our definition of the Pure States

Definition 11. (Filtration of S). Let define the subsets

$$S_{\alpha_1 \dots \alpha_{L+1}} = \{\alpha(\alpha_1 \dots \alpha_{L+1})\} \quad (5.3)$$

each composed by only one element of S mapped onto the Tree Index of Definition 10 by some bijection θ_S . Then we call filtration of S the sequence of refinements

$$S = \{S_{\alpha_1 \dots \alpha_{\ell}} : 1 \leq \alpha_{\ell} \leq s_{\ell-1}, 0 \leq \ell \leq L + 1\}, \quad (5.4)$$

obtained from joining the subsets $S_{\alpha_1 \dots \alpha_{\ell}} \subseteq S$ from the last layer $S_{\alpha_1 \dots \alpha_{L+1}}$ via the recursive relations

$$S_{\alpha_1 \dots \alpha_{\ell}} = \bigcup_{\alpha_{\ell+1}=1}^{s_{\ell}} S_{\alpha_1 \dots \alpha_{\ell+1}}, \quad (5.5)$$

down to the root level $\ell = 0$. The root level

$$S = \bigcup_{\alpha_1=1}^{s_0} \dots \bigcup_{\alpha_{L+1}=1}^{s_L} S_{\alpha_1 \dots \alpha_{L+1}}, \quad (5.6)$$

is associated to the set S itself.

Notice that by definition there is a bijection between $\alpha \in S$ and the leaves indexes $\alpha_1 \dots \alpha_{L+1}$ that establish which state is placed in which set

$$\theta_S : \bigotimes_{\ell'=1}^{L+1} \{1, \dots, s_{\ell'}\} \rightarrow \{1, \dots, n\}. \quad (5.7)$$

This is a free parameter of the construction, and can be eventually tuned for our aims. In fact, the filtration above is completely identified by the set of parameters $(L, s, \theta_S) \in \mathcal{A}$. For notation convenience we will write

$$\alpha(\alpha_1 \dots \alpha_{L+1}) = \alpha_1 \dots \alpha_{L+1}(\alpha), \quad (5.8)$$

to explicitly indicate the map θ_S , so that it clearly exploits the correspondence from $\alpha \in S$ to $\alpha_1 \dots \alpha_{L+1}$. Analogously, we write $\alpha_1 \dots \alpha_\ell(\alpha)$ to indicate the first ℓ indexes of $\alpha_1 \dots \alpha_{L+1}(\alpha)$. By construction also holds that

$$S_{\alpha_1 \dots \alpha_\ell} \subseteq S_{\alpha_1 \dots \alpha_{\ell-1}} \quad (5.9)$$

for all $1 \leq \ell \leq L+1$, and we can eventually rewrite the sets in the equivalent form

$$S_{\alpha_1 \dots \alpha_\ell} = \bigcup_{\alpha_{\ell+1}=1}^{s_\ell} \dots \bigcup_{\alpha_{L+1}=1}^{s_L} S_{\alpha_1 \dots \alpha_{L+1}}, \quad (5.10)$$

Their sizes are $n_L = |S_{\alpha_1 \dots \alpha_{L+1}}| = 1$ for the last layer $\ell = L+1$ and

$$n_\ell = |S_{\alpha_1 \dots \alpha_\ell}| = \prod_{\ell'=\ell}^L s_{\ell'}, \quad (5.11)$$

for all the other layers, up to $S = n$ for the root level $\ell = 0$.

We can already identify this construction with a filtration for the subset algebra of S , uniquely identified through the parameters (L, s, θ_S) . It only remains to introduce the kernel variables.

Definition 12. (*Magnetization Averages*) Let m^α , $\alpha \in S$, be the columns of the Magnetization Matrix M and let \mathcal{S} be a filtration of parameters (L, s, θ_S) . We introduce a notation for the averages of the field components on the states $S_{\alpha_1 \dots \alpha_\ell}$. The last layer $\ell = L+1$ is associated to the external nodes

$$\Omega^V \ni m^{\alpha_1 \dots \alpha_{L+1}} = m^{\alpha(\alpha_1 \dots \alpha_{L+1})}. \quad (5.12)$$

For all the other layers we define

$$[0, 1]^V \ni m^{\alpha_1 \dots \alpha_\ell} = \frac{1}{|S_{\alpha_1 \dots \alpha_\ell}|} \sum_{\alpha \in S_{\alpha_1 \dots \alpha_\ell}} m^\alpha, \quad (5.13)$$

down to the root level $\ell = 0$, for which we drop the tree index and use the simple notation

$$m = \frac{1}{n} \sum_{\alpha \in S} m^\alpha. \quad (5.14)$$

An equivalent definition that uses the tree indexing only can be

$$m^{\alpha_1 \dots \alpha_\ell} = \frac{1}{s_\ell} \sum_{\alpha_{\ell+1}=1}^{s_\ell} \dots \frac{1}{s_L} \sum_{\alpha_{L+1}=1}^{s_L} m^{\alpha_1 \dots \alpha_{L+1}}. \quad (5.15)$$

and it is important to notice that all these kernel quantities are implicitly dependent from the parameters (L, s, θ_S) of the filtration from which are defined. For any $\alpha \in S$ and a given level ℓ we indicate with

$$m^{\alpha_1 \dots \alpha_\ell(\alpha)} = \frac{1}{|S_{\alpha_1 \dots \alpha_\ell(\alpha)}|} \sum_{\alpha' \in S_{\alpha_1 \dots \alpha_\ell(\alpha)}} m^{\alpha'} \quad (5.16)$$

the averaged magnetization vector conditioned to the state $S_{\alpha_1 \dots \alpha_\ell(\alpha)}$ to which α belongs.

Definition 13. (*Magnetization Increments*) Let $0 \leq \ell \leq L + 1$, then starting from the Magnetization Averages we define the parameters

$$\delta m^{\alpha_1 \dots \alpha_\ell} = \frac{1}{|S_{\alpha_1 \dots \alpha_\ell}|} \sum_{\alpha \in S_{\alpha_1 \dots \alpha_\ell}} m^\alpha - \frac{1}{|S_{\alpha_1 \dots \alpha_{\ell-1}}|} \sum_{\alpha' \in S_{\alpha_1 \dots \alpha_{\ell-1}}} m^{\alpha'} \quad (5.17)$$

that indicate the increment of the averaged magnetization from $m^{\alpha_1 \dots \alpha_{\ell-1}}$ of the layer $\ell - 1$ to $m^{\alpha_1 \dots \alpha_\ell}$ of the layer ℓ . For the the root level $\ell = 0$ we simply write $\delta m = m$, while the first generation $\ell = 1$ is $\delta m^{\alpha_1} = m^{\alpha_1} - m$. The other levels are obtained iterating

$$\delta m^{\alpha_1 \dots \alpha_\ell} = m^{\alpha_1 \dots \alpha_\ell} - m^{\alpha_1 \dots \alpha_{\ell-1}} = m^{\alpha_1 \dots \alpha_\ell} - \frac{1}{s_{\ell-1}} \sum_{\alpha'_\ell=1}^{s_{\ell-1}} m^{\alpha_1 \dots \alpha_{\ell-1} \alpha'_\ell} \quad (5.18)$$

down to the last level $\delta m^{\alpha_1 \dots \alpha_{L+1}} = m^{\alpha_1 \dots \alpha_{L+1}} - m^{\alpha_1 \dots \alpha_L}$ associated to the external nodes.

We remark that, by construction, the average of any $\delta m^{\alpha_1 \dots \alpha_\ell}$ respect to the index α_ℓ is zero

$$\frac{1}{s_{\ell-1}} \sum_{\alpha_\ell=1}^{s_{\ell-1}} \delta m^{\alpha_1 \dots \alpha_\ell} = 0 \quad (5.19)$$

and any averaged magnetization $m^{\alpha_1 \dots \alpha_\ell}$ can be reconstructed from the increments according to the formula

$$m^{\alpha_1 \dots \alpha_\ell} = m + \delta m^{\alpha_1} + \delta m^{\alpha_1 \alpha_2} + \dots + \delta m^{\alpha_1 \dots \alpha_\ell} = \sum_{\ell'=0}^{\ell} \delta m^{\alpha_1 \dots \alpha_{\ell'}} \quad (5.20)$$

Clearly also the above quantities are implicitly dependent from the filtration parameters (L, s, θ_S) . For any $\alpha \in S$ there is a correspondence with the magnetizations of S ,

$$m^\alpha = m^{\alpha_1 \dots \alpha_{L+1}(\alpha)} = \sum_{\ell=0}^{L+1} \delta m^{\alpha_1 \dots \alpha_\ell(\alpha)} \in \Omega \quad (5.21)$$

while we can write

$$m^{\alpha_1 \dots \alpha_\ell(\alpha)} = \sum_{\ell'=0}^{\ell} \delta m^{\alpha_1 \dots \alpha_{\ell'}}(\alpha) \quad (5.22)$$

to indicate the average magnetizations of the state $S_{\alpha_1 \dots \alpha_\ell(\alpha)}$ to which the index α belongs. We can now give our interpretation of the Pure States as intended in the RSB ansatz

Definition 14. (*Pure States*) Let $0 \leq \ell \leq L+1$ and let M be a magnetization matrix $M : S \otimes V \rightarrow \Omega$ of N rows and n columns, then we define the sequence of magnetization matrices

$$M_\ell : S \otimes V \rightarrow [0, 1] \quad (5.23)$$

where the columns are the average magnetization vectors

$$M_\ell = \{m^{\alpha_1 \dots \alpha_\ell(\alpha)} \in [0, 1]^V : 1 \leq \alpha \leq n\}, \quad (5.24)$$

given in Definition 12 and $M_{L+1} = M$. Then, for each $1 \leq \ell \leq L$ we call Pure States of η_M the average magnetization vectors associated to the internal nodes of the filtration. We also introduce the matrices sequence

$$\delta M_\ell : S \otimes V \rightarrow [0, 1] \quad (5.25)$$

such that the matrix elements of

$$\delta M_\ell = \{\delta m^{\alpha_1 \dots \alpha_\ell(\alpha)} \in [0, 1]^V : 1 \leq \alpha \leq n\}, \quad (5.26)$$

are the increments of Definition 13, and

$$M = M_{L+1} = \sum_{\ell=0}^{L+1} \delta M_\ell. \quad (5.27)$$

Although the invariance of the sum order guarantees that the observables don't change under the action of the map θ_S of Eq. (5.7), when we switch configurations between the states this can change both the magnetization averages and the magnetization increments. In essence, we interpret the Pure States as hidden parameters that we can use to control a Lebesgue approximation of a Riemann integrable kernel (the Magnetization kernel of a finite volume spin system is always Riemann integrable⁷).

The above Definition 14 is intended to give a rigorous formulation for the *finite volume pure states* argued by Marinari et al. in [5], except for the ultrametricity assumption that we discuss in the next section. Since we are working with finite systems here we avoid a comparison with the usual DLR states of the infinite systems⁸.

⁷In cut distance approximation we would do the opposite. If we use a Szemerédi Partition to construct a stepfunction approximation of the kernel in the thermodynamic limit, then we are performing a Riemann approximation of a Lebesgue integrable kernel.

⁸Concerning the thermodynamic limit of our finite volume pure states, notice that the support Ω^N of a finite N -spin system can be represented by the vertex of an N -dimensional hypercube. In the continuous kernel representation we embed the hypercube into a compact space, where it is possible to define a convex envelope for it, following the prescriptions of the Krein-Milman theorem [8].

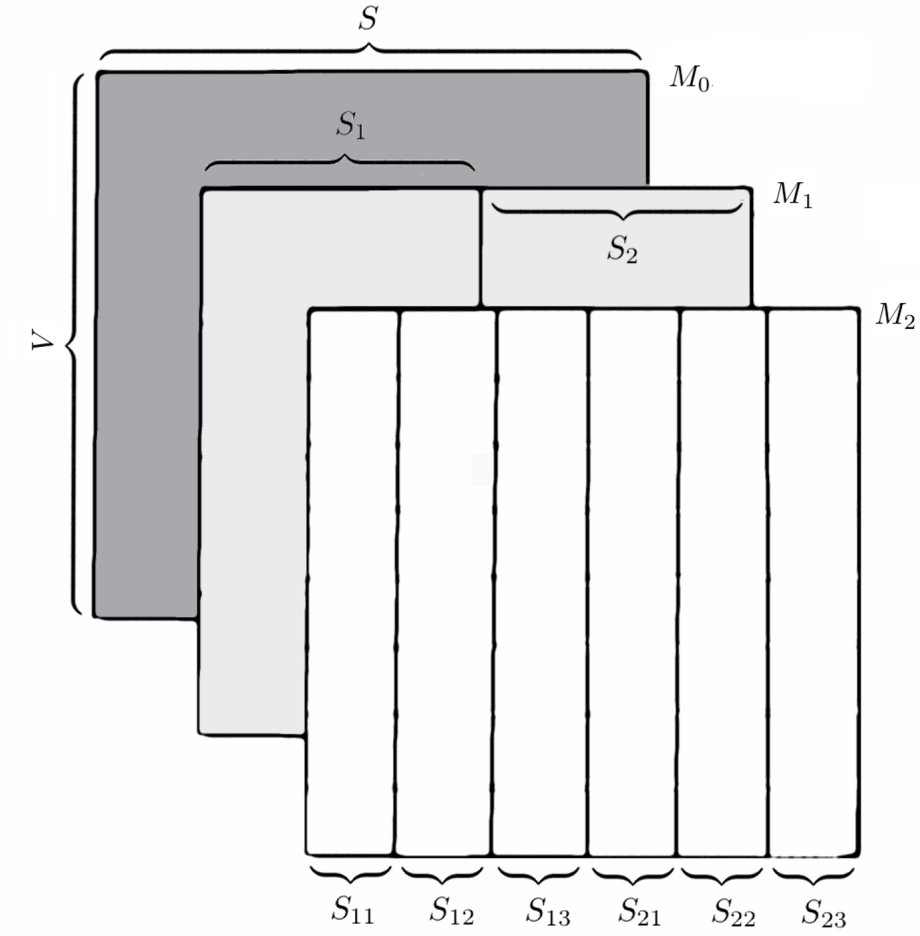


Figure 5.1: Pure states representation of Definition (14) for a non-random 2-RSB system ($L = 2$) and size parameters $n_0 = 2$ and $n_1 = 3$. The vertical lines highlight the pure states of each layer M_0 , M_1 and M_2 of the kernel M . The last kernel $M_3 = M$ not shown for better visualization. See also Figure 6.1.

Notice that the refinement sequence \mathcal{S} is indeed a filtered algebra and can be used to define martingale processes. In particular, it is possible to associate this construction with a Doob martingale (see below). This property has been first observed by Guerra et al. in [16, 17, 18], where it is proven using different techniques that the pure states of the RSB ansatz admit a martingale representation.

The Doob martingale [15, 19] is a martingale that approximates any random variable according to a given filtration. Given a random set

$$\mathcal{J}^T \ni \mathbf{X} = \{\mathbf{X}_t \in \mathcal{J} : 0 \leq t \leq T\} \sim \xi : \mathcal{J}^T \rightarrow [0, 1] \quad (5.28)$$

and some $f : \mathcal{J}^T \rightarrow \mathbb{R}$, then we call Doob Martingale \mathbf{f} of initial condition

$$f_0 = \mathbb{E}_\xi (f(\mathbf{X})) \quad (5.29)$$

stopped at time T the stochastic process

$$\mathbf{f}_t = \mathbb{E}_\xi (f(\mathbf{X}) | \mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_t), \quad (5.30)$$

where the average is taken on last $T - t$ variables only. We can use the filtration \mathcal{S} to construct a Doob martingale having η_M as initial condition.

First we need to introduce the distributions of the increments, and we remark that the next definition will be instrumental for the identification of the Pure States

Definition 15. (*Distribution of the increments*) *Let consider a magnetization matrix M of N rows and n columns. Let*

$$p^{\alpha_1 \dots \alpha_\ell} : [0, 1]^V \rightarrow [0, 1]; \quad 0 \leq \ell \leq L \quad (5.31)$$

be a sequence of atomic probability distributions defined as

$$p^{\alpha_1 \dots \alpha_\ell}(\phi) = \frac{1}{s_\ell} \sum_{\alpha_{\ell+1}=1}^{s_\ell} \prod_{i=1}^N \delta(\phi_i - \delta m_i^{\alpha_1 \dots \alpha_{\ell+1}}) \quad (5.32)$$

where the parameters $\delta m^{\alpha_1 \dots \alpha_\ell}$ are the increments of magnetization in Definition 13 and ϕ the associated dummy variable

$$[0, 1]^V \ni \phi = \{\phi_i \in [0, 1] : i \in V\}. \quad (5.33)$$

Then, the Doob Martingale stopped at time $L+1$ having η_M as initial condition is given by the following theorem

Theorem 16. (*Doob Martingale approximation of η_M*) *For any choice of the filtration parameters (L, s, θ_S) the distribution η_M can be represented as follows*

$$\eta_M(\sigma) = \mathbb{E}_p \mathbb{E}_{p^{\alpha_1}} \dots \mathbb{E}_{p^{\alpha_1 \dots \alpha_L}} \boldsymbol{\eta}^{\alpha_1 \dots \alpha_{L+1}}(\sigma), \quad (5.34)$$

with $\boldsymbol{\eta}^{\alpha_1 \dots \alpha_{L+1}}$ random probability measure

$$\boldsymbol{\eta}^{\alpha_1 \dots \alpha_{L+1}}(\sigma) = 2^{-N} \prod_{i=1}^N \left(1 + \sum_{0 \leq \ell \leq L+1} \delta \mathbf{m}_i^{\alpha_1 \dots \alpha_\ell} \sigma_i \right) \quad (5.35)$$

dependent on the random vectors $\delta \mathbf{m}^{\alpha_1 \dots \alpha_\ell} \sim p^{\alpha_1 \dots \alpha_\ell}$ given in Definition 15. The sequence of random probability measures

$$\boldsymbol{\eta}^{\alpha_1 \dots \alpha_\ell}(\sigma) : [0, 1]^V \rightarrow [0, 1]; \quad 0 \leq \ell \leq L+1 \quad (5.36)$$

defined through the recursion

$$\boldsymbol{\eta}^{\alpha_1 \dots \alpha_\ell}(\sigma) = \mathbb{E}_{p^{\alpha_1 \dots \alpha_\ell}} \boldsymbol{\eta}^{\alpha_1 \dots \alpha_{\ell+1}}(\sigma) \quad (5.37)$$

is the Doob martingale of $\boldsymbol{\eta}^{\alpha_1 \dots \alpha_{L+1}}$ stopped at $L+1$ with initial condition

$$\eta_M = \mathbb{E}_p \boldsymbol{\eta}^{\alpha_1}(\sigma). \quad (5.38)$$

Proof. We need to prove the first equation only, being the rest simply a definition for the Doob Martingale (see [15], also Section 12.11 of [19] for the related Doob Representation). The claim becomes evident if we rewrite the definition

$$\eta_M(\sigma) = \frac{1}{|S|} \sum_{\alpha \in S} \prod_{i \in V} \left(\frac{1+m_i^\alpha \sigma_i}{2} \right) \quad (5.39)$$

using the tree indexing and the kernel variables above. We can write m_i^α in term of the increments

$$m_i^\alpha \rightarrow \sum_{\ell=0}^{L+1} \delta m_i^{\alpha_1 \dots \alpha_\ell} \quad (5.40)$$

and change to the appropriate sum operator

$$\frac{1}{|S|} \sum_{i \in S} \rightarrow \frac{1}{|S|} \sum_{\alpha_1=1}^{s_0} \dots \sum_{\alpha_{L+1}=1}^{s_L} \rightarrow \frac{1}{s_0} \sum_{\alpha_1=1}^{s_0} \dots \frac{1}{s_L} \sum_{\alpha_{L+1}=1}^{s_L}, \quad (5.41)$$

up to find the iteration

$$\eta_M(\sigma) = \frac{1}{s_0} \sum_{\alpha_1=1}^{s_0} \dots \frac{1}{s_L} \sum_{\alpha_{L+1}=1}^{s_L} \boldsymbol{\eta}^{\alpha_1 \dots \alpha_{L+1}}(\sigma), \quad (5.42)$$

where the initial condition is

$$\boldsymbol{\eta}^{\alpha_1 \dots \alpha_{L+1}}(\sigma) = 2^{-N} \prod_{i=1}^N \left(1 + \sum_{0 \leq \ell \leq L+1} \delta m_i^{\alpha_1 \dots \alpha_\ell} \sigma_i \right). \quad (5.43)$$

Applying the definition of the sequence $p^{\alpha_1 \dots \alpha_\ell}$ we also change

$$\frac{1}{s_\ell} \sum_{\alpha_{\ell+1}=1}^{s_\ell} \rightarrow \mathbb{E}_{p^{\alpha_1 \dots \alpha_\ell}} \quad (5.44)$$

and the increments $\delta m^{\alpha_1 \dots \alpha_\ell}$ with their random version $\delta \mathbf{m}^{\alpha_1 \dots \alpha_\ell}$. Doing this we automatically obtain

$$\eta^{\alpha_1 \dots \alpha_{L+1}}(\sigma) \rightarrow \boldsymbol{\eta}^{\alpha_1 \dots \alpha_{L+1}}(\sigma), \quad (5.45)$$

from which the claim follows. \square

Before discussing the Parisi Ansatz it worth notice that although the above martingale is here described as a forward process, the construction of its filtration has been obtained by joining the set backward, from the finer level that corresponds to the stopping time $\ell = L + 1$ to the initial condition $\ell = 0$. Then, from the probabilistic point of view this construction should rather be viewed as a backward process, and if we want to consider it a forward process we should keep in mind that the final conditions $m^{\alpha_1 \dots \alpha_{L+1}(\alpha)}$ are the microscopic states of M , and then must always be binary vectors for each state.

6 Parisi Ansatz

The martingale representation of Eq. (5.34) holds for any choice of (L, s, θ_S) by construction, then we have a lot of freedom in tuning the parameters to elaborate further.

Indeed, we believe that this is the point in which the ultrametricity assumption of the Parisi Ansatz find its mathematical ratio, ie in the possibility of tuning the free parameters (L, s, θ_S) in such a way that the increments $\delta m^{\alpha_1 \dots \alpha_\ell}$ can be treated as i.i.d. random variables (at least on average, if we deal with a random magnetization matrix \mathbf{M}). To make this argument more precise we can proceed as follows. First, consider the overlap support of η_M ,

$$Q = \{q(m^\alpha, m^{\alpha'}) : \alpha, \alpha' \in S\}, \quad (6.1)$$

then by Lemma 8 we have that

$$q(m^\alpha, m^{\alpha'}) = \frac{1}{N} m^\alpha m^{\alpha'} = \frac{1}{N} \sum_{i \in V} m_i^\alpha m_i^{\alpha'}, \quad (6.2)$$

Let see what happen when we rewrite it using the following tree indexing

$$\alpha = \alpha_1 \dots \alpha_{L+1}(\alpha) = \alpha_1 \dots \alpha_{L+1} = \underline{\alpha}_{L+1}, \quad (6.3)$$

$$\alpha_1 \dots \alpha_\ell(\alpha) = \alpha_1 \dots \alpha_\ell = \underline{\alpha}_\ell. \quad (6.4)$$

By definition holds that

$$\begin{aligned} Nq(m^\alpha, m^{\alpha'}) &= q(m^{\alpha_1 \dots \alpha_{L+1}}, m^{\alpha'_1 \dots \alpha'_{L+1}}) = \\ &= m^{\alpha_1 \dots \alpha_{L+1}} m^{\alpha'_1 \dots \alpha'_{L+1}} = \sum_{\ell=0}^{L+1} \sum_{\ell'=0}^{L+1} \delta m^{\alpha_1 \dots \alpha_\ell} \delta m^{\alpha'_1 \dots \alpha'_{\ell'}} \end{aligned}$$

and after some simple algebra we arrive to the expression

$$\begin{aligned}
& \sum_{\ell=0}^{L+1} \sum_{\ell'=0}^{L+1} \delta m^{\alpha_1 \dots \alpha_\ell} \delta m^{\alpha'_1 \dots \alpha'_{\ell'}} = \\
& = \sum_{\ell=0}^{L+1} \delta m^{\alpha_1 \dots \alpha_\ell} \delta m^{\alpha'_1 \dots \alpha'_{\ell}} + \sum_{\ell \neq \ell'} \delta m^{\alpha_1 \dots \alpha_\ell} \delta m^{\alpha'_1 \dots \alpha'_{\ell'}} = \\
& = \sum_{\ell=0}^{\ell^*(\alpha, \alpha')} |\delta m^{\alpha_1 \dots \alpha_\ell}|^2 + \sum_{\ell=\ell^*(\alpha, \alpha')+1}^{L+1} \delta m^{\alpha_1 \dots \alpha_\ell} \delta m^{\alpha'_1 \dots \alpha'_{\ell}} + \\
& \quad + \sum_{\ell \neq \ell'} \delta m^{\alpha_1 \dots \alpha_\ell} \delta m^{\alpha'_1 \dots \alpha'_{\ell'}} \quad (6.5)
\end{aligned}$$

where in the last equation we introduced the index

$$\ell^*(\alpha, \alpha') = \sup \{ \ell \in \mathbb{N} : \alpha_1 \dots \alpha_\ell = \alpha'_1 \dots \alpha'_{\ell} \} = \sum_{\ell=1}^{L+1} \prod_{\ell'=1}^{\ell} \mathbb{I}_{\{\alpha_{\ell'} = \alpha'_{\ell'}\}}. \quad (6.6)$$

In practice, we separated the off diagonal terms of the overlap matrix of the increments, that bring us to the following expression for the overlap support:

$$Ng(m^\alpha, m^{\alpha'}) = \sum_{\ell=0}^{\ell^*(\alpha, \alpha')} |\delta m^{\alpha_1 \dots \alpha_\ell}|^2 + R(\alpha, \alpha'), \quad (6.7)$$

where we collected the reminders into the estimator

$$\begin{aligned}
R(\alpha, \alpha') &= R_1(\alpha, \alpha') + R_2(\alpha, \alpha') = \\
&= \sum_{\ell=\ell^*(\alpha, \alpha')+1}^{L+1} \delta m^{\alpha_1 \dots \alpha_\ell} \delta m^{\alpha'_1 \dots \alpha'_{\ell}} + \sum_{\ell \neq \ell'} \delta m^{\alpha_1 \dots \alpha_\ell} \delta m^{\alpha'_1 \dots \alpha'_{\ell'}}. \quad (6.8)
\end{aligned}$$

It is clear from the above expression that if we could find a configuration of (L, s, θ_S) for which $R(\alpha, \alpha')$ is small this would be quite like having independence between the increments, and this would induce a hierarchical topology for the martingale evolution (although not necessarily ultrametric as we shall see in short).

Based on the above considerations, if we restrict our analysis to the case

$$\mathcal{L}_2 = \{ \mathbf{M} \sim \psi : \mathbb{E}_\psi \mathbf{m}^\alpha = 0, \mathbb{E}_\psi |\mathbf{m}^\alpha|^2 \leq C_{\max} N \} \quad (6.9)$$

of \mathcal{L}_2 -bounded random matrices with vanishing magnetizations we can give a straightforward sufficient condition for this property to hold on average

Lemma 17. (*Tree Topology Condition, TTC*). *Let consider some random magnetization matrix $\mathbf{M} \in \mathcal{L}_2$ of N rows and n columns. Then, let define the average overlap matrix of the increments,*

$$C_\psi(\underline{\alpha}_\ell, \underline{\alpha}'_{\ell'}) = \mathbb{E}_\psi \delta \mathbf{m}^{\alpha_1 \dots \alpha_\ell} \delta \mathbf{m}^{\alpha'_1 \dots \alpha'_{\ell'}} \quad (6.10)$$

computed from \mathbf{M} , and call \mathcal{A} the set of choices for the parameters (L, s, θ_S) . Given the (possibly empty) subset

$$\mathcal{A}_\psi^*(\epsilon) = \{(L, s, \theta_S) \in \mathcal{A} : \sup_{\alpha, \ell \neq \alpha', \ell'} |C_\psi(\underline{\alpha}_\ell, \underline{\alpha}'_{\ell'})| \leq \epsilon N / L^2\}, \quad (6.11)$$

then $\forall (L, s, \theta_S) \in \mathcal{A}_\psi^*(\epsilon)$ there is a collection of bounded real numbers

$$\tilde{\Gamma}(L, s) \ni \tilde{\gamma} = \{\tilde{\gamma}_{\alpha_1 \dots \alpha_\ell} \in [0, C_{\max}^{1/2}] : 1 \leq \alpha_\ell \leq s_{\ell-1}, 0 \leq \ell \leq L+1\} \quad (6.12)$$

such that the following condition holds for any pair $\underline{\alpha}_\ell, \underline{\alpha}'_{\ell'}$

$$|\mathbb{E}_\psi \delta \mathbf{m}^{\alpha_1 \dots \alpha_\ell} \delta \mathbf{m}^{\alpha'_1 \dots \alpha'_{\ell'}} - N \tilde{\gamma}_{\alpha_1 \dots \alpha_\ell} \mathbb{I}_{\{\alpha_1 \dots \alpha_\ell = \alpha'_1 \dots \alpha'_{\ell'}\}}| \leq 2\epsilon N. \quad (6.13)$$

We say that \mathbf{M} has an ϵ -tree topology if $\mathcal{A}_\psi^*(\epsilon)$ contains at least one element.

Proof. We only need to prove the existence of $\tilde{\gamma}$ when $(L, s, \theta_S) \in \mathcal{A}_\psi^*(\epsilon)$. Since by hypothesis hold

$$\sup_{\alpha, \ell \neq \alpha', \ell'} |C_\psi(\underline{\alpha}_\ell, \underline{\alpha}'_{\ell'})| \leq \epsilon N / L^2 \quad (6.14)$$

we can start from Eq. (6.8) to obtain the following trivial bound

$$\begin{aligned} \mathbb{E}_\psi \mathbf{R}(\alpha, \alpha') &\leq \sum_{\ell \geq \ell^*(\alpha, \alpha')} |C_\psi(\underline{\alpha}_\ell, \underline{\alpha}'_{\ell'})| + \sum_{\ell \neq \ell'} |C_\psi(\underline{\alpha}_\ell, \underline{\alpha}'_{\ell'})| \leq \\ &\leq 2L^2 \sup_{\alpha, \ell \neq \alpha', \ell'} |C_\psi(\underline{\alpha}_\ell, \underline{\alpha}'_{\ell'})| \leq 2\epsilon N \end{aligned} \quad (6.15)$$

and then define $\tilde{\gamma}$ from $\tilde{\gamma}_{\alpha_1 \dots \alpha_\ell}^2 = C_\psi(\underline{\alpha}_\ell, \underline{\alpha}_\ell) / N \in [0, C_{\max}]$. \square

Strictly speaking, the TTC above is equivalent to ask that the overlap matrix of the increments is almost diagonal, and seems to describe some kind of decorrelation process where the states $\alpha, \alpha', \alpha'', \dots$ evolves together up to some abstract time $\ell^*(\alpha, \alpha', \alpha'', \dots)$ at which they are separated by the filtration process. Essentially, each kernel column evolves like a Markov Process in ℓ , but the evolution of different states is coupled until they are in the same group, then their evolutions become independent.

We believe that this generalized approximation could find useful applications in studying models that exhibit a non self-averaging free energy density respect to the disorder. Notice also that many intriguing questions may come from the interpretation of the RSB parameter ℓ . Mathematically, we presented it as an abstract time for the evolution forward of the Doob martingale, but we wonder if some physical interpretation is possible in term of block renormalization, or other scaling operations of some kind.

To avoid distracting technicalities, in the present work we don't even try to establish for which ψ the subset $\mathcal{A}_\psi^*(\epsilon)$ is non-empty in the limit of infinitesimal ϵ , nor we attempt to apply these ideas to practical situations, but we remark that more refined (and useful, depending on the case) estimates can be developed starting from the concept of TTC above.

Anyway, it is important to understand that TTC alone does not imply ultrametricity. In fact, the averaged support of the overlap is given by

$$\mathbb{E}_\psi q(\mathbf{m}^\alpha, \mathbf{m}^{\alpha'}) = \sum_{\ell=0}^{\ell^*(\alpha, \alpha')} \tilde{\gamma}_{\alpha_1 \dots \alpha_\ell}^2. \quad (6.16)$$

and it is not necessarily ultrametric. Even if the dependence from $\ell^*(\alpha, \alpha')$ imply a hierarchical organization of the overlap support, we also need to assume equality in distribution of the increments within each layer ℓ , ie

$$\delta \mathbf{m}^{\alpha_1 \dots \alpha_\ell} \stackrel{d}{=} \delta \mathbf{m}^{\alpha'_1 \dots \alpha'_\ell}, \quad (6.17)$$

to get all $\tilde{\gamma}_{\alpha_1 \dots \alpha_\ell}$ equal to some γ_ℓ . We can give a simple formal condition for this property to hold, independent from the TTC above. We conjecture, but not prove here, that for $\mathbf{M} \in \mathcal{L}_2$ the State Equivalence described below should correspond to the Replica Equivalence of [20] and the Stochastic Stability introduced in [21]. In general, it seems much less exotic than TTC.

Definition 18. (*States Equivalence Condition, SEC*) Let γ be a collection of $L + 1$ bounded real parameters

$$\Gamma(L) \ni \gamma = \{\gamma_\ell \in [0, C_{\max}^{1/2}] : 0 \leq \ell \leq L + 1\}, \quad (6.18)$$

let $\Gamma(L)$ be the space of these parameters and let $\mathcal{A} \ni (L, s, \theta_S)$. For some random magnetization matrix $\mathbf{M} \in \mathcal{L}_2$ we define the index

$$\Delta(\alpha, \gamma_\ell) = C_\psi(\underline{\alpha}_\ell, \underline{\alpha}_\ell) - N\gamma_\ell^2. \quad (6.19)$$

computed respect to each layer ℓ of filtration. Then, we say that the Pure States of \mathbf{M} are ϵ -equivalent if the subset

$$\mathcal{A}_\psi^\dagger(\epsilon) = \bigcup_{\gamma \in \Gamma(L)} \{(L, s, \theta_S) \in \mathcal{A} : \sup_{\alpha, \ell} |\Delta(\alpha, \gamma_\ell)| \leq \epsilon N\} \quad (6.20)$$

has at least one element.

If TTC and SEC hold for some kernel \mathbf{M} , then the associated overlap distribution is ultrametric. That said, we can finally introduce the class of Parisi Kernels, that we expect to include the Magnetization Matrix associated the SK model.

Theorem 19. (*Parisi Kernel*) We call Parisi Type Matrix an array $\mathbf{M} \in \mathcal{L}_2$ such that for some choice of $(L, s, \theta_S) \in \mathcal{A}$ and $\gamma \in \Gamma(L)$ both TTC and SEC hold exactly,

$$\mathbb{E}_\psi \delta \mathbf{m}^{\alpha_1 \dots \alpha_\ell} \delta \mathbf{m}^{\alpha'_1 \dots \alpha'_{\ell'}} = N\gamma_\ell^2 \mathbb{I}_{\{\alpha_1 \dots \alpha_\ell = \alpha'_1 \dots \alpha'_{\ell'}\}}. \quad (6.21)$$

Then, the support of the overlap is ultrametric on average

$$\mathbb{E}_\psi q(\mathbf{m}^\alpha, \mathbf{m}^{\alpha'}) = \sum_{\ell=0}^{\ell^*(\alpha, \alpha')} \gamma_\ell^2. \quad (6.22)$$

We call Parisi Distribution $\eta_{\mathbf{M}}$ the probability distribution associated to \mathbf{M} .

Proof. It suffices to take the average of Eq.s (6.7), (6.8) and use the hypothesis in Eq. (6.21) to find $\mathbb{E}_\psi \mathbf{R}(\alpha, \alpha') = 0$, corresponding to TTC, and $\tilde{\gamma}_{\alpha_1 \dots \alpha_\ell} = \gamma_\ell$, that is the SEC. Then Eq. (6.22) follows. \square

Notice that if we also assume concentration of the random overlap support \mathbf{Q} on its averaged version

$$\mathbb{E}_\psi \mathbf{Q} = \{ \mathbb{E}_\psi q(\mathbf{m}^\alpha, \mathbf{m}^{\alpha'}) : \alpha, \alpha' \in S \}, \quad (6.23)$$

(in general, a property still dependent from the choice of ψ) we can compare with the usual expressions of the averaged overlap distribution from Replica Theory. By the following change of variable

$$\gamma_{\ell+1} = \sqrt{q_{\ell+1} - q_\ell}, \quad (6.24)$$

we obtain the expression

$$\mathbb{E}_\psi q(\mathbf{m}^\alpha, \mathbf{m}^{\alpha'}) = \sum_{\ell'=1}^{\ell^*(\alpha, \alpha')} (q_{\ell'} - q_{\ell'-1}) = q_{\ell^*(\alpha, \alpha')} \quad (6.25)$$

and by assuming concentration we get

$$\mathbb{E}_\psi P_M(q) = \frac{1}{n^2} \sum_{\alpha=1}^n \sum_{\alpha'=1}^n \delta(q - q_{\ell^*(\alpha, \alpha')}). \quad (6.26)$$

Concerning the sizes n_ℓ of the blocks in which the step function $\ell^*(\alpha, \alpha')$ is divided, as suggested by Lemma 9 and considerations afterward they must be proportional to the blocks $\lambda_{\ell-1}$ of the RSB ansatz in the Replica Formulation, then we find $\lambda_{\ell-1}/\lambda_\ell = s_\ell$.

To avoid long detours, we won't discuss the connection with the classical probabilistic Point Process description of [6], although it is not hard to prove that assuming concentration the random parameters of the above measure will be distributed according to a Ruelle Cascade, with rates of each layer corresponding to our filtration parameters s_ℓ .

We also avoid to give a detailed physical interpretation at this point, because for this purpose we will need to introduce a joint filtration for S and V , and construct a more refined martingale approximation to capture the factorization of the pure states of Lemma 2. By the way, notice that usual mean field theories approximate the target functional, for example the free energy density, by suppressing the correlations between the spin sites. The RSB ansatz seems instead a mean field theory that suppresses correlations between the states (in a rather nontrivial way if we take $L > 1$), and this can be interpreted as a suppression of time correlations between different quenched measurements.

In Figure 6.1 we show the Magnetization kernel of a non-random 2-RSB measure, where a measure preserving map has been applied also to V to highlight the factorization of the pure states. For each state m^α the stripes $V_k \subset V$ define collections of spin configurations

$$m_{V_k}^\alpha = \{ m_i^\alpha \in \Omega : i \in V_k \} \quad (6.27)$$

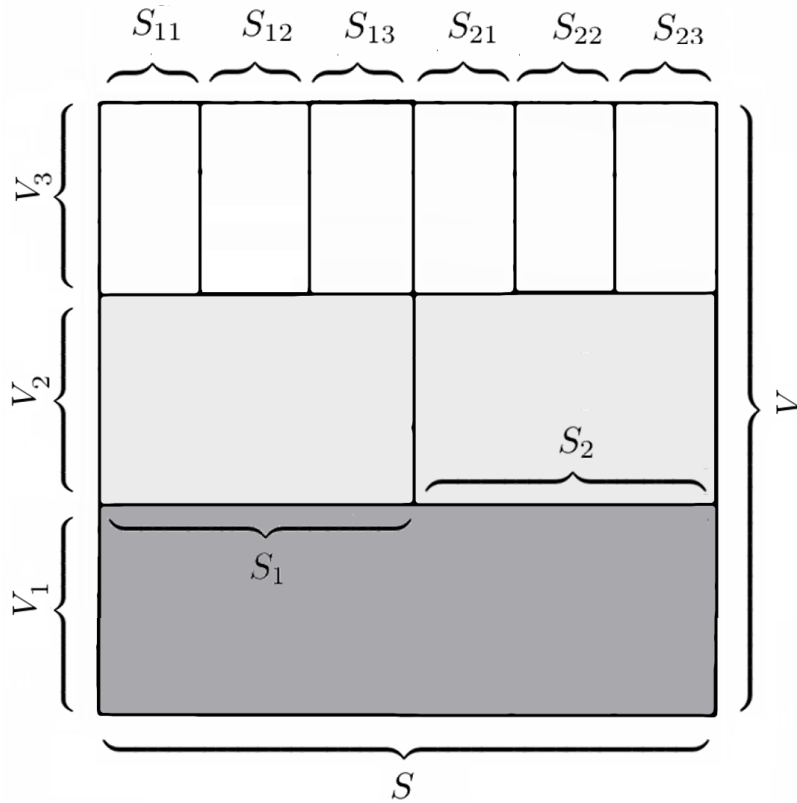


Figure 6.1: Magnetization kernel of a non-random 2-RSB measure. For each V_k the state components $m_{V_k}^\alpha$ of Eq.(6.27) are dependent inside the same state and orthogonal between different ones. Using an opportune measure preserving mapping of V to match with the Figure 5.1 it is possible to show that the above kernel satisfy both TTC and SEC.

that are dependent inside the same state and orthogonal between different states. As shown in Fig. 6.1, if $\alpha_1 \dots \alpha_{k-1} \neq \alpha'_1 \dots \alpha'_{k-1}$ and

$$\alpha \in S_{\alpha_1 \dots \alpha_{k-1}}, \alpha' \in S_{\alpha'_1 \dots \alpha'_{k-1}} \quad (6.28)$$

then $m_{V_k}^\alpha m_{V_k}^{\alpha'} = 0$, while if $\alpha, \alpha' \in S_{\alpha_1 \dots \alpha_{k-1}}$ then $m_{V_k}^\alpha m_{V_k}^{\alpha'} > 0$. If we interpret the kernel columns of Figure 6.1 as the results of an infinite series of consecutive physical measurements of the spin configuration $\sigma \sim \mu$, then this kernel describes a spin system in which the spin space V is partitioned into zones that evolves almost independently⁹, but according to different self-correlation times for each V_k (here V_1 is the slowest and V_3 is the fastest). This will be investigated in the next paper, where we show how to deal with the Cavity method by applying these ideas not to the magnetization kernel, but to the kernel of the Cavity fields.

As final remark, we insist that almost all the arguments presented here are not limited to the Magnetization kernel. As we shall see, we can build kernels from any bounded random variable, and then chose the one that best fit to our problem.

7 Acknowledgments

I wish to thank Amin Coja-Oghlan for explaining the ideas presented in[9], and their connection with Graph Theory and Replica Symmetry Breaking. I also wish to thank Giorgio Parisi, Francesco Guerra, Pietro Caputo, Nicola Kistler, Pan Liming, Francesco Concetti and Riccardo Balzan for interesting discussions and suggestions, and the anonymous Referee for his question on the Doob martingale.

This project has received funding from the European Research Council (ERC) under the European Union’s Seventh Framework Programme (grant agreement No [278857]) and the European Union’s Horizon 2020 research and innovation programme (grant agreement No [694925]).

References

- [1] G. Parisi, *A sequence of approximate solutions to the S-K model for spin glasses*, Journal of Physics A 13 (1980).
- [2] G. Parisi, M. Mezard, M. Virasoro, *Spin Glass theory and Beyond*, World Scientific (1987).

⁹Actually, the zones show a peculiar correlation structure since there is a hierarchy between the evolution of the volumes V_k . In fact, when the magnetization of V_k jumps to a new configuration then all volumes with $V_j, j > k$ also change. This is an expression of the fractal landscape, in which to escape from a certain cluster one must also escape from all its sub-clusters.

- [3] F. Guerra, *Broken replica symmetry bounds in the mean field spin glass model*, Communications in Mathematical Physics 233 (2003).
- [4] M. Talagrand, *The Parisi Formula*, Annals of Mathematics 163 (2006).
- [5] E. Marinari, G. Parisi, F. Ricci-Tersenghi, J. Ruiz-Lorenzo, F. Zuliani, *Replica Symmetry Breaking in Short-Range Spin Glasses: Theoretical Foundations and Numerical Evidences*, Journal of Statistical Physics 98 (2000). Updated 2014 version: <https://arxiv.org/abs/cond-mat/9906076>
- [6] D. Panchenko, *The Sherrington-Kirkpatrick Model*, Springer (2013).
- [7] D. Panchenko, *Spin glass models from the point of view of spin distributions*, Annals of Probability 41 (2013).
- [8] M. Krein, D. Milman, *On extreme points of regular convex sets*, Studia Mathematica 9 (1940).
- [9] V. Bapst, A. Coja-Oghlan, *Harnessing the Bethe Free Energy*, Random Structures and Algorithms 49 (2016).
- [10] A. Coja-Oghlan, W. Perkins, K. Skubch, *Limits of discrete distributions and Gibbs measures on random graphs*, European Journal of Combinatorics 66 (2017).
- [11] A. Coja-Oghlan, W. Perkins, *Belief propagation on replica symmetric random factor graph models*, to appear in Annales de L’Institut Henri Poincaré D (2017).
- [12] P. Diaconis, S. Janson, *Graph limits and exchangeable random graphs*, Rendiconti di Matematica e delle sue Applicazioni VII 28 (2008).
- [13] E. Szemerédi, *Regular partitions of graphs*, Problèmes combinatoires et théorie des graphes, Colloques Internationaux du CNRS 260, Paris (1978).
- [14] L. Lovasz, *Large Networks and Graph Limits*, AMS Colloquim Publications 60 (2012).
- [15] J. L. Doob, *Regularity properties of certain families of chance variables*, Transactions of the American Mathematical Society 47 (1940)
- [16] F. Guerra, *Fluctuations and thermodynamic variables in mean field spin glass models*, Stochastic Processes, Physics and Geometry II, World Scientific Singapore (1995).
- [17] F. Guerra, *Functional order parameters for the quenched free energy in mean field spin glass models*, Field Theory and Collective Phenomena, World Scientific Singapore (1995).
- [18] A. Barra, A. Di Biasio, F. Guerra, *Replica symmetry breaking in mean-field spin glasses through the Hamilton–Jacobi technique*, Journal of Statistical Mechanics 09 (2010).

- [19] D. Williams, *Probability with Martingales*, Cambridge University Press (1991).
- [20] G. Parisi, F. Ricci-Tersenghi, *On the origin of ultrametricity*, Journal of Physics A 33 (2000).
- [21] M. Aizenman, P. Contucci, *On the Stability of the Quenched state in Mean Field Spin Glass Models*, Journal of Statistical Physics 92 (1998).