University of Groningen

Metastates in Finite-type Mean-field Models

Iacobelli, Giulio; Kuelske, Christof

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

[Link to publication in University of Groningen/UMCG research database](#)

Metastates in Finite-type Mean-field Models: Visibility, Invisibility, and Random Restoration of Symmetry

Giulio Iacobelli · Christof Külske

Abstract We consider a general class of disordered mean-field models where both the spin variables and disorder variables η take finitely many values. To investigate the size-dependence in the phase-transition regime we construct the metastate describing the probabilities to find a large system close to a particular convex combination of the pure infinite-volume states. We show that, under a non-degeneracy assumption, only pure states j are seen, with non-random probability weights w_j for which we derive explicit expressions in terms of interactions and distributions of the disorder variables. We provide a geometric construction distinguishing invisible states (having $w_j = 0$) from visible ones. As a further consequence we show that, in the case where precisely two pure states are available, these must necessarily occur with the same weight, even if the model has no obvious symmetry relating the two.

1 Introduction

Dealing with phase transitions in the theory of Gibbs measures of disordered systems is usually not an easy task. First of all, one likes to understand which are the possible phases and how do they depend on the realization of the disorder. This can be a formidable task, even for mean-field models, as the history of the SK-model shows. Secondly, even if we suppose that the phases are identified, it is not a priori clear what role they will play for the

G. Iacobelli
Department of Mathematics and Computing Sciences, University of Groningen, Nijenborgh 9,
9747 AC Groningen, The Netherlands
e-mail: G.Iacobelli@rug.nl

C. Külske (✉)
Ruhr-Universität Bochum, Fakultät für Mathematik, Universitätsstraße 150, 44780 Bochum, Germany
e-mail: Christof.Kuelske@rub.de
url: http://www.ruhr-uni-bochum.de/ffm/Lehrstuehle/Kuelske/kuelske.html

typical behavior of a large but finite system. Indeed, in a regime where there are competing extremal phases (say a plus and a minus phase in a random field model) it may depend on the realization of the disorder variables which of the convex combinations the system will be close to in equilibrium. Some of the possible infinite-volume equilibrium states might not even show up in a typical large volume. To make sense of these questions, the concept of a metastate has been invented by Aizenman and Wehr [1], Newman and Stein [19–21], being a probability measure which gives the weights in the large-volume asymptotics to find a system close to one of the possible candidates among the Gibbs measures. We stress that the metastate is a notion describing purely the equilibrium behavior. Yet another type of questions is of dynamic nature: Suppose a system undergoes a Glauber dynamics, how much time will it need to go from an initial state to it global free energy minimum? For some initial states this time is exceptionally long, a phenomenon called metastability (not to be confused with the notion of a metastate), and again, to derive precise asymptotic saying how long, examples of mean-field systems have been very instructive model systems [4, 12].

On the lattice the metastate has been shown to be a useful concept in spinglasses by Newman and Stein [18] and Arguin et al. [3] who showed that there is only one groundstate pair in the two-dimensional Edwards-Anderson model in the half-plane, using translation-ergodicity and Burton-Keane type of arguments.

Explicit constructions for lattice models are difficult (see however [11] where the influence of random boundary conditions on an Ising model was analyzed), but possible in mean-field models. Previously treated examples are given in very specific models, namely the symmetric random field Ising model and Hopfield model with a finite or a growing number of patterns [7, 8, 13–15].

In this paper we aim for completeness in a particular direction, namely disordered mean-field models with finitely many values for both spin and disorder variables. Such models include in particular the random-field Curie-Weiss Ising model and Potts-type Curie-Weiss random field models with or without symmetries in Hamiltonians or random field distributions. What we aim for is the abstract construction of the phase diagram, embellished with probability weights giving us the appearance of the candidate states. That is, we first say *which* states are available. This, for disordered mean-field models comes from an investigation of the corresponding free energy (resp. rate functions) and is a standard thing. Next and new in our paper is the additional information on the *weights* with which they occur, and the proof of the validity of a corresponding approximate extreme decomposition, asymptotically for large volumes. This is then cast in the metastate formulation. The weights are obtained by studying the distribution of the free energy fluctuations w.r.t. to the disorder variables entering. Will the same type of results be true for corresponding lattice models at low temperatures at phase coexistence? We believe yes, but a proof will have to build around sophisticated expansion techniques and be technically rather challenging. One would need to show first the coexistence of states (as it was done for the random field Ising model in [5]), and then the dominance of one of the available states over the others for typical realizations of the disorder. The mean-field results should provide guidance for that, and moreover we believe that they are a rather nice complete example for a limit theorem in statistical mechanics.

1.1 The Models: Mean-field Models with Local Disorder

These are the models we consider. At each site $i = 1, \ldots, n$ there is a spin variable $\sigma(i)$ taking values in a finite set E and a disorder variable $\eta(i)$ taking values in the finite (possibly different) set E'. We write $\mathcal{P}(E)$ for the set of probability measures on E, and use similar notation for other spaces. We write $L_n = \frac{1}{n} \sum_{i=1}^{n} \delta_{\sigma(i)} \in \mathcal{P}(E)$ for the (total) empirical

measure of the spins and consider a twice continuously differentiable function F on $\mathcal{P}(E)$. The influence of the disorder variables on the Gibbs measures for the spins is through the local a priori measures $\alpha[b] \in \mathcal{P}(E)$, for any possible type of the disorder $b \in E'$. Hence the present analysis excludes models with disorder entering the interaction such as e.g. the Hopfield model treated in [8, 13].

Definition 1.1 The mean-field model with Hamiltonian $nF(\nu)$ and a priori measures $\alpha[b] \in \mathcal{P}(E)$, for all $b \in E'$, is given by the disorder-dependent finite-volume Gibbs measures

$$\mu_{F,n}[\eta(1), \ldots, \eta(n)](\sigma(1) = \omega(1), \ldots, \sigma(n) = \omega(n))$$

$$= \frac{1}{Z_{F,n}[\eta(1), \ldots, \eta(n)]} \exp\left(-nF\left(L_n^\omega\right)\right) \prod_{i=1}^{n} \alpha[\eta_i](\omega_i) \tag{1}$$

together with the prescription of a probability distribution $\pi \in \mathcal{P}(E')$ for the disorder variables according to which they are chosen independently over the sites. We assume $\pi(b) > 0$ for all $b \in E'$.

To summarize, our model depends on the triple of parameters (F, α, π) of: mean-field interaction F, a priori measures $\alpha = (\alpha[b])_{b \in E'}$, and disorder distribution π.

We need to introduce more notations. Given η, we write

$$\Lambda_n(b) = \{i \in \{1, 2, \ldots, n\}; \eta(i) = b\}$$

for all $b \in E'$, for the b-like sites. Write

$$\hat{\pi}_n(b) = \frac{|\Lambda_n(b)|}{n}$$

for the frequency of the b-like sites (empirical distribution of random field types). Write

$$\hat{L}_n(b) = \frac{1}{|\Lambda_n(b)|} \sum_{i \in \Lambda_n(b)} \delta_{\sigma(i)}$$

for the empirical spin-distribution on the b-like sites. Write $\hat{L}_n = (\hat{L}_n(b))_{b \in E'}$ for the vector of empirical distributions. The total empirical distribution is then the scalar product of $\hat{\pi}_n$ with the vector of empirical spin distributions

$$L_n = \sum_{b \in E'} \hat{\pi}_n(b)\hat{L}_n(b) \tag{2}$$

1.2 The Metastate on the Level of the States

Let us jump into the following definition of a metastate, obtained by a conditioning procedure, which was given first by Aizenman and Wehr [1]. There are different constructions of a metastate, but the present one will be the only one considered in the paper. This Aizenman-Wehr construction is related to a different and more intuitive construction as empirical averages of Gibbs measures along volume-(sub-)sequences by Newman and Stein. We refer to the monographs [6, 19].

Definition 1.2 Assume that, for every bounded continuous $\Xi : \mathcal{P}(E^\infty) \times (E')^\infty \to \mathbb{R}$ the limit

$$\lim_{n \uparrow \infty} \int \mathbb{P}(d\eta) \Xi(\mu_n[\eta], \eta) = \int J(d\mu, d\eta) \Xi(\mu, \eta) \tag{3}$$

exists. Then the conditional distribution $\kappa[\eta](d\mu) := J(d\mu|\eta)$ is called the *AW-metastate on the level of the states*.

As is common, continuity is meant in the following sense: A function on an infinite product of a finite space is continuous (w.r.t. local topology) if it is a uniform limit of local functions. For probability measures on $\mathcal{P}(E^\infty)$ we use the weak topology (according to which a sequence of measures converges iff it converges on continuous test-functions), and for $\mathcal{P}(E^\infty) \times (E')^\infty$, we use the product topology.

1.3 Main Theorem

How do we get the possible equilibrium states of the system? They are obtained as solutions to the following minimization problem.

Definition 1.3 Consider the free energy minimization problem

$$\hat{v} \mapsto \Phi[\pi](\hat{v}) \tag{4}$$

on $\mathcal{P}(E)^{E'}$, with the free energy functional

$$\Phi : \mathcal{P}(E') \times \mathcal{P}(E)^{E'} \mapsto \mathbb{R}$$

$$\Phi[\hat{\pi}](\hat{v}) = F\left(\sum_{b \in E'} \hat{\pi}(b)\hat{v}(b)\right) + \sum_b \hat{\pi}(b) S(\hat{v}(b)|\alpha[b]) \tag{5}$$

where $S(p_1|p_2) = \sum_{a \in E} p_1(a) \log \frac{p_1(a)}{p_2(a)}$ is the relative entropy. We say that the random mean-field system obeys the *non-degeneracy condition (1)* if $\hat{v} \mapsto \Phi[\pi](\hat{v})$ has a finite set of minimizers $M^* = M^*(F, \alpha, \pi)$ where all the eigenvalues of the Hessian are strictly positive.

It is very hard for a system not to satisfy this condition and we will assume in the following that it is satisfied. If it is true the vector of the empirical spin distributions of the system, $\hat{L}_n$, will concentrate around the set M^*. More than that, it may even concentrate on a smaller set. The following theorem about the metastate will tell us how this concentration will take place and get the weights w_j.

Let $\hat{v}_j$ be a fixed element in M^*. Let us consider the linearization of the free energy functional at the fixed minimizers as a function of π, which reads

$$\Phi[\tilde{\pi}](\hat{v}_j) - \Phi[\pi](\hat{v}_j) = -B_j[\tilde{\pi} - \pi] + o(\|\tilde{\pi} - \pi\|) \tag{6}$$

where

$$B_j[\tilde{\pi} - \pi] = -\left(dF_{\pi \cdot \hat{v}_j}\left(\sum_b (\tilde{\pi}(b) - \pi(b))\hat{v}_j(b)\right) + \sum_b [\tilde{\pi}(b) - \pi(b)] S(\hat{v}_j(b)|\alpha[b])\right) \tag{7}$$

This defines an affine function on the tangent space of field type measures $T\mathcal{P}(E')$ (i.e. vectors which sum up to zero), for any j.

Definition 1.4 We call B_j the stability vector of $\hat{v}_j$. We call

$$R_j := \left\{ x \in T\mathcal{P}(E'), \langle x, B_j \rangle > \max_{k \neq j} \langle x, B_k \rangle \right\} \tag{8}$$

the stability region of $\hat{v}_j$.

Now comes our second condition.

Definition 1.5 We say the vector $B = (B_1, \ldots, B_k)$ satisfies the *non-degeneracy condition (2)* if no different minimizers j, j' have the same $B_j = B_{j'}$

In other words the randomness lifts all symmetries. The Hopfield model for instance (which was excluded already before because of the structure of the randomness) would also be excluded at this state since it has spin-flip symmetry for all realizations, and so minimizers in the free energy would come in pairs.

Note that the condition in the definition implies that $(\bigcup_{j=1,\ldots,k} R_j)^c$ has zero Lebesgue measure in $T\mathcal{P}(E')$. Indeed, if the map $j \mapsto \langle x, B_j \rangle$ has no unique maximizer for fixed x, then, for some pair $j \neq k$ we have that $\langle x, B_j - B_k \rangle = 0$. For fixed j, k this set of x's is a hyperplane (hence a measure zero set) since, by assumption, $B_j \neq B_k$.

We note the following simple but important geometric lemma.

Lemma 1.6 $R_j \neq \emptyset \Leftrightarrow B_j \in \text{ex}(\mathcal{H}_{\text{conv}}\{B_1, \ldots, B_k\})$.

Here, for a subset $A \subset \mathbb{R}^d$, $\mathcal{H}_{\text{conv}}(A)$ denotes the *convex hull* of A, that is the smallest convex set which contains A. $\text{ex}(C)$, for a convex set C, denotes the extremal points of C, that is those points which can not be written as a non-trivial convex combination with points from C. In our case $\mathcal{H}_{\text{conv}}\{B_1, \ldots, B_k\}$ is a convex polyhedron and $\text{ex}(\mathcal{H}_{\text{conv}}\{B_1, \ldots, B_k\})$ is the smallest set of points which generates it.

Proof of Lemma 1.6 We prove the implication "$\Rightarrow$" by contradiction. Suppose that B_j is not an extremal point. Then it can be written as a non-trivial convex combination $B_j = \sum_i \alpha_i B_i$ with $\sum_{i=1}^k \alpha_i = 1$, where $\alpha_i \geq 0$ and non-zero only for $B_i \in \text{ex}(\mathcal{H}_{\text{conv}}\{B_1, \ldots, B_k\})$. Any vector $x \in R_j$ satisfies $\langle x, B_j \rangle > \langle x, B_i \rangle$ for all $i \neq j$ and hence $\langle x, B_j \rangle = \sum_i \alpha_i \langle x, B_i \rangle > \sum_{i=1}^k \alpha_i \langle x, B_i \rangle = \langle x, B_j \rangle$. This is a contradiction and hence $R_j = \emptyset$.

To prove the opposite implication "$\Leftarrow$" let us consider an extremal point B_j and note the following: If $B_j \notin \mathcal{H}_{\text{conv}}\{B_1, \ldots, B_{j-1}, B_{j+1}, \ldots, B_k\}$ then, after a suitable translation and rotation, we can find coordinates such that the vectors take the form $B_j = (0, \ldots, 0, B_{j,d})$ and $B_i = (B_i', B_{i,d})$ with $B_{j,d} > 0$ and $B_{i,d} \leq 0$ for $i \neq j$. (The latter statement follows from the fact that there is a separating hyperplane between $\mathcal{H}_{\text{conv}}\{B_1, \ldots, B_{j-1}, B_{j+1}, \ldots, B_k\}$ and the point B_j. This finite-dimensional version of the Hahn-Banach theorem is a classical result in geometry, see Theorem 1.2.4 in [17]. Having this separating hyperplane we choose the origin as the orthogonal projection of B_j to this plane, the first coordinates as orthogonal coordinates inside the plane, and the last coordinate axis pointing in the direction of B_j.) The proof relies on the last two inequalities. Indeed we have, with the general notation $x = (x', x_d) \in \mathbb{R}^{d-1} \times \mathbb{R}$ that

$$R_j = \{x \in \mathbb{R}^d : \forall i \neq j \text{ holds } \langle x, B_j - B_i \rangle > 0\}$$

$$= \{x \in \mathbb{R}^d : \forall i \neq j \text{ holds } \langle x', B_j' - B_i' \rangle + x_d(B_{j,d} - B_{i,d}) > 0\}$$

$$= \left\{ x \in \mathbb{R}^d : x_d > \max_{i:i \neq j} \frac{\langle x', B_i' - B_j' \rangle}{B_{j,d} - B_{i,d}} \right\} \neq \emptyset \tag{9}$$

$\square$

Before we state our theorem let us introduce the kernels

$$\gamma[b](a|\nu) = \frac{e^{-dF_\nu(a)}\alpha[b](a)}{\sum_{\bar{a} \in E} e^{-dF_\nu(\bar{a})}\alpha[b](\bar{a})} \tag{10}$$

with $\nu \in \mathcal{P}(E)$. These are the limiting local distributions of a spin at a site with a disorder variable in the state b if the empirical spin-average of the rest of the system is given by the measure ν. The products over all sites of these quantities, for $\nu = \pi \hat{\nu}_j$, will play the role of pure measures.

We are now in the position to give our main result.

Theorem 1.7 *Assume that the model satisfies the* non-degeneracy assumptions (1) *and* (2). *Define the weights*

$$w_j := \mathbb{P}_\pi(G \in R_j) \tag{11}$$

where $G \in T\mathcal{P}(E')$ is a centered Gaussian variable with the same covariance as $\sqrt{n}(\hat{\pi}_n - \pi)$ which is given by the expression $C_\pi(b, b') = \pi(b)1_{b=b'} - \pi(b)\pi(b')$.

Then $\sum_{j=1}^k w_j = 1$ and the metastate on the level of the states equals

$$\kappa[\eta](d\mu) = \sum_{j=1}^k w_j \delta_{\mu_j[\eta]}(d\mu) \tag{12}$$

where $\mu_j[\eta] := \prod_{i=1}^\infty \gamma[\eta(i)](\cdot |\pi \hat{\nu}_j)$.

Comment. We like to reformulate our result on the visibility or invisibility of the phases in the following way. Let us denote by $M^{**} = \{\hat{\nu} \in M^* : w_{\hat{\nu}} > 0\}$ the subset of *visible pure phases* in the pure phases M^*. Let us use the symbol $B_.$ for the bijection (under our hypothesis)

$$B_. : M^* \to T\mathcal{P}(E')$$

$$\hat{\nu} \mapsto B_{\hat{\nu}}$$

Then we can write in short

$$M^{**} = (B_.)^{-1}(\text{ex}(\mathcal{H}_{\text{conv}}(B_.(M^*))))$$

Let us derive the following immediate consequence which provides a symmetry, due to the randomness (the symmetry of the Gaussian, obtained via the CLT).

Corollary 1.8 *Suppose that the system admits precisely two pure phases, i.e. $|M^*| = 2$. Then the metastate is the symmetric mixture between the two, i.e.*

$$\kappa[\eta](d\mu) = \frac{1}{2}\delta_{\mu_1[\eta]}(d\mu) + \frac{1}{2}\delta_{\mu_1[\eta]}(d\mu) \tag{13}$$

The corollary is clear from the theorem since in that case $R_1 = -R_2$ and this implies by the *non-degeneracy assumption (2)* that $w_1 = w_2$.

Corollary 1.9 *Suppose that the random-field is two-valued, i.e. $|E'| = 2$, and the number of pure phases $|M^*| \geq 2$ arbitrary. Then the set of visible states has two elements and $w(\hat{v}) = \frac{1}{2}$ for both elements $\hat{v} \in M^{**}$.*

The corollary is clear from the theorem since any convex polyhedron in one dimension has only two extremal points.

For illustrational purposes recall the situation in the mean-field random field Ising model with two-valued symmetrically distributed random field with coupling strength ε and temperature β^{-1}. In this model the β^{-1}, ε-plane contains a bounded open region for which $|M^*(\beta^{-1}, \varepsilon)| = 2$. The boundary of this region is a curve which splits into a part for which $|M^*(\beta^{-1}, \varepsilon)| = 3$ and a part for which $|M^*(\beta^{-1}, \varepsilon)| = 1$. In the complement of the union of those previous regions we have $|M^*(\beta^{-1}, \varepsilon)| = 1$. The situation of the second corollary is met on the curve where $|M^*(\beta^{-1}, \varepsilon)| = 3$.

1.3.1 Exploiting the Mean-field Equation

Using variational calculus and assuming differentiability of F one sees that the minimizers of the variational problem above must satisfy the consistency (mean-field) equations

$$\hat{v}[b](a) = \gamma[b](a|\pi \cdot \hat{v}) \tag{14}$$

which are coupled over $b \in E'$. Summing over these indices one gets the mean-field equation for the total empirical mean $v = \pi \cdot \hat{v}$ of the form

$$v(a) = \sum_{b \in E'} \pi(b)\gamma[b](a|v) \tag{15}$$

We note the following lemma.

Lemma 1.10 *Define the function $\hat{\Gamma} : \mathcal{P}(E) \to \mathcal{P}(E)^{E'}$ by the r.h.s. of the mean field equation, namely*

$$\hat{\Gamma}(v) = (\gamma[b](\cdot|v))_{b \in E'} \tag{16}$$

Define the function $\hat{B} : \mathcal{P}(E) \to T\mathcal{P}(E')$ by

$$\hat{B}_v[b] = \log \sum_{a \in E} e^{-dF_v(a)} \alpha[b](a) - C$$

$$C = \frac{1}{|E'|} \sum_{b \in E'} \log \sum_{a \in E} e^{-dF_v(a)} \alpha[b](a) \tag{17}$$

Then, for all $\hat{v} \in M^*$ we have that

$$\hat{v} = \hat{\Gamma}(\pi\hat{v})$$

$$B_{\hat{v}} = \hat{B}_{\pi\hat{v}} \tag{18}$$

For all $v \in \pi M^*$ we have that the free energy can be written as

$$\Phi[\pi](\hat{\Gamma}(v)) = F(v) - \langle dF_v, v \rangle - \langle \hat{B}_v, \pi \rangle + C \tag{19}$$

The first statement is just a rephrasing of the mean-field equation. It serves us to see that there is a bijection between $\pi M^* = \{\pi\hat{v}|\hat{v} \in M^*\} \subset \mathcal{P}(E)$ (a subset in a space of measures with dimension $|E| - 1$) and M^* (a subset in a space of measures with dimension $(|E| - 1)^{|E'|}$). The second part means that the logarithm of the normalization factor ("little partition function") of the mean-field kernels in the total empirical distribution v of type b produces the b'th component of the stability vector corresponding to the minimizer with total empirical mean v.

The interesting feature is that the form of π does not enter at all into this formula (it enters however through the question which minimizer $\hat{v}$ and hence also v appears).

Proof The first part is obvious. To prove the second part, for $\hat{v} \in M^*$ we write, with a constant C' to be determined

$$-B_{\hat{v}}[b] = \sum_{a \in E} dF_{\pi \cdot \hat{v}}(a)\hat{v}[b](a) + S(\hat{v}[b]|\alpha[b]) - C'$$

$$\equiv \hat{v}[b](dF_{\pi \cdot \hat{v}}(\cdot)) + S(\hat{v}[b]|\alpha[b]) - C'$$

$$= -\log \sum_{a \in E} e^{-dF_{\pi\hat{v}}(a)}\alpha[b](a) - C' \tag{20}$$

where the last equality follows from the mean-field equation. This proves the second claim. The last claim follows from the first equality of the last display multiplying with $\pi(b)$ and summing over $b \in E'$. $\qquad\square$

1.4 Ising Random-field Examples

Let us take the Ising model with $F(v) = -\beta(v(+)^2 + v(-)^2)$.

Any possible local single-site measure α can be described as an $\alpha[h](\sigma_i) = \frac{e^{h\sigma_i}}{2\cosh h}$. Any $v = v_m$ can be described in terms of its mean value $v_m(+) - v_m(-) = m$.

So we can write

$$\hat{B}_{v_m}[h] \equiv \hat{B}_{v_m}[\alpha[h]] = \log \frac{e^{\beta 2\frac{1+m}{2}+h} + e^{\beta 2\frac{1-m}{2}-h}}{2\cosh h} - C$$

$$= \beta + \log \frac{\cosh(\beta m + h)}{\cosh h} - C \tag{21}$$

Let us now fix $E' = \text{supp}(\pi) = \{\alpha_h : h \in \{h_1, h_2, \dots, h_L\}\}$ as the set of allowed local measures. This gives us the normalized vector in the tangent space $T\mathcal{P}(E')$ with entries

$$\hat{B}_{v_m}[h_i] := \log \frac{\cosh(\beta m + h_i)}{\cosh h_i} - \frac{1}{L}\sum_{j=1}^{L} \log \frac{\cosh(\beta m + h_j)}{\cosh h_j} \tag{22}$$

Writing a vector with $L = |E'|$ components we have

$$\hat{B}_{v_m} = \begin{pmatrix} \log \frac{\cosh(\beta m + h_1)}{\cosh h_1} \\ \dots \\ \log \frac{\cosh(\beta m + h_L)}{\cosh h_L} \end{pmatrix} - \frac{1}{L}\sum_{j=1}^{L} \log \frac{\cosh(\beta m + h_j)}{\cosh h_j} \begin{pmatrix} 1 \\ \dots \\ 1 \end{pmatrix}$$

Lemma 1.11 *Let $E' \subset \mathbb{R}$, $2 \le |E'| < \infty$. Then the map $m \mapsto \hat{B}_{v_m}$ is injective.*

Proof We have at least two elements, $h_1 < h_2$ (after possible change of indices) in E'. Let $v_m, v_{\tilde{m}}$ be given with $\hat{B}_{v_m} = \hat{B}_{v_{\tilde{m}}}$. By easy manipulations looking at the first two components of $\hat{B}$ the latter implies that

$$\frac{\cosh(\beta m + h_1)}{\cosh(\beta m + h_2)} = \frac{\cosh(\beta \tilde{m} + h_1)}{\cosh(\beta \tilde{m} + h_2)} \tag{23}$$

From this follows $m = \tilde{m}$ by injectivity of the function $x \mapsto \frac{\cosh x}{\cosh(x+1)}$. $\square$

Let us extend the random-field Ising model to a non-quadratic Hamiltonian $F(v) = G(v(+) - v(-))$ and general local measures $\alpha = (\alpha[h])_{h \in E'}$ with a finite set E' just as above in the quadratic case.

Then the mean field equation becomes

$$m = \sum_{i=1}^{L} \pi(h_i) \tanh(-G'(m) + h_i) \tag{24}$$

The stability vector becomes

$$\hat{B}_{v_m}[h_i] := \log \frac{\cosh(-G'(m) + h_i)}{\cosh h_i} - \frac{1}{L} \sum_{j=1}^{L} \log \frac{\cosh(-G'(m) + h_j)}{\cosh h_j} \tag{25}$$

Then the injectivity of the map $m \mapsto \hat{B}_{v_m}$ holds under the assumption that $m \mapsto G'(m)$ is injective, by the same proof, replacing m by $-G'(m)$ in (23).

We have thus proved the following statement.

Proposition 1.12 *For a random-field Ising model with Hamiltonian $F(v) = G(v(+) - v(-))$ and G' injective the non-degeneracy assumption (2) is automatically satisfied, for any distribution of random fields with finite support.*

It is easy to create a two-minima situation where there is no symmetry, by looking at the equal-depth condition for the free energy

$$\Phi[\pi](\hat{\Gamma}(v_m)) = F(v_m) - \sum_{a \in E} d F_{v_m}(a) v_m(a) - \sum_{b \in E'} \pi(b) \log \sum_{a \in E} e^{-d F_{v_m}(a)} \alpha[b](a)$$

$$= G(m) - m G'(m) - \sum_{i=1}^{L} \pi(h_i) \log \frac{\cosh(-G'(m) + h_i)}{\cosh h_i}$$

where both minima would get the same weight in the metastate necessarily.

In fact, a situation with precisely two minimizers not related by symmetry was proved to occur (even) for the (symmetric) model $G(m) = -\frac{\beta m^2}{2}$, $E = E' = \{1, -1\}$, $\pi(1) = \frac{1+\alpha}{2} = 1 - \pi(-1)$, $\alpha[b](a) = \frac{e^{\beta \varepsilon a b}}{\cosh \beta \varepsilon}$, for the region R_{34} in the $(\beta^{-1}, \varepsilon)$-plane characterized in [16] and depicted below, for a suitable choice of $\alpha = \alpha(\beta, \varepsilon) > 0$. We note that the part of the phase diagram corresponding to the case $\alpha = 0$ was given already in [2].

EXTENDED PHASE DIAGRAM
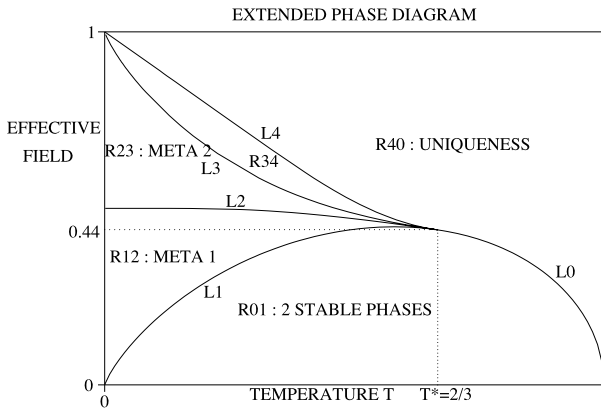
1.5 Potts Random Field Examples

Let us take the Potts model with quadratic interaction

$$F(v) = -\frac{\beta}{2}(v(1)^2 + \cdots + v(q)^2)$$

in the presence of the local single-site measures $\alpha[b](\sigma_i)$ (specified below) where we write

$$E' = \mathrm{supp}(\pi) = \{\alpha[b] : b \in \{b_1, b_2, \ldots, b_L\}\}$$

Then we have for the stability vector

$$\hat{B}_v = \begin{pmatrix} \log \sum_{a=1}^{q} e^{\beta v(a)} \alpha[b_1](a) \\ \cdots \\ \log \sum_{a=1}^{q} e^{\beta v(a)} \alpha[b_L](a) \end{pmatrix} - \frac{1}{L} \sum_{j=1}^{L} \log \sum_{a=1}^{q} e^{\beta v(a)} \alpha[b_j](a) \begin{pmatrix} 1 \\ \cdots \\ 1 \end{pmatrix}$$

Remark The map $\hat{B}. : \mathcal{P}(E) \to T\mathcal{P}(E')$ is a map between spaces of dimension $|E| - 1$ and $|E'| - 1$. It has a chance to be injective as such (on the whole space $\mathcal{P}(E)$) only when $|E'| \geq |E|$.

Let us take $E \equiv E'$ and π to be the equidistribution and switch to the specific case $\alpha[b](a) = \frac{e^{B 1_{b=a}}}{e^B + q - 1}$ (random field with homogeneous intensity). The kernels become

$$\gamma[b](a|v) = \frac{e^{\beta v(a) + B 1_{a=b}}}{\sum_{\bar{a} \in E} e^{\beta v(\bar{a}) + B 1_{\bar{a}=b}}}$$

We will be looking at measures in $v_{j,u} \in \mathcal{P}(E)$ of the form $v_{j,u}(j) = \frac{1 + u(q-1)}{q}$, $v_{j,u}(i) = \frac{1-u}{q}$ for $i \neq j$. The stability vector for $v_{1,u}$ is given by

$$\hat{B}_{v_{1,u}} = \begin{pmatrix} \frac{q-1}{q} \log \frac{e^{\beta u + B} + q - 1}{e^{\beta u} + e^B + q - 2} \\ -\frac{1}{q} \log \frac{e^{\beta u + B} + q - 1}{e^{\beta u} + e^B + q - 2} \\ \cdots \\ -\frac{1}{q} \log \frac{e^{\beta u + B} + q - 1}{e^{\beta u} + e^B + q - 2} \end{pmatrix}$$

the other ones are related by symmetry. We note that the first entry is strictly positive while the other entries are negative (for $B > 0$ and $u > 0$).

We have the mean-field equation for u of the form

$$u = \frac{e^{\beta u}}{e^{\beta u} + e^B + (q-2)} - \frac{1}{e^{\beta u+B} + (q-1)} \tag{26}$$

We notice that $u = 0$ is always a solution, and for $B = 0$ we obtain exactly the known mean-field equation for Potts without disorder. The latter model shows a first-order transition as a function of temperature at critical temperature $\beta_c = \frac{2(q-1)}{q-2} \log(q-1)$ [10].

The r.h.s. of (26) is always positive, as a computation shows. This gives rise to a non-trivial solution u, in a certain range of parameters. Note that this non-trivial solution is not always the one to be chosen. It is to be chosen iff $\Phi[\pi](\hat{\Gamma}(\nu_{j,u})) < \Phi[\pi](\hat{\Gamma}(\nu_{j,u=0}))$. So, the first order transition point is given by equality of the last equation. Forgetting a u-independent term we have, independently of the direction j,

$$\Phi[\pi](\hat{\Gamma}(\nu_{j,u})) = \log \frac{e^B + q - 1}{e^{\beta u} + e^B + q - 2} + \frac{\beta(q-1)}{2q}u^2 + \frac{\beta}{q}u - \frac{1}{q}\log\frac{e^{\beta u+B} + q - 1}{e^{\beta u} + e^B + q - 2} \tag{27}$$

with the property that $\Phi[\pi](\hat{\Gamma}(\nu_{j,u=0})) = 0$. For illustrational purposes let us focus on the case $q = 3$. We don't provide a complete bifurcation analysis here, but just outline the picture. The case $B = 0$ is perfectly understood and we know that there is a first order transition at the critical inverse temperature $\beta = 4\log 2$. The nature of the transition stays the same when B takes small enough positive values and there is a line in the space of temperature and coupling strength B of an equal-depth minimum at $u = 0$ and a positive value of $u = u^*(\beta, q)$. (See Fig. 1 for a numerical example.) Along this line the set of Gibbs measures is strictly bigger then the set of states which are seen under the metastate.

The Plot shows the graph of $u \mapsto \Phi[\pi](\hat{\Gamma}(\nu_{j,u}))$ for $B = 0.3, q = 3, \beta = 4\log 2 + 0.03203$ at which there is the first order transition.
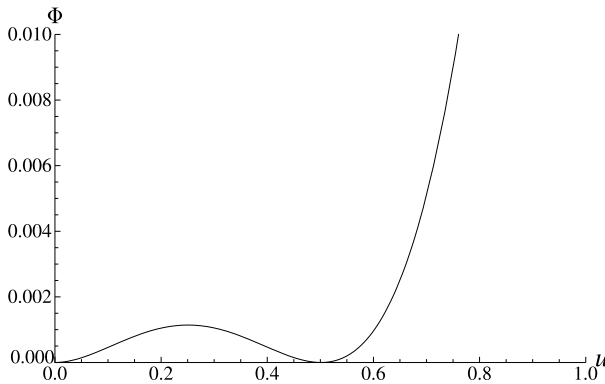
The metastate becomes $\kappa[\eta](d\mu) = \frac{1}{3}\sum_{j=1}^3 \delta_{\mu_j[\eta]}$ with $\mu_j[\eta] = \prod_{i=1}^\infty \gamma[\eta(i)](\cdot|\nu_{j,u=u^*(\beta,q)})$. This follows from the form of the stability vector using that $\hat{B}_{\nu_{1,u=0}} = 0$ and hence lies in the convex hull of the three others.

1.6 Strategy of Proof, Non-degeneracy Assumption and Concentration

The outline of the remainder of the paper is as follows.

We begin in Sect. 2 with a discussion of another related result, namely the metastate on the level of the empirical spin-distributions. The theorem is quite analogous, the same weights w_j appear, and the proof is slightly easier than that of the full theorem. To arrive at the proof of this theorem we will discuss the concentration property of the vector of the empirical distributions for good realizations of the disorder which will force the system to be in one definite state. In particular it will show how the non-degeneracy assumptions (1) and (2) are naturally used in that argument and this will explain how the CLT for empirical distributions of disorder variables translates into the form of the weights w_j.

Then we will turn in Sect. 3 to the proof of the metastate theorem on the level of states and conclude.

2 The Metastate on the Level of the Empirical Spin-distribution

Two ways of looking at the spin-distributions of disordered mean-field systems are natural. In the first one, described in the introduction, we focus on measures on the spins themselves, and evaluate them on local observables. In the second one, we focus on aggregate properties of the system, and look at functions of the empirical spin-distribution of the whole system. From the second point of view it is natural to make the following definition of a metastate on the level of the empirical spin-distribution.

Denote by $\rho[\eta](n) := \mu_{F,n}[\eta](L_n)$ the image of the finite-volume Gibbs-measure under the empirical distribution. This defines a disorder-dependent element in $\mathcal{P}(\mathcal{P}(E))$. Under our assumptions these measures will concentrate on the finite set $\pi M^* = \{\pi \hat{v}_j, j = 1, \ldots, k\}$. It is useful to introduce a metastate which tells us more precisely how this concentration takes place. This is the reason for the following definition.

Definition 2.1 Assume that, for every bounded continuous $\Psi : \mathcal{P}(\mathcal{P}(E)) \times (E')^{\infty}$ the limit

$$\lim_{n \uparrow \infty} \int \mathbb{P}(d\eta) \Psi(\rho[\eta](n), \eta) = \int K(d\rho, d\eta) \Psi(\rho, \eta) \tag{28}$$

exists. Then the conditional distribution $\bar{\kappa}[\eta](d\rho) := K(d\rho|\eta)$ is called the *metastate on the level of the empirical spin-distribution.*

Believing in the first theorem it is not surprising that this metastate takes the following form.

Theorem 2.2 *Under the* non-degeneracy assumptions (1) *and* (2), *we have*

$$\bar{\kappa}[\eta](d\rho) = \sum_{j=1}^{k} w_j \delta_{\delta_{\pi \hat{v}_j}}(d\rho) \tag{29}$$

for $\mathbb{P}_{\pi}$*-a.e.* η.

As a difference with respect to the first theorem let us point out that in this case the dependence on the disorder has vanished on the r.h.s.

Proof of Theorem 2.2 For $n_1 < n_2$ integers, let's define

$$X_{[n_1, n_2]}[\eta] = \frac{1}{\sqrt{n_2 - n_1 + 1}} \sum_{i=n_1}^{n_2} \delta_{\eta_i} - \sqrt{n_2 - n_1 + 1}\pi \tag{30}$$

Define n, l-dependent *good-sets* $\mathcal{H}_{n,l}^{\tau}$ of the realization of the randomness as follows

$$\mathcal{H}_{i,n,l}^{\tau} := \left\{ \eta \in (E')^{n-l} : X_{[l+1,n]}[\eta] \in R_{i,n}^{\tau} \right\}$$

$$\mathcal{H}_{n,l}^{\tau} := \bigcup_{i=1}^{k} \mathcal{H}_{i,n,l}^{\tau} \tag{31}$$

where $R_{i,n}^{\tau} := \{x \in T\mathcal{P}(E') : \langle x, B_i \rangle - \max_{k \neq i} \langle x, B_k \rangle > n^{-\frac{1}{2}+\tau}, \|x\| \leq n^{\frac{\tau}{4}}\}$, where $0 < \tau < \frac{1}{2}$. For the sake of clarity set $\delta_n = n^{-\frac{1}{2}+\tau}$. The chosen range of τ ensures that $\delta_n \downarrow 0$, but not too fast, namely in such a way that $\sqrt{n}\delta_n \uparrow \infty$.

$\mathcal{H}_{i,n,l}^{\tau}$ is a region of the disorder random variables which allows us to deduce that the measure on the empirical distribution will be with large probability inside a ball around $\pi \hat{v}_i^*$.

Remark We need $\delta_n \downarrow 0$ because we want to cover all of the corresponding *stability-region* R_i (8) in the large-n limit. The condition regarding the velocity with which δ_n is going to 0 ensures the concentration of the measure around a particular minimizer, in other words it will enable us to see the breaking of the degeneracy of the minimizers caused by the fluctuations of $\hat{\pi}_n$. The relevance of the cutoff $\|x\| \leq n^{\frac{\tau}{4}}$ will be seen later.

Lemma 2.3 *Let us assume that* $\eta \in \mathcal{H}_{i,n,0}^{\tau}$. *Then*

$$\mu_{F,n}[\eta(1), \ldots, \eta(n)](d(L_n, \pi \hat{v}_i^*) \leq \varepsilon) \geq 1 - \bar{r}(\varepsilon, n) \tag{32}$$

where $\lim_{n \uparrow \infty} \bar{r}(\varepsilon, n) = 0$ *for all* $\varepsilon > 0$.

Proof Call $M_n := \{v \in \mathcal{P}(E) : \exists \omega \in E^n \text{ such that } L_n^{\omega} = v\}$. To every element $v \in M_n$ correspond several possible values of the empirical distribution vectors $\hat{L}_n \in \mathcal{P}(E)^{E'}$, given $\hat{\pi}_n$. We call this set $\hat{M}_n := \{\hat{v} \in \mathcal{P}(E)^{E'} : \exists \omega \in E^n \text{ such that } \hat{v} = \hat{L}_n\}$. Let's define $\rho^{\varepsilon}[\eta](n) \in \mathcal{P}(\pi M^*)$ assigning probability weights to the ε-balls by

$$\rho^{\varepsilon}[\eta](n)(\pi \hat{v}_i^*) := \frac{\mu_{F,n}[\eta(1), \ldots, \eta(n)](L_n \in B(\varepsilon, \pi \hat{v}_i^*))}{\sum_{j=1}^{k} \mu_{F,n}[\eta(1), \ldots, \eta(n)](L_n \in B(\varepsilon, \pi \hat{v}_j^*))} \tag{33}$$

At this stage the measures appearing in the former definition involve a sum over $v \in M_n \cap B(\varepsilon, \pi \hat{v}_i^*)$ and for the correspondence formerly mentioned we can write

$$\rho^{\varepsilon}[\eta](n)(\pi \hat{v}_i^*) = \frac{\sum_{\hat{v}:\hat{\pi}_n \hat{v} \in B(\varepsilon, \pi \hat{v}_i^*)} \mu_{F,n}[\eta(1), \ldots, \eta(n)](\hat{L}_n = \hat{v})}{\sum_{j=1}^{k} \sum_{\hat{v}:\hat{\pi}_n \hat{v} \in B(\varepsilon, \pi \hat{v}_j^*)} \mu_{F,n}[\eta(1), \ldots, \eta(n)](\hat{L}_n = \hat{v})}$$

$$= \frac{\sum_{\hat{v}:\hat{\pi}_n \hat{v} \in B(\varepsilon, \pi \hat{v}_i^*)} \frac{\sum_{\sigma \in E^n : \hat{L}_n^{\sigma} = \hat{v}} e^{-nF(\hat{\pi}_n \hat{v})} \prod_{i=1}^{n} \alpha[\eta_i](\sigma_i)}{\sum_{\bar{v} \in \hat{M}_n} \sum_{\sigma \in E^n : \hat{L}_n^{\sigma} = \bar{v}} e^{-nF(\hat{\pi}_n \bar{v})} \prod_{i=1}^{n} \alpha[\eta_i](\bar{\sigma}_i)}}{\sum_{j=1}^{k} \sum_{\substack{\hat{v}: \\ \hat{\pi}_n \hat{v} \in B(\varepsilon, \pi \hat{v}_j^*)}} \frac{\sum_{\sigma \in E^n : \hat{L}_n^{\sigma} = \hat{v}} e^{-nF(\hat{\pi}_n \hat{v})} \prod_{i=1}^{n} \alpha[\eta_i](\sigma_i)}{\sum_{\bar{v} \in \hat{M}_n} \sum_{\sigma \in E^n : \hat{L}_n^{\sigma} = \bar{v}} e^{-nF(\hat{\pi}_n \bar{v})} \prod_{i=1}^{n} \alpha[\eta_i](\bar{\sigma}_i)}} \tag{34}$$

Decomposing the spin-sums into sums over possible values of the vector of empirical distributions on the b-like sites we can rewrite the last expression as

$$\frac{\sum_{\hat{v}:\hat{\pi}_n\hat{v}\in B(\varepsilon,\pi\hat{v}_i^*)} e^{-nF(\hat{\pi}_n\hat{v})} \prod_{b=1}^{|E'|} \alpha[b]^{|\Lambda_n(b)|}(\Omega_{|\Lambda_n(b)|}(\hat{v}(b)))}{\sum_{j=1}^{k} \sum_{\hat{v}:\hat{\pi}_n\hat{v}\in B(\varepsilon,\pi\hat{v}_j^*)} e^{-nF(\hat{\pi}_n\hat{v})} \prod_{b=1}^{|E'|} \alpha[b]^{|\Lambda_n(b)|}(\Omega_{|\Lambda_n(b)|}(\hat{v}(b)))} \tag{35}$$

where $\Omega_{|\Lambda_n(b)|}(\hat{v}(b)) = \{\sigma \in E^{|\Lambda_n(b)|} : \hat{L}_n^\sigma(b) = \hat{v}(b)\}$, and $\alpha[b]^{|\Lambda_n(b)|}(\cdot)$ is the product measure on the b-like sites. For sake of clarity let us recall the finite-volume finite-alphabet version of Sanov's theorem which is stated as Lemma 2.1.8 in [9], which we will make use of in the next step.

Lemma 2.4 *Let v be a probability measure on a finite state space E. For fixed n define the set of* microstates *compatible with v by*

$$\Omega(v) := \{\omega \in E^n | L_n^\omega = v\} \tag{36}$$

Then, if $nv(x)$ is integer-valued for all $x \in E$ we have the upper and lower large deviation bounds

$$(n+1)^{-|E|} e^{-nS(v|\mu)} \le \mu(\Omega(v)) = \mu(\{\omega : L_n^\omega = v\}) \le e^{-nS(v|\mu)} \tag{37}$$

Using (37) we get a lower bound to (35) of the form

$$\frac{1}{1 + \sum_{j\neq i}^{k} \dfrac{\sum_{\hat{v}:\hat{\pi}_n\hat{v}\in B(\varepsilon,\pi\hat{v}_j^*)} e^{-n\Phi[\hat{\pi}_n](\hat{v})}}{\sum_{\hat{v}:\hat{\pi}_n\hat{v}\in B(\varepsilon,\pi\hat{v}_i^*)} e^{-n\Phi[\hat{\pi}_n](\hat{v})} \prod_{b\in E'}(|\Lambda_n(b)|+1)^{-|E'|}}} \tag{38}$$

Let us notice that, in the last expression, the free energy (5) has appeared. However it does not involve yet the minimizer $\hat{v}_i^*$ in an explicit way. What we would like to do next, is to understand the $\hat{\pi}_n$-dependence of the minima in the different balls. Differences in the depths of the minima would not be present for $\hat{\pi}_n = \pi$ but will be created by the fluctuations of $\hat{\pi}_n$.

In order to achieve this we first need to compare the values that the $\hat{\pi}_n$-dependent free energy takes on the ball with the one corresponding to the center. As we will see in Proposition 2.5 this can be done uniformly with respect to the centers (π-minimizers). Secondly we will compare, for any fixed minimizer, the difference between the $\hat{\pi}_n$-dependent free energy and the π-dependent one; this will be done using the linearization procedure (6). Let us emphasize the fact that the definition of the *good-sets* $\mathcal{H}_{i,n,l}^\tau$ has been chosen *ad hoc* to guarantee, in the limit $n \uparrow \infty$, that the i-th stability vector will "dominate" the others, and thus the concentration around $\hat{v}_i^*$ will take place. We also need an upper bound on $\|\hat{\pi}_n - \pi\|$ for that procedure to work which is the reason for the cutoff in the definition of the *good-sets*.

The next proposition formalizes the first step.

Proposition 2.5 *Under the non-degeneracy assumption* (1) *there exists an $\varepsilon_0 > 0$ and a positive constant K such that for all $\varepsilon \le \varepsilon_0$ and for n sufficiently large*

$$\frac{-K\|\hat{\pi}_n - \pi\|^2}{2} \le \inf_{\hat{v}\in\hat{M}_n\cap B(\varepsilon,\pi\hat{v}_j^*)} (\Phi[\hat{\pi}_n](\hat{v}) - \Phi[\hat{\pi}_n](\hat{v}_j^*)) \tag{39}$$

for all minimizers $\hat{v}_j^*$.

Proof We will show that there exists a positive K such that

$$\frac{-K\|\hat{\pi}_n - \pi\|^2}{2} \leq \inf_{\hat{v} \in B(\varepsilon, \pi\hat{v}_j^*)} (\Phi[\hat{\pi}_n](\hat{v}) - \Phi[\hat{\pi}_n](\hat{v}_j^*)) \tag{40}$$

holds, hence the proposition will follow using the inequality $\inf_{\hat{v} \in \hat{M}_n \cap B(\varepsilon, \pi\hat{v}_j^*)} \leq \inf_{\hat{v} \in \hat{B}(\varepsilon, \pi\hat{v}_j^*)}$.

Let us take a Taylor expansion of $\Phi[\hat{\pi}_n](\hat{v})$ around $\hat{v}_j^*$, namely

$$\Phi[\hat{\pi}_n](\hat{v}) = \Phi[\hat{\pi}_n](\hat{v}_j^*) + \langle \nabla\Phi[\hat{\pi}_n](\hat{v}_j^*), \hat{v} - \hat{v}_j^* \rangle$$

$$+ \frac{1}{2}\langle \hat{v} - \hat{v}_j^*, H\Phi[\hat{\pi}_n](\hat{v}_j^*)(\hat{v} - \hat{v}_j^*) \rangle + \|\hat{v} - \hat{v}_j^*\|^2 R(\hat{v}, \hat{v}_j^*) \tag{41}$$

where $R(\hat{v}, \hat{v}_j^*)$ is a continuous function at $\hat{v} = \hat{v}_j^*$ with $R(\hat{v}_j^*, \hat{v}_j^*) = 0$, and H is the Hessian.
So we have

$$\Phi[\hat{\pi}_n](\hat{v}) - \Phi[\hat{\pi}_n](\hat{v}_j^*) \geq \langle \nabla\Phi[\hat{\pi}_n](\hat{v}_j^*), \hat{v} - \hat{v}_j^* \rangle + \frac{1}{2}\langle \hat{v} - \hat{v}_j^*, (H\Phi[\hat{\pi}_n](\hat{v}_j^*) - C_1)(\hat{v} - \hat{v}_j^*) \rangle \tag{42}$$

where C_1 is a non-negative constant which can be chosen arbitrarily close to zero when we restrict to balls with sufficiently small radii ε.

The inf for the previous r.h.s. is obtained at the point $\hat{v}^* = -(H\Phi[\hat{\pi}_n](\hat{v}_j^*) - C_1)^{-1} \times \nabla\Phi[\hat{\pi}_n](\hat{v}_j^*) + \hat{v}_j^*$.

So we have

$$\inf_{\hat{v} \in B(\varepsilon, \pi\hat{v}_j^*)} (\Phi[\hat{\pi}_n](\hat{v}) - \Phi[\hat{\pi}_n](\hat{v}_j^*)) \geq -\frac{1}{2}\langle \nabla\Phi[\hat{\pi}_n](\hat{v}_j^*), (H\Phi[\hat{\pi}_n](\hat{v}_j^*) - C_1)^{-1}\nabla\Phi[\hat{\pi}_n](\hat{v}_j^*) \rangle \tag{43}$$

Non-degeneracy assumption (1) implies together with the twice continuous differentiability of F that there exists a positive constant $\tilde{K}$ such that

$$\langle x, H\Phi[\xi](\hat{v}_j^*)x \rangle \geq \tilde{K}\|x\|^2 \tag{44}$$

for all ξ in a neighborhood of π. Noticing that $\|\nabla\Phi[\hat{\pi}_n](\hat{v}_j^*)\| \leq c\|\hat{\pi}_n - \pi\|$ we have

$$-\frac{K\|\hat{\pi}_n - \pi\|^2}{2} \leq \inf_{\hat{v} \in B(\varepsilon, \pi\hat{v}_j^*)} (\Phi[\hat{\pi}_n](\hat{v}) - \Phi[\hat{\pi}_n](\hat{v}_j^*)) \leq 0 \tag{45}$$

with $K = \frac{c^2}{\tilde{K} - C_1}$ which is positive for ε_0 sufficiently small. $\qquad\square$

From the last right-hand side of (38) we have

$$\frac{1}{1 + \sum_{j \neq i}^{k} \frac{\sum_{\hat{v}:\hat{\pi}_n\hat{v} \in B(\varepsilon, \pi\hat{v}_j^*)} e^{-n(\Phi[\hat{\pi}_n](\hat{v}) - \Phi[\hat{\pi}_n](\hat{v}_j^*))} e^{-n\Phi[\hat{\pi}_n](\hat{v}_j^*)}}{\sum_{\hat{v}:\hat{\pi}_n\hat{v} \in B(\varepsilon, \pi\hat{v}_i^*)} e^{-n(\Phi[\hat{\pi}_n](\hat{v}) - \Phi[\hat{\pi}_n](\hat{v}_i^*))} e^{-n\Phi[\hat{\pi}_n](\hat{v}_i^*)} \prod_{b \in E'}(|\Lambda_n(b)|+1)^{-|E'|}}}$$

$$\geq \frac{1}{1 + \sum_{j \neq i}^{k} \frac{\sum_{\hat{v}:\hat{\pi}_n\hat{v} \in B(\varepsilon, \pi\hat{v}_j^*)} e^{-n(\inf_{\hat{v} \in \hat{M}_n \cap B(\varepsilon, \pi\hat{v}_j^*)} (\Phi[\hat{\pi}_n](\hat{v}) - \Phi[\hat{\pi}_n](\hat{v}_j^*)))} e^{-n\Phi[\hat{\pi}_n](\hat{v}_j^*)}}{e^{-n(\Phi[\hat{\pi}_n](\hat{v}) - \Phi[\hat{\pi}_n](\hat{v}_i^*))} e^{-n\Phi[\hat{\pi}_n](\hat{v}_i^*)} \prod_{b \in E'}(|\Lambda_n(b)|+1)^{-|E'|}}}$$

$$\geq \frac{1}{1 + \sum_{j \neq i}^{k} \frac{\sum_{\hat{v}: \hat{\pi}_n \hat{v} \in B(\varepsilon, \pi \hat{v}_j^*)} e^{\frac{Kn\|\hat{\pi}_n - \pi\|^2}{2}} e^{-n\Phi[\hat{\pi}_n](\hat{v}_j^*)}}{e^{-C} e^{-n\Phi[\hat{\pi}_n](\hat{v}_i^*)} \prod_{b \in E'}(|\Lambda_n(b)|+1)^{-|E'|}}} \tag{46}$$

In the first inequality we have chosen $\tilde{v}$ as a best-approximation of $\hat{v}_i^*$ in $\hat{M}_n$ to get rid of the sum in the denominator of the denominator. In the second inequality we have used Proposition 2.5, and moreover the bound on the corresponding discretization error of the order $1/n$ and the uniform boundedness of the first derivative of Φ. The sums over measures in balls only give rise to polynomial constants which are swallowed by the terms in the exponential (as we will see, because the random terms lifting the degeneracy between the minimizers will be of order squareroot).

Now to the lowest order in $\hat{\pi}_n - \pi$, we have

$$\Phi[\hat{\pi}_n](\hat{v}_i^*) = \Phi[\pi](\hat{v}_i^*) + \Phi_\pi[\pi](\hat{v}_i^*)(\hat{\pi}_n - \pi) + o(\|\hat{\pi}_n - \pi\|) \tag{47}$$

So the last right-hand side of (46) becomes

$$\geq \frac{1}{1 + \sum_{j \neq i}^{k} e^{\frac{Kn\|\hat{\pi}_n - \pi\|^2}{2} + C} e^{-n\langle B_{\hat{v}_i^*} - B_{\hat{v}_j^*}, \hat{\pi}_n - \pi \rangle} e^{-n \cdot o(\|\hat{\pi}_n - \pi\|)} \prod_{b \in E'}(|\Lambda_n(b)|+1)^{2|E|}} \tag{48}$$

We are considering n sufficiently large such that there is at least one element in $\{\hat{v} : \hat{\pi}_n \hat{v} \in B(\varepsilon, \pi \hat{v}_i^*)\}$.

For $\eta \in \mathcal{H}_{i,n,0}^\tau$ we have that

$$\rho^\varepsilon[\eta](n)(\pi \hat{v}_i^*) > 1 - r(n) \tag{49}$$

Indeed we defined the *good-set* $\mathcal{H}_{i,n,0}^\tau$ in such a way that $n\|\hat{\pi}_n - \pi\|^2 \leq n^{\frac{\tau}{2}}$ and $n\langle B_{\hat{v}_i^*} - B_{\hat{v}_j^*}, \hat{\pi}_n - \pi \rangle \geq n^\tau$. Here we see the reason for the choice of the cutoff.

In order to prove Lemma 2.3 let us write

$$\mu_{F,n}[\eta(1), \dots, \eta(n)](L_n \in B(\varepsilon, \pi \hat{v}_i^*))$$
$$= \rho^\varepsilon[\eta](n)(\pi \hat{v}_i^*)(1 - \mu_{F,n}[\eta(1), \dots, \eta(n)](d(L_n, \pi M^*) \geq \varepsilon)) \tag{50}$$

Now we can use the concentration property for the empirical distribution saying that $\forall \varepsilon > 0$ and for all $\eta \in \mathcal{H}_{i,n,0}^\tau$ we have

$$\mu_{F,n}[\eta(1), \dots, \eta(n)](d(L_n, \pi M^*) \geq \varepsilon) \leq \hat{r}(n, \varepsilon) \tag{51}$$

with $\lim_{n \uparrow \infty} \hat{r}(n, \varepsilon) = 0$ for all positive ε.

This concentration property is a consequence of the bound

$$\mu_{F,n}[\eta(1), \dots, \eta(n)](d(L_n, \pi M^*) \geq \varepsilon)$$
$$\leq \prod_{b \in E'}(n\hat{\pi}_n(b)+1)^{2|E|} \exp\left(-n \inf_{\substack{\hat{v} \in \hat{M}_n: \\ d(\hat{\pi}_n \hat{v}, \pi M^*) \geq \varepsilon}} \Phi[\hat{\pi}_n](\hat{v}) + n \inf_{\hat{v}' \in \hat{M}_n} \Phi[\hat{\pi}_n](\hat{v}')\right)$$
$$\leq \prod_{b \in E'}(n\hat{\pi}_n(b)+1)^{2|E|} e^{K_2 n\|\hat{\pi}_n - \pi\| + C_2} \exp\left(-n \inf_{\substack{\hat{v} \in \mathcal{P}(E)^{|E'|}: \\ d(\pi \hat{v}, \pi M^*) \geq \varepsilon}} \Phi[\pi](\hat{v}) + n \inf_{\hat{v}' \in \mathcal{P}(E)^{|E'|}} \Phi[\pi](\hat{v}')\right)$$
$$\tag{52}$$

where in the second inequality we have used the Lipschitz property of Φ w.r.t. π and the control of the discretization error. On the *good-sets* we have $n\|\hat{\pi}_n - \pi\| \leq n^{\frac{1}{2}+\frac{\tau}{4}}$, while the quadratic nature of the minima gives us a term of exponential decay in n from the rightmost exponential, for any fixed $\varepsilon > 0$. This proves the concentration property. So the lemma follows from (49), (50) and (51). $\qquad\square$

Having proved, for a particular choice of the disorder variables, the concentration of the empirical distribution around a given minimizer, the following lemma represents the natural extension to averages.

Lemma 2.6 *For any real-valued continuous function g on $\mathcal{P}(E)$ the following holds*:

$$|\rho[\eta](n)(g) - g(\pi\hat{v}_j^*)| \leq \tilde{r}(n), \quad \forall \eta \in \mathcal{H}_{j,n,0}^{\tau} \tag{53}$$

where $\lim_{n\uparrow\infty} \tilde{r}(n) = 0$.

Proof Let $B(\varepsilon, \pi\hat{v}_j^*)$ be an ε-ball around the measure $\pi\hat{v}_j^*$. Then for any $\varepsilon > 0$ and integer n,

$$|\rho[\eta](n)(g) - g(\pi\hat{v}_j^*)|$$
$$= |\rho[\eta](n)(1_{B(\varepsilon,\pi\hat{v}_j^*)}(g - g(\pi\hat{v}_j^*))) + \rho[\eta](n)(1_{B^c(\varepsilon,\pi\hat{v}_j^*)}(g - g(\pi\hat{v}_j^*)))|$$
$$\leq \sup_{v\in B(\varepsilon,\pi\hat{v}_j^*)} |g(v) - g(\pi\hat{v}_j^*)| + 2\|g\|_\infty\rho[\eta](n)(B^c(\varepsilon, \pi\hat{v}_j^*))$$

holds. Choosing first ε sufficiently small and then n sufficiently large proves the lemma. $\square$

Now comes the study of how the probability of the *good-sets* $\mathcal{H}_{j,n,l}^{\tau}$ behaves in the limit $n \uparrow \infty$. Out of this analysis the weights (11) will arise. The fundamental step is that the limit will not depend on any finite number l of coordinates η, while the corresponding tail will provide, using CLT, the longed-for weights. This together with the Stone-Weierstrass theorem and Lemma 2.6 are the overriding tools for proving Theorem 2.2.

Let us start the analysis looking at the n, l-dependent *good-sets* $\mathcal{H}_{i,n,l}^{\tau}$ in a slightly different way.

For any $l < n$, we have

$$\mathcal{H}_{i,n,0}^{\tau} = \left\{\eta \in (E')^n : \frac{\sqrt{nl}}{n}X_{[1,l]}[\eta] + \frac{\sqrt{n(n-l)}}{n}X_{[l+1,n]}[\eta] \in R_{i,n}^{\tau}\right\} \tag{54}$$

Saying that $X_{[1,n]}[\eta] \in R_{i,n}^{\tau}$ it means

$$a_n\langle X_{[1,l]}[\eta], B_i\rangle + b_n\langle X_{[l+1,n]}[\eta], B_i\rangle - \max_{k\neq i}(a_n\langle X_{[1,l]}[\eta], B_k\rangle + b_n\langle X_{[l+1,n]}[\eta], B_k\rangle) > \delta_n \tag{55}$$

and $\|X_{[1,n]}[\eta]\| \leq n^{\frac{\tau}{4}}$, where $a_n = \frac{\sqrt{nl}}{n}$ and $b_n = \frac{\sqrt{n(n-l)}}{n}$.

Now let us define a subregion of $\mathcal{H}_{i,n,0}^{\tau}$, namely $\mathcal{H}_{i,n,0}^{\tau}(l)$ as follows

$$\mathcal{H}_{i,n,0}^{\tau}(l) := \left\{\eta \in (E')^n : a_n\langle X_{[1,l]}[\eta], B_i\rangle + b_n\langle X_{[l+1,n]}[\eta], B_i\rangle - \max_{k\neq i}(a_n\langle X_{[1,l]}[\eta], B_k\rangle)\right.$$

$$\left. - \max_{k\neq i}(b_n\langle X_{[l+1,n]}[\eta], B_k\rangle) > \delta_n, \text{ and } \|X_{[1,n]}[\eta]\| \leq n^{\frac{\tau}{4}}\right\} \tag{56}$$

Remark While $\mathcal{H}_{i,n,0}^{\tau}$ does not depend on l , $\mathcal{H}_{i,n,0}^{\tau}(l)$ does, indeed the partitioning might change the *max-value*.

It is worthwhile mentioning the following results.

Lemma 2.7 *For any integer l, $\mathbb{P}(\mathcal{H}_{i,n,0}^{\tau} \setminus \mathcal{H}_{i,n,0}^{\tau}(l))$ goes to zero in the limit $n \uparrow \infty$.*

Proof Note that

$$\mathcal{H}_{i,n,0}^{\tau} \setminus \mathcal{H}_{i,n,0}^{\tau}(l) \subseteq \left\{ \eta \in (E')^n : \max_{k \neq i}\langle a_n X_{[1,l]}[\eta], B_k\rangle + \max_{k \neq i}\langle b_n X_{[l+1,n]}[\eta], B_k\rangle + \delta_n \right.$$
$$\geq \langle a_n X_{[1,l]}[\eta] + b_n X_{[l+1,n]}[\eta], B_i\rangle$$
$$\left. > \max_{k \neq i}\langle a_n X_{[1,l]}[\eta] + b_n X_{[l+1,n]}[\eta], B_k\rangle + \delta_n \right\} \tag{57}$$

Now

$$\mathcal{H}_{i,n,0}^{\tau} \setminus \mathcal{H}_{i,n,0}^{\tau}(l) \subseteq \left\{ \eta : \frac{C(l)}{\sqrt{n}} + \max_{k \neq i}\langle b_n X_{[l+1,n]}[\eta], B_k\rangle + \delta_n \right.$$
$$\geq \langle a_n X_{[1,l]}[\eta] + b_n X_{[l+1,n]}[\eta], B_i\rangle$$
$$\left. > -\frac{C(l)}{\sqrt{n}} + \max_{k \neq i}\langle b_n X_{[l+1,n]}[\eta], B_k\rangle + \delta_n \right\} \tag{58}$$

where $C(l) = \sqrt{l} \max_{\eta} \max_k |\langle X_{[1,l]}[\eta], B_k\rangle| \leq \sqrt{l} \max_k \|B_k\|_{\infty}$.

The set on the right-hand side of (58) can be written as

$$\left\{ \eta : \frac{C(l)}{\sqrt{n}} + \max_{k \neq i}\langle b_n X_{[l+1,n]}[\eta], B_k\rangle + \delta_n \geq a_n\langle X_{[1,l]}[\eta], B_i\rangle + b_n\langle X_{[l+1,n]}[\eta], B_i\rangle \right\}$$
$$\cap \left\{ \eta : -\frac{C(l)}{\sqrt{n}} + \max_{k \neq i}\langle b_n X_{[l+1,n]}[\eta], B_k\rangle + \delta_n < a_n\langle X_{[1,l]}[\eta], B_i\rangle + b_n\langle X_{[l+1,n]}[\eta], B_i\rangle \right\}$$
$$= \left\{ \eta : \left(\frac{C(l)}{\sqrt{n}} - a_n\langle X_{[1,l]}[\eta], B_i\rangle + \delta_n\right)b_n^{-1} \geq \langle X_{[l+1,n]}[\eta], B_i\rangle - \max_{k \neq i}\langle X_{[l+1,n]}[\eta], B_k\rangle \right\}$$
$$\cap \left\{ \eta : \left(-\frac{C(l)}{\sqrt{n}} - a_n\langle X_{[1,l]}[\eta], B_i\rangle + \delta_n\right)b_n^{-1} \right.$$
$$\left. < \langle X_{[l+1,n]}[\eta], B_i\rangle - \max_{k \neq i}\langle X_{[l+1,n]}[\eta], B_k\rangle \right\}$$
$$\subset \left\{ \eta : \left(\frac{2C(l)}{\sqrt{n}} + \delta_n\right)b_n^{-1} \geq \langle X_{[l+1,n]}[\eta], B_i\rangle - \max_{k \neq i}\langle X_{[l+1,n]}[\eta], B_k\rangle \right\}$$
$$\cap \left\{ \eta : \left(-\frac{2C(l)}{\sqrt{n}} + \delta_n\right)b_n^{-1} < \langle X_{[l+1,n]}[\eta], B_i\rangle - \max_{k \neq i}\langle X_{[l+1,n]}[\eta], B_k\rangle \right\} \tag{59}$$

Let's define

$$\varphi^i(X_{[l+1,n]}[\eta]) := \langle X_{[l+1,n]}[\eta], B_i\rangle - \max_{k \neq i}\langle X_{[l+1,n]}[\eta], B_k\rangle \tag{60}$$

So we have

$$\mathbb{P}\big(\mathcal{H}^{\tau}_{i,n,0} \setminus \mathcal{H}^{\tau}_{i,n,0}(l)\big) \leq \mathbb{P}\left(\left\{\eta : \varphi^i(X_{[l+1,n]}[\eta]) \in b_n^{-1}\left(\delta_n - \frac{2C(l)}{\sqrt{n}}, \delta_n + \frac{2C(l)}{\sqrt{n}}\right)\right\}\right) \quad (61)$$

To take care of the n-dependence of the interval it's enough to notice that, $\forall \varepsilon > 0 \, \exists \bar{n}(\varepsilon)$ such that, for all $n > \bar{n}(\varepsilon)$ the following holds

$$b_n^{-1}\left(\delta_n - \frac{2C(l)}{\sqrt{n}}, \delta_n + \frac{2C(l)}{\sqrt{n}}\right) \subset (-\varepsilon, \varepsilon) \quad (62)$$

So

$$\lim_{n\uparrow\infty} \mathbb{P}\left(\left\{\eta : \varphi^i(X_{[l+1,n]}[\eta]) \in b_n^{-1}\left(\delta_n - \frac{2C(l)}{\sqrt{n}}, \delta_n + \frac{2C(l)}{\sqrt{n}}\right)\right\}\right)$$

$$\leq \lim_{n\uparrow\infty} \mathbb{P}(\{\eta : \varphi^i(X_{[l+1,n]}[\eta]) \in (-\varepsilon, \varepsilon)\}) \quad (63)$$

By the multidimensional CLT we have

$$\lim_{n\uparrow\infty} \mathbb{P}(\{\eta : \varphi^i(X_{[l+1,n]}[\eta]) \in (-\varepsilon, \varepsilon)\}) = \mathbb{P}_\pi(\varphi^i(G) \in (-\varepsilon, \varepsilon)) \quad (64)$$

where G is a centered Gaussian variable.

Taking the limit $\varepsilon \downarrow 0$ and using the *non-degeneracy assumption (2)*, the lemma is proved. $\qquad\square$

We have just seen that, for any fixed integer l, there is a subregion of the *good-set* which will not play any role in the limit $n \uparrow \infty$. We focus now on the probability of the main part of the *good-set*, especially on how its limit does not depend on any finite number of η-coordinates. Let us formalize the previous heuristic.

The condition (56) defining $\mathcal{H}^{\tau}_{i,n,0}(l)$ can also be written as

$$\langle X_{[l+1,n]}[\eta], B_i \rangle - \max_{k \neq i} \langle X_{[l+1,n]}[\eta], B_k \rangle > \frac{a_n}{b_n}\left(\max_{k \neq i} \langle X_{[1,l]}[\eta], B_k \rangle - \langle X_{[1,l]}[\eta], B_i \rangle\right) + b_n^{-1}\delta_n$$

$$\text{and} \quad \|X_{[1,n]}[\eta]\| \leq n^{\frac{\tau}{4}}. \quad (65)$$

Define the following sets

$$A^{\tau}_{i,n}(l) := \left\{\eta : \langle X_{[l+1,n]}[\eta], B_i \rangle - \max_{k \neq i} \langle X_{[l+1,n]}[\eta], B_k \rangle \right.$$

$$\left. > -\frac{a_n}{b_n} C_2(l) + b_n^{-1}\delta_n, \|X_{[l+1,n]}[\eta]\| \leq b_n^{-1} n^{\frac{\tau}{4}}\right\}$$

$$B^{\tau}_{i,n}(l) := \left\{\eta : \langle X_{[l+1,n]}[\eta], B_i \rangle - \max_{k \neq i} \langle X_{[l+1,n]}[\eta], B_k \rangle \right.$$

$$\left. > \frac{a_n}{b_n} C_2(l) + b_n^{-1}\delta_n, \|X_{[l+1,n]}[\eta]\| \leq b_n^{-1}(n^{\frac{\tau}{4}} - \tilde{C}_2(l))\right\} \quad (66)$$

where $C_2(l) = \max_\eta |\max_{k \neq i} \langle X_{[1,l]}[\eta], B_k \rangle - \langle X_{[1,l]}[\eta], B_i \rangle| \leq 2 \max_k \|B_k\|_\infty$, and $\tilde{C}_2(l) = \max_\eta \|X_{[1,l]}[\eta]\|$. These maxima give us the intended independence of the set from $\eta \in (E')^l$

and we have

$$B_{i,n}^{\tau}(l) = (E')^l \times \mathcal{H}_{i,n,l}^{1,\tau}$$

$$A_{i,n}^{\tau}(l) = (E')^l \times \mathcal{H}_{i,n,l}^{2,\tau}$$

(67)

where

$$\mathcal{H}_{i,n,l}^{1,\tau} = \left\{ \eta \in (E')^{n-l} : \varphi^i(X_{[l+1,n]}[\eta]) > b_n^{-1}(\delta_n + a_n C_2(l)), \right.$$

$$\left. \text{and } \|X_{[l+1,n]}[\eta]\| \le b_n^{-1}(n^{\frac{\tau}{4}} - \tilde{C}_2(l)) \right\}$$

$$\mathcal{H}_{i,n,l}^{2,\tau} = \left\{ \eta \in (E')^{n-l} : \varphi^i(X_{[l+1,n]}[\eta]) > b_n^{-1}(\delta_n - a_n C_2(l)), \text{ and } \|X_{[l+1,n]}[\eta]\| \le b_n^{-1} n^{\frac{\tau}{4}} \right\}$$

(68)

The following holds

$$A_{i,n}^{\tau}(l) \supseteq \mathcal{H}_{i,n,0}^{\tau}(l) \supseteq B_{i,n}^{\tau}(l)$$

$$\mathcal{H}_{i,n,l}^{1,\tau} \subseteq \mathcal{H}_{i,n,l}^{2,\tau}$$

(69)

Lemma 2.8 *For any integer l, $\mathbb{P}(\mathcal{H}_{i,n,l}^{2,\tau} \setminus \mathcal{H}_{i,n,l}^{1,\tau})$ goes to zero in the limit $n \uparrow \infty$.*

Proof

$$\mathcal{H}_{i,n,l}^{2,\tau} \setminus \mathcal{H}_{i,n,l}^{1,\tau} \subseteq \left\{ \eta : \frac{a_n}{b_n} C_2(l) + b_n^{-1}\delta_n \ge \langle X_{[l+1,n]}[\eta], B_i \rangle - \max_{k \ne i} \langle X_{[l+1,n]}[\eta], B_k \rangle \right.$$

$$\left. > -\frac{a_n}{b_n} C_2(l) + b_n^{-1}\delta_n \right\}$$

$$= \left\{ \eta : b_n^{-1}(\delta_n + a_n C_2(l)) \ge \varphi^i(X_{[l+1,n]}[\eta]) > b_n^{-1}(\delta_n - a_n C_2(l)) \right\} \quad (70)$$

By the same argument we have used in Lemma 2.7, we have $\mathbb{P}(\mathcal{H}_{i,n,l}^{2,\tau} \setminus \mathcal{H}_{i,n,l}^{1,\tau}) \longrightarrow 0$ in the limit $n \uparrow \infty$. □

Lemma 2.9 *For any integer l, $\lim_{n\uparrow\infty} \mathbb{P}(\mathcal{H}_{i,n,l}^{1,\tau}) = \mathbb{P}_{\pi}(G \in R_i)$ where $G \sim \mathcal{N}(0, \Sigma)$.*

Proof From the previous lemma we know that

$$\lim_{n\uparrow\infty} \mathbb{P}(\mathcal{H}_{i,n,0}^{\tau}(l)) = \lim_{n\uparrow\infty} \mathbb{P}(B_{i,n}^{\tau}(l)) \quad \text{and}$$

$$\lim_{n\uparrow\infty} \mathbb{P}(\mathcal{H}_{i,n,l}^{2,\tau}) = \lim_{n\uparrow\infty} \mathbb{P}(\mathcal{H}_{i,n,l}^{1,\tau})$$

(71)

Now $\forall \varepsilon > 0 \; \exists n_0(\varepsilon)$ such that for all $n > n_0(\varepsilon)$ the following holds

$$\gamma_{-\varepsilon}^n \supset \mathcal{H}_{i,n,l}^{1,\tau} \supset \gamma_{\varepsilon}^n$$

(72)

where $\gamma_{\varepsilon}^n = \{\eta : \varphi^i(X_{[l+1,n]}[\eta]) > \varepsilon\}$. Therefore

$$\lim_{n\uparrow\infty} \mathbb{P}(\gamma_{-\varepsilon}^n) \ge \lim_{n\uparrow\infty} \mathbb{P}(\mathcal{H}_{i,n,l}^{1,\tau}) \ge \lim_{n\uparrow\infty} \mathbb{P}(\gamma_{\varepsilon}^n)$$

(73)

Applying the CLT to both the right and the left-hand side and taking the limit for $\varepsilon \downarrow 0$ we have

$$\lim_{n \uparrow \infty} \mathbb{P}\left(\mathcal{H}_{i,n,l}^{1,\tau}\right) = \mathbb{P}_\pi(G \in R_i) \tag{74}$$

where $G \sim \mathcal{N}(0, \Sigma)$. $\qquad\square$

Let us now summarize what we have done above for the decompositions of the various regions of the η-configuration space.

$$1_{\mathcal{H}_{i,n,0}^\tau} = 1_{\mathcal{H}_{i,n,0}^\tau(l)} + 1_{\mathcal{H}_{i,n,0}^\tau \setminus \mathcal{H}_{i,n,0}^\tau}(l)$$

$$1_{B_{i,n}^\tau(l)} = 1_{(E')^l} 1_{\mathcal{H}_{i,n,l}^{1,\tau}} \tag{75}$$

$$1_{A_{i,n}^\tau(l)} = 1_{(E')^l} 1_{\mathcal{H}_{i,n,l}^{2,\tau}}$$

To state our next result let us fix one more notation.

We let Ψ be a continuous real-valued function on $\mathcal{P}(\mathcal{P}(E)) \times (E')^m$, for some positive integer m.

Lemma 2.10 *Suppose Ψ is as above. Then under the* non-degeneracy assumptions (1) *and* (2) *the following holds*:

$$\lim_{n \uparrow \infty} \int_{\mathcal{H}_{i,n,0}^\tau} \mathbb{P}_\pi(d\eta) \Psi(\rho[\eta](n), \eta) = w_i \int_{(E')^m} \pi^{\otimes m}(d\eta) \Psi(\delta_{\pi \hat{v}_i}, \eta), \tag{76}$$

where $\pi^{\otimes m}(d\eta) = \prod_{k=1}^m \pi(d\eta_k)$, $w_i = \mathbb{P}(G \in R_i)$ *with* $G \sim \mathcal{N}(0, \Sigma)$.

Proof Set $l = m$,

$$\int_{\mathcal{H}_{i,n,0}^\tau} \mathbb{P}_\pi(d\eta) \Psi(\rho[\eta](n), \eta) = \int_{\mathcal{H}_{i,n,0}^\tau} \mathbb{P}_\pi(d\eta)(\Psi(\rho[\eta](n), \eta) - \Psi(\delta_{\pi \hat{v}_i}, \eta))$$

$$+ \int_{\mathcal{H}_{i,n,0}^\tau} \mathbb{P}_\pi(d\eta) \Psi(\delta_{\pi \hat{v}_i}, \eta) \tag{77}$$

We can assume that Ψ is of the form $\Psi(\rho, \eta) = \tilde{\Psi}(\rho(g_1), \ldots, \rho(g_l), \eta_{[1,m]})$ for a finite l with continuous and bounded g_i's, and continuous $\tilde{\Psi}$. So, together with Lemma 2.6 we have that the first term in the left-hand side is going to 0 in the limit $n \uparrow \infty$.

Now from the first equality of (75)

$$\int_{\mathcal{H}_{i,n,0}^\tau} \mathbb{P}_\pi(d\eta) \Psi(\delta_{\pi \hat{v}_i}, \eta) = \int_{\mathcal{H}_{i,n,0}^\tau(l)} \mathbb{P}_\pi(d\eta) \Psi(\delta_{\pi \hat{v}_i}, \eta) + \int_{\mathcal{H}_{i,n,0}^\tau \setminus \mathcal{H}_{j,n,0}^\tau(l)} \mathbb{P}_\pi(d\eta) \Psi(\delta_{\pi \hat{v}_i}, \eta).$$

Under the *non-degeneracy assumption (2)* the second term on the right-hand side of the above equation plays no role in the limit, indeed

$$\int_{\mathcal{H}_{i,n,0}^\tau \setminus \mathcal{H}_{i,n,0}^\tau(l)} \mathbb{P}_\pi(d\eta) \Psi(\delta_{\pi \hat{v}_i}, \eta) \leq \|\Psi\|_\infty \mathbb{P}\left(\mathcal{H}_{i,n,0}^\tau \setminus \mathcal{H}_{i,n,0}^\tau(l)\right) \tag{78}$$

and from Lemma 2.7 $\mathbb{P}(\mathcal{H}_{i,n,0}^\tau \setminus \mathcal{H}_{i,n,0}^\tau(l))$ goes to zero in the limit $n \uparrow \infty$.

Observe from the first inclusion relation of (69) that

$$\int_{B_{i,n}^\tau(l)} \mathbb{P}_\pi(d\eta)\Psi(\delta_{\pi\hat{v}_i},\eta) \le \int_{\mathcal{H}_{i,n,0}^\tau(l)} \mathbb{P}_\pi(d\eta)\Psi(\delta_{\pi\hat{v}_i},\eta) \le \int_{A_{i,n}^\tau(l)} \mathbb{P}_\pi(d\eta)\Psi(\delta_{\pi\hat{v}_i},\eta) \quad (79)$$

Next, observe from (67) that

$$\int_{(E')^l} \pi^{\otimes l}(d\eta)\Psi(\delta_{\pi\hat{v}_i},\eta) \int_{\mathcal{H}_{i,n,l}^{1,\tau}} \pi^{\otimes n-l}(d\eta) \le \int_{\mathcal{H}_{i,n,0}^\tau(l)} \mathbb{P}_\pi(d\eta)\Psi(\delta_{\pi\hat{v}_i},\eta)$$

$$\int_{\mathcal{H}_{i,n,0}^\tau(l)} \mathbb{P}_\pi(d\eta)\Psi(\delta_{\pi\hat{v}_i},\eta) \le \int_{(E')^l} \pi^{\otimes l}(d\eta)\Psi(\delta_{\pi\hat{v}_i},\eta) \int_{\mathcal{H}_{i,n,l}^{2,\tau}} \pi^{\otimes n-l}(d\eta) \tag{80}$$

Taking the limit $n \uparrow \infty$ we obtain

$$\lim_{n\uparrow\infty} \int_{\mathcal{H}_{i,n,0}^\tau(l)} \mathbb{P}_\pi(d\eta)\Psi(\delta_{\pi\hat{v}_i},\eta) = \int_{(E')^l} \pi^{\otimes l}(d\eta)\Psi(\delta_{\pi\hat{v}_i},\eta) \lim_{n\uparrow\infty} \int_{\mathcal{H}_{i,n,l}^{2,\tau}} \pi^{\otimes n-l}(d\eta) \tag{81}$$

and using Lemma 2.9 we are done. $\qquad\square$

Now we have provided all the ingredients, and so the proof of Theorem 2.2 is straightforward.

$$\int \mathbb{P}_\pi(d\eta)\Psi(\rho[\eta](n),\eta) = \sum_{i=1}^k \int \mathbb{P}_\pi(d\eta)\Psi(\rho[\eta](n),\eta) 1_{\mathcal{H}_{i,n,0}^\tau}(\eta)$$

$$+ \int \mathbb{P}_\pi(d\eta)\Psi(\rho[\eta](n),\eta) 1_{(\mathcal{H}_{n,0}^\tau)^c}(\eta) \tag{82}$$

Clearly for bounded Ψ one has

$$\left| \int \mathbb{P}_\pi(d\eta)\Psi(\rho[\eta](n),\eta) 1_{(\mathcal{H}_{n,0}^\tau)^c}(\eta) \right| \le \|\Psi\|_\infty \mathbb{P}_\pi\left(\left(\mathcal{H}_{n,0}^\tau\right)^c(\eta) \right) \tag{83}$$

and the *non-degeneracy assumption (2)* tells us that this term will not play any role in the limit $n \uparrow \infty$. For every summand of the first term, by Lemma 2.10 we have

$$\lim_{n\uparrow\infty} \int_{\mathcal{H}_{i,n,0}^\tau} \mathbb{P}_\pi(d\eta)\Psi(\rho[\eta](n),\eta) = w_i \int_{(E')^m} \pi^{\otimes m}(d\eta)\Psi(\delta_{\pi\hat{v}_i},\eta), \tag{84}$$

where $w_i = \mathbb{P}(G \in R_i)$ with $G \sim \mathcal{N}(0,\Sigma)$.

Therefore

$$\lim_{n\uparrow\infty} \int \mathbb{P}_\pi(d\eta)\Psi(\rho[\eta](n),\eta) = \sum_{i=1}^k \int \mathbb{P}_\pi(d\eta)\Psi(\rho,\eta) w_i \delta_{\delta_{\pi\hat{v}_i}}(d\rho) \tag{85}$$

Looking now at the definition of the AW-metastate, we can identify the joint distribution K we are interested in

$$K(d\rho, d\eta) = \sum_{i=1}^{k} \mathbb{P}_\pi(d\eta) w_i \delta_{\delta_{\pi\hat{v}_i}}(d\rho)$$

$$\implies \quad K(d\rho|\eta) = \sum_{i=1}^{k} w_i \delta_{\delta_{\pi\hat{v}_i}}(d\rho) \tag{86}$$

$\square$

3 The Metastate on the Level of States

Let us go from the global perspective (talking about the empirical mean) to the local view (talking about finitely many variables $\sigma_1, \ldots, \sigma_k$). In different words, we are fixing a sub-population of finite size, and we are asking how it will behave when we couple it to a large system whose size n will be let tend to infinity. Let us introduce a metric on the space of probability measures $\mu, \mu' \in \mathcal{P}(E^\infty)$ by

$$d(\mu, \mu') = \sum_{i=1}^{\infty} 2^{-i} \|\mu - \mu'\|_i \tag{87}$$

where

$$\|\mu - \mu'\|_i := \frac{1}{2} \sum_{\omega_1, \ldots, \omega_i} |\mu(\omega_1, \ldots, \omega_i) - \mu'(\omega_1, \ldots, \omega_i)| \tag{88}$$

is the total variation norm of the restriction of the measure to the first i coordinates.

The statement about the metastate promised in the main theorem implies in particular that, for all $\varepsilon > 0$

$$\lim_{n\uparrow\infty} \mathbb{P}(d(\mu_{F,n}[\eta], \text{ext}(\mathcal{G}[\eta])) > \varepsilon) = 0 \tag{89}$$

where

$$\mathcal{G}[\eta] = \left\{ \sum_{\hat{v} \in M^*} p_{\hat{v}} \mu_{\hat{v}}[\eta], \ p_{\hat{v}} \in \mathcal{P}(\pi M^*) \right\} \tag{90}$$

and $\mu_{\hat{v}}[\eta](\cdot) = \prod_{i=1}^{\infty} \gamma[\eta(i)](\cdot | \pi\hat{v})$. Throughout this section we identify $\mu_{F,n}[\eta]$ with the infinite-volume measure which is obtained by tensorization with the equidistribution for sites outside of $\{1, \ldots, n\}$.

We will in fact prove that

$$\lim_{n\uparrow\infty} \sup_{\eta \in \mathcal{H}_{i,n,0}^\tau} d(\mu_{F,n}[\eta], \mu_{\hat{v}_i^*}[\eta]) = 0 \tag{91}$$

where $\mathcal{H}_{i,n,0}^\tau$ are the disorder sets ensuring the dominance of the i-th minimizer. Let us remark that it can not be expected in general that $\lim_{n\uparrow\infty} d(\mu_{F,n}[\eta], \text{ext}(\mathcal{G}[\eta])) = 0$ for $\mathbb{P}$-a.e. η holds, as already the example of the random field Ising model discussed in [13] shows, due to the empirical distribution $\hat{\pi}_n$ passing regions of "ties" outside of the good sets infinitely often.

So we are about to prove that the possible limiting distributions will be product measures of a particular sort. These limiting measures will depend on which region of the disorder variables we are restricting ourselves to.

Let us look at the k-marginal

$$\mu_{F,n}[\eta](\sigma_1, \ldots, \sigma_k) = \sum_{\omega_{k+1}, \ldots, \omega_n} \mu_{F,n}[\eta](\sigma_1, \ldots, \sigma_k, \omega_{k+1}, \ldots, \omega_n)$$

$$= \sum_{\omega_{k+1}, \ldots, \omega_n} \frac{e^{-nF(L_n^{\sigma_{[1,k]}, \omega_{[k+1,n]}})} \prod_{i=1}^k \alpha[\eta_i](\sigma_i) \prod_{j=k+1}^n \alpha[\eta_j](\omega_j)}{\sum_{\bar{\sigma} \in E^n} e^{-nF(L_n^{\bar{\sigma}})} \prod_{i=1}^n \alpha[\eta_i](\bar{\sigma}_i)} \quad (92)$$

Let us now introduce the suitable decomposition of the empirical distribution, obtained by dividing the volume $\{1, \ldots, n\}$ in two subvolumes $\{1, \ldots, k\}$ and $\{k+1, \ldots, n\}$, where k is the size of the marginal we are considering; then we focus on the respective b-like sites for both of the subvolumes.

$$L_n^{\sigma_{[1,k]}, \omega_{[k+1,n]}} = \frac{k}{n} \frac{1}{k} \sum_{i=1}^k \delta_{\sigma_i} + \frac{n-k}{n} \frac{1}{n-k} \sum_{i=k+1}^n \delta_{\omega_i}$$

$$= \frac{k}{n} \sum_{b \in E'} \hat{\pi}_{[1,k]}(b) \hat{L}_{[1,k]}(b) + \frac{n-k}{n} \sum_{b \in E'} \hat{\pi}_{[k+1,n]}(b) \hat{L}_{[k+1,n]}(b) \quad (93)$$

In the process to carry out (93), we have also made use of the following definitions

$$\Lambda_{[1,k]}(b) = \{i \in \{1, \ldots, k\} : \eta(i) = b\}$$

$$\Lambda_{[k+1,n]}(b) = \{i \in \{k+1, \ldots, n\} : \eta(i) = b\} \quad (94)$$

$$\hat{\pi}_{[1,k]}(b) = \frac{|\Lambda_{[1,k]}(b)|}{k}, \qquad \hat{\pi}_{[k+1,n]}(b) = \frac{|\Lambda_{[k+1,n]}(b)|}{n-k}$$

Proof of Theorem 1.7 Let us start providing the key result, namely the *weak convergence* of the measure $\mu_{F,n}[\eta]$ to $\mu_j[\eta] = \prod_{i=1}^\infty \gamma[\eta(i)](\cdot | \pi \hat{v}_j)$ conditional on the suitable region of the disorder.

The following lemma is the short view (local topology) version of Lemma 2.6.

Lemma 3.1 *For any event A which depends only on the first k coordinates the following holds*

$$|\mu_{F,n}[\eta](A) - \mu_j[\eta](A)| \leq \tilde{r}(n) \quad \forall \eta \in \mathcal{H}_{j,n,k}^\tau \quad (95)$$

where $\lim_{n \uparrow \infty} \tilde{r}(n) = 0$.

Proof It suffices to consider the event A which fixes the first k coordinates and write

$$\mu_{F,n}[\eta](\sigma_1, \ldots, \sigma_k)$$

$$= \sum_{\omega_{k+1}, \ldots, \omega_n} \frac{e^{-nF(\frac{k}{n} \sum_{b \in E'} \hat{\pi}_{[1,k]}(b) \hat{L}_{[1,k]}^\sigma(b) + \frac{n-k}{n} \sum_{b \in E'} \hat{\pi}_{[k+1,n]}(b) \hat{L}_{[k+1,n]}^\omega(b))} \prod_{i=1}^k \alpha[\eta_i](\sigma_i) \prod_{j=k+1}^n \alpha[\eta_j](\omega_j)}{\sum_{\bar{\sigma} \in E^n} e^{-nF(L_n^{\bar{\sigma}})} \prod_{i=1}^n \alpha[\eta_i](\bar{\sigma}_i)}$$

$$= \frac{\sum_{\hat{v} \in \hat{M}_{[k+1,n]}} \sum_{\substack{\omega_{k+1},\ldots,\omega_n: \\ \hat{L}^{\tilde{\omega}}_{[k+1,n]}(\cdot)=\hat{v}(\cdot)}} e^{-nF(\frac{k}{n}\langle \hat{\pi}_{[1,k]}, \hat{L}^{\sigma}_{[1,k]}\rangle + \frac{n-k}{n}\langle \hat{\pi}_{[k+1,n]}, \hat{v}\rangle)} \prod_{i=1}^k \alpha[\eta_i](\sigma_i) \prod_{j=k+1}^n \alpha[\eta_j](\omega_j)}{\sum_{\tilde{v} \in \hat{M}_n} \sum_{\substack{\tilde{\sigma} \in E^n: \\ \hat{L}^{\tilde{\sigma}}_n = \tilde{v}}} e^{-nF(L^{\tilde{\sigma}}_n)} \prod_{i=1}^n \alpha[\eta_i](\tilde{\sigma}_i)} \tag{96}$$

where we have introduced the following space

$$\hat{M}_{[k+1,n]} := \left\{ \hat{v} \in \mathcal{P}(E)^{E'} : \exists \tilde{\omega} = (\tilde{\omega}_{k+1}, \ldots, \tilde{\omega}_n) : \hat{L}^{\tilde{\omega}}_{[k+1,n]}(b) = \hat{v}(b), \forall b \in E' \right\} \tag{97}$$

Using the partition induced by the disorder variables η on the sub-volume $\{k+1, \ldots, n\}$, we have

$$\sum_{\substack{\omega_{k+1},\ldots,\omega_n: \\ \hat{L}^{\omega}_{[k+1,n]}(\cdot)=\hat{v}(\cdot)}} \prod_{j=k+1}^n \alpha[\eta_j](\omega_j) = \prod_{b \in E'} \alpha[b]^{|\Lambda_{[k+1,n]}(b)|}(\Omega_{|\Lambda_{[k+1,n]}(b)|}(\hat{v}(b))) \tag{98}$$

and with this we obtain

$$\mu_{F,n}[\eta](\sigma_1, \ldots, \sigma_k)$$

$$= \frac{\prod_{i=1}^k \alpha[\eta_i](\sigma_i) \sum_{\hat{v} \in \hat{M}_{[k+1,n]}} e^{-nF(\frac{k}{n}\langle \hat{\pi}_{[1,k]}, \hat{L}^{\sigma}_{[1,k]}\rangle + \frac{n-k}{n}\langle \hat{\pi}_{[k+1,n]}, \hat{v}\rangle)} \prod_{b \in E'} \alpha[b]^{|\Lambda_{[k+1,n]}(b)|}(\Omega_{|\Lambda_{[k+1,n]}(b)|}(\hat{v}(b)))}{\sum_{v' \in \hat{M}_{[1,k]}} \sum_{\hat{v} \in \hat{M}_{[k+1,n]}} \sum_{\substack{\tilde{\sigma} \in E^k: \\ \hat{L}^{\tilde{\sigma}}_{[1,k]}=v'}} e^{-nF(\frac{k}{n}\langle \hat{\pi}_{[1,k]}, \hat{L}^{\tilde{\sigma}}_{[1,k]}\rangle + \frac{n-k}{n}\langle \hat{\pi}_{[k+1,n]}, \hat{v}\rangle)} \prod_{i=1}^k \alpha[\eta_i](\tilde{\sigma}_i) \sum_{\substack{\tilde{\sigma} \in E^{n-k}: \\ \hat{L}^{\tilde{\sigma}}_{[k+1,n]}=v'}} \prod_{j=k+1}^n \alpha[\eta_j](\tilde{\sigma}_j)} \tag{99}$$

where naturally $\prod_{b \in E'} \alpha[b]^{|\Lambda_{[k+1,n]}(b)|}(\Omega_{|\Lambda_{[k+1,n]}(b)|}(\hat{v}(b))) = \sum_{\substack{\tilde{\sigma} \in E^{n-k}: \\ \hat{L}^{\tilde{\sigma}}_{[k+1,n]}=v'}} \prod_{j=k+1}^n \alpha[\eta_j](\tilde{\sigma}_j)$.

Now multiplying and dividing, both numerator and denominator, by $e^{-nF(\langle \hat{\pi}_{[k+1,n]}, \hat{v}\rangle)}$ we arrive at:

$$\mu_{F,n}[\eta](\sigma_1, \ldots, \sigma_k)$$

$$= \frac{\prod_{i=1}^k \alpha[\eta_i](\sigma_i) \sum_{\hat{v} \in \hat{M}_{[k+1,n]}} e^{-n[F(\frac{k}{n}\langle \hat{\pi}_{[1,k]}, \hat{L}^{\sigma}_{[1,k]}\rangle + \frac{n-k}{n}\langle \hat{\pi}_{[k+1,n]}, \hat{v}\rangle) - F(\langle \hat{\pi}_{[k+1,n]}, \hat{v}\rangle)]} \rho_{F,n,k}[\eta](\hat{v})}{\sum_{v' \in \hat{M}_{[1,k]}} \sum_{\hat{v} \in \hat{M}_{[k+1,n]}} \sum_{\substack{\tilde{\sigma} \in E^k: \\ \hat{L}^{\tilde{\sigma}}_{[1,k]}=v'}} e^{-n[F(\frac{k}{n}\langle \hat{\pi}_{[1,k]}, \hat{L}^{\tilde{\sigma}}_{[1,k]}\rangle + \frac{n-k}{n}\langle \hat{\pi}_{[k+1,n]}, \hat{v}\rangle) - F(\langle \hat{\pi}_{[k+1,n]}, \hat{v}\rangle)]} \prod_{i=1}^k \alpha[\eta_i](\tilde{\sigma}_i) \rho_{F,n,k}[\eta](\hat{v})} \tag{100}$$

where $\rho_{F,n,k}[\eta] \in \mathcal{P}(\hat{M}_{[k+1,n]})$ is defined as

$$\rho_{F,n,k}[\eta](\hat{v}) := \frac{e^{-nF(\langle \hat{\pi}_{[k+1,n]}, \hat{v}\rangle)} \prod_{b \in E'} \alpha[b]^{|\Lambda_{[k+1,n]}(b)|}(\Omega_{|\Lambda_{[k+1,n]}(b)|}(\hat{v}(b)))}{\sum_{\tilde{v} \in \hat{M}_{[k+1,n]}} e^{-nF(\langle \hat{\pi}_{[k+1,n]}, \hat{v}\rangle)} \prod_{b \in E'} \alpha[b]^{|\Lambda_{[k+1,n]}(b)|}(\Omega_{|\Lambda_{[k+1,n]}(b)|}(\tilde{v}(b)))} \tag{101}$$

Note that the measure $\rho_{F,n,k}[\eta]$ depends on the disorder variables just in the subvolume $\{k+1, \ldots, n\}$.

Recall that a function $F : \mathcal{P}(E) \to \mathbb{R}$ is differentiable if, for all $\alpha \in \mathcal{P}(E)$ there is a linear map $dF_\alpha : T(\mathcal{P}(E)) \to \mathbb{R}$ on the tangent space such that

$$F(\alpha') = F(\alpha) + dF_\alpha(\alpha' - \alpha) + \|\alpha' - \alpha\| r(\alpha', \alpha) \tag{102}$$

where $\alpha' \to r(\alpha', \alpha)$ is continuous at $\alpha' = \alpha$ with $r(\alpha, \alpha) = 0$. Then uniformly in α, α' we have

$$\sup_{\alpha, \alpha'} |F(\alpha + p(\alpha' - \alpha)) - F(\alpha) - pdF_\alpha(\alpha' - \alpha)| \leq Cpr(p) \tag{103}$$

where $r(p) \downarrow 0$ with $p \downarrow 0$. The uniformity in α, α' follows by the compactness of $\mathcal{P}(E)$. In our set up we will have

$$\left| F\left(\frac{k}{n}L_{[1,k]}(\sigma) + \frac{n-k}{n}\nu\right) - F(\nu) - dF_\nu \frac{k}{n}(L_{[1,k]}(\sigma) - \nu) \right| \leq C\frac{k}{n}r\left(\frac{k}{n}\right) \qquad (104)$$

where we have set $\nu = \langle \hat{\pi}_{[k+1,n]}, \hat{\nu} \rangle$. This gives, recognizing that $\langle \hat{\pi}_{[1,k]}, \hat{L}^\sigma_{[1,k]} \rangle$ and ν are both elements in $\mathcal{P}(E)$, the upper bound

$$\mu_{F,n}[\eta](\sigma_1, \ldots, \sigma_k)$$

$$\leq \frac{e^{2Ckr(\frac{k}{n})} \prod_{i=1}^k \alpha[\eta_i](\sigma_i) \sum_{\hat{\nu} \in \hat{M}_{[k+1,n]}} \prod_{i=1}^k e^{-dF_{\langle \hat{\pi}_{[k+1,n]}, \hat{\nu} \rangle}(\delta_{\sigma_i} - \langle \hat{\pi}_{[k+1,n]}, \hat{\nu} \rangle)} \rho_{F,n,k}[\eta](\hat{\nu})}{\sum_{\hat{\nu} \in \hat{M}_{[k+1,n]}} \rho_{F,n,k}[\eta](\hat{\nu}) \prod_{i=1}^k \sum_{\bar{\sigma}_i \in E} \alpha[\eta_i](\bar{\sigma}_i) e^{-dF_{\langle \hat{\pi}_{[k+1,n]}, \hat{\nu} \rangle}(\delta_{\bar{\sigma}_i} - \langle \hat{\pi}_{[k+1,n]}, \hat{\nu} \rangle)}} \qquad (105)$$

and the corresponding lower bound which is obtained from the last r.h.s. by replacing $C > 0$ by $-C$. The measure $\rho_{F,n,k}[\eta]$ can be written in the form

$$\rho_{F,n,k}[\eta](\hat{\nu}) = \frac{\rho_{F,n-k}[\eta](\hat{\nu}) e^{-kF(\langle \hat{\pi}_{[k+1,n]}, \hat{\nu} \rangle)}}{\sum_{\bar{\nu} \in \hat{M}_{[k+1,n]}} \rho_{F,n-k}[\eta](\bar{\nu}) e^{-kF(\langle \hat{\pi}_{[k+1,n]}, \bar{\nu} \rangle)}} \qquad (106)$$

where we have recovered the proper random mean-field measure on the empirical distribution of size $n - k$. Note once again that $\langle \hat{\pi}_{[k+1,n]}, \hat{\nu} \rangle \in \mathcal{P}(E)$, and when $\hat{\nu}$ moves in $\hat{M}_{[k+1,n]}$ the corresponding measure $\langle \hat{\pi}_{[k+1,n]}, \hat{\nu} \rangle$ is moving among the possible empirical measures of size $n - k$. Let us write $C(A, \varepsilon) := \{\nu \in \mathcal{P}(E)^{E'} : d(\nu, A) \leq \varepsilon\}$, for the ε-ball of a set A. By definition $\rho_{F,n}[\eta]$ is said to concentrate on the set A iff $\rho_{F,n}[\eta](C(A, \varepsilon)^c) \downarrow 0$, $\forall \varepsilon > 0$. So whenever $\rho_{F,n}[\eta]$ concentrates on a finite set, so does $\rho_{F,n,k}[\eta]$, by the boundedness of F.

We remark that $\rho_{F,n-k}[\eta](\hat{\nu}) = \mu_{F,n-k}[\eta](L_{[k+1,n]} = \langle \hat{\pi}_{[k+1,n]}, \hat{\nu} \rangle)$ has the property of concentrating around the minimizers, as we know from the analysis of the previous chapter.

To study the bound (105) and the corresponding lower bound let's introduce the quantities

$$\xi_N = \sum_{\hat{\nu} \in \hat{M}_{[k+1,n]}} \rho_{F,n,k}[\eta](\hat{\nu}) \prod_{i=1}^k \alpha[\eta_i](\sigma_i) \prod_{i=1}^k e^{-dF_{\langle \hat{\pi}_{[k+1,n]}, \hat{\nu} \rangle}(\delta_{\sigma_i} - \langle \hat{\pi}_{[k+1,n]}, \hat{\nu} \rangle)}$$

$$\xi_D = \sum_{\hat{\nu} \in \hat{M}_{[k+1,n]}} \rho_{F,n,k}[\eta](\hat{\nu}) \prod_{i=1}^k \sum_{\bar{\sigma}_i \in E} \alpha[\eta_i](\bar{\sigma}_i) \prod_{i=1}^k e^{-dF_{\langle \hat{\pi}_{[k+1,n]}, \hat{\nu} \rangle}(\delta_{\bar{\sigma}_i} - \langle \hat{\pi}_{[k+1,n]}, \hat{\nu} \rangle)}$$

$$\qquad (107)$$

Let us decompose the $\sum_{\hat{\nu} \in \hat{M}_{[k+1,n]}}$ over $C(M^*, \varepsilon)$ and its complement, and compare the terms with their values at the midpoints:

$$\xi_N = \sum_{\hat{\nu}^* \in M^*} \sum_{\substack{\hat{\nu} \in \\ B(\varepsilon, \hat{\nu}^*) \cap \hat{M}_{[k+1,n]}}} \rho_{F,n,k}[\eta](\hat{\nu}) \prod_{i=1}^k \alpha[\eta_i](\sigma_i) \prod_{i=1}^k e^{-dF_{\pi\hat{\nu}^*}(\delta_{\sigma_i} - \pi\hat{\nu}^*)}$$

$$+ \sum_{\hat{\nu}^* \in M^*} \sum_{\substack{\hat{\nu} \in \\ B(\varepsilon, \hat{\nu}^*) \cap \hat{M}_{[k+1,n]}}} \rho_{F,n,k}[\eta](\hat{\nu}) \prod_{i=1}^k \alpha[\eta_i](\sigma_i)$$

$$\times \left[\prod_{i=1}^{k} e^{-dF_{\langle \hat{\pi}_{[k+1,n]}, \hat{\nu} \rangle}(\delta_{\sigma_i} - \langle \hat{\pi}_{[k+1,n]}, \hat{\nu} \rangle)} - \prod_{i=1}^{k} e^{-dF_{\pi \hat{\nu}^*}(\delta_{\sigma_i} - \pi \hat{\nu}^*)}\right]$$

$$+ \sum_{\substack{\hat{\nu} \in \\ C(M^*, \varepsilon)^c \cap \hat{M}_{[k+1,n]}}} \rho_{F,n,k}[\eta](\hat{\nu}) \prod_{i=1}^{k} \alpha[\eta_i](\sigma_i) \prod_{i=1}^{k} e^{-dF_{\langle \hat{\pi}_{[k+1,n]}, \hat{\nu} \rangle}(\delta_{\sigma_i} - \langle \hat{\pi}_{[k+1,n]}, \hat{\nu} \rangle)} \quad (108)$$

The sum in the last line is bounded by a function $r_{cp}(\varepsilon, n)$, where $\lim_{n \uparrow \infty} r_{cp}(\varepsilon, n) = 0$, when η is in the union of the *good-sets*. This holds by the concentration property of the empirical distribution given in (51) applied to the measure for sites $\geq k$, using the boundedness of the first derivative of F.

The second line is bounded in modulus by a function $\gamma(\varepsilon)$ where $\lim_{\varepsilon \downarrow 0} \gamma(\varepsilon) = 0$ by the twice continuous differentiability of F.

This implies the bounds

$$\left| \xi_N - \sum_{\hat{\nu}^* \in M^*} \rho_{F,n,k}[\eta](\tilde{B}(\varepsilon, \hat{\nu}^*)) \prod_{i=1}^{k} \alpha[\eta_i](\sigma_i) e^{-dF_{\pi \hat{\nu}^*}(\delta_{\sigma_i} - \pi \hat{\nu}^*)} \right| \leq \gamma(\varepsilon) + r_{cp}(\varepsilon, n) \quad (109)$$

assuming that η is in the union of the *good-sets*.

Summing over the finitely many values of $\sigma_1, \ldots, \sigma_k$ we obtain the same type of bounds (with possibly worse functions $\gamma(\varepsilon), r_{cp}(\varepsilon, n)$) for ξ_D.

Recall the definition of the kernels (10) and choose the disorder variable $\eta \in \mathcal{H}_{i,n,k}^{\tau}$ in the part of the union of the *good-sets* which ensures the dominance of the i-th minimizer. This gives that

$$\left| \mu_{F,n}[\eta](\sigma_1, \ldots, \sigma_k) - \prod_{j=1}^{k} \gamma[\eta_j](\sigma_j | \pi \hat{\nu}_i^*) \right| \leq \zeta(\varepsilon) + \chi(\varepsilon, n) \quad (110)$$

where $\lim_{n \uparrow \infty} \chi(\varepsilon, n) = 0$ and $\lim_{\varepsilon \downarrow 0} \zeta(\varepsilon) = 0$. This proves the lemma and also the statement (91).

Lemma 3.2 *Let Ξ be a continuous real-valued function on $\mathcal{P}(E^{\infty}) \times (E')^{\infty}$. Then under the* non-degeneracy assumptions (1) *and* (2) *the following holds*:

$$\lim_{n \uparrow \infty} \int_{\mathcal{H}_{i,n}^{\delta_n}} \mathbb{P}_{\pi}(d\eta) \Xi(\mu_{F,n}[\eta], \eta) = w_i \int \mathbb{P}_{\pi}(d\eta)(d\eta) \Xi(\mu_i[\eta], \eta) \quad (111)$$

$w_i = \mathbb{P}_{\pi}(G \in R_i)$.

The proof of this lemma, thanks to the continuity of Ξ which allows finite dimensional approximation and to Lemma 3.1, follows the trail drawn by the proof of Lemma 2.10.

Using all the tools we have provided, we find

$$\lim_{n \uparrow \infty} \int \mathbb{P}_{\pi}(\eta) \Xi(\mu_{F,n}[\eta], \eta) = \sum_{j=1}^{k} \int \mathbb{P}_{\pi}(\eta) \Xi(\mu, \eta) w_j \delta_{\mu_j[\eta]}(d\mu) \quad (112)$$

where we can identify

$$J(d\mu, d\eta) = \sum_{j=1}^{k} \mathbb{P}_\pi(d\eta) w_j \delta_{\mu_j[\eta]}(d\mu)$$

$$\implies \quad J(d\mu|\eta) = \sum_{j=1}^{k} w_j \delta_{\mu_j[\eta]}(d\mu) \tag{113}$$

This finishes the proof of the Main Theorem 1.7 and concludes the paper. □

References

1. Aizenman, M., Wehr, J.: Rounding effects of quenched randomness on first-order phase transitions. Commun. Math. Phys. **130**, 489–528 (1990)
2. Amaro de Matos, J.M.G., Patrick, A.E., Zagrebnov, V.A.: Random infinite-volume Gibbs states for the Curie-Weiss random field Ising model. J. Stat. Phys. **66**, 139–164 (1992)
3. Arguin, L.-P., Damron, M., Newman, C.M., Stein, D.L.: Uniqueness of ground states for short-range spin glasses in the half-plane. arXiv:0911.4201 (2009)
4. Bianchi, A., Bovier, A., Ioffe, D.: Sharp asymptotics for metastability in the random field Curie-Weiss model. Electron. J. Probab. **14**, 1541–1603 (2009)
5. Bricmont, J., Kupiainen, A.: Phase transition in the 3d random field Ising model. Commun. Math. Phys. **116**(4), 539–572 (1988)
6. Bovier, A.: Statistical Mechanics of Disordered Systems. A Mathematical Perspective. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge (2006)
7. Bovier, A., van Enter, A., Niederhauser, B.: Stochastic symmetry-breaking in a Gaussian Hopfield model. J. Stat. Phys. **95**, 181–213 (1999)
8. Bovier, A., Gayrard, V.: Metastates in the Hopfield model in the replica symmetric regime. Math. Phys. Anal. Geom. **1**, 107–144 (1998)
9. Dembo, A., Zeitouni, O.: Large Deviations Techniques and Applications. Stochastic Modelling and Applied Probability, vol. 38. Springer-Verlag, Berlin (2010)
10. Ellis, R., Wang, K.: Limit theorems for the empirical vector of the Curie-Weiss-Potts model. Stoch. Process. Appl. **35**, 59–79 (1990)
11. van Enter, A., Netočný, K., Schaap, H.G.: Incoherent boundary conditions and metastates. In: Dynamics & Stochastics. IMS Lecture Notes Monogr. Ser., vol. 48, pp. 144–153. Inst. Math. Statist., Beachwood (2006)
12. Dai Pra, P., den Hollander, F.: McKean-Vlasov limit for interacting random processes in random media. J. Stat. Phys. **84**, 735–772 (1996)
13. Külske, C.: Metastates in disordered mean-field models: Random field and Hopfield models. J. Stat. Phys. **88**, 1257–1293 (1997)
14. Külske, C.: Metastates in disordered mean-field models. II. The superstates. J. Stat. Phys. **91**, 155–176 (1998)
15. Külske, C.: Limiting behavior of random Gibbs measures: Metastates in some disordered mean field models. In: Mathematical Aspects of Spin Glasses and Neural Networks. Progr. Probab., vol. 41, pp. 151–160. Birkhäuser, Boston (1998)
16. Külske, C., LeNy, A.: Spin-flip dynamics of the Curie-Weiss model: Loss of Gibbsianness with possibly broken symmetry. Commun. Math. Phys. **271**, 431–454 (2007)
17. Matoušek, J.: Lectures on Discrete Geometry. Graduate Texts in Mathematics, vol. 212. Springer-Verlag, New York (2002)

18. Newman, C.M., Stein, D.L.: Are there incongruent ground states in 2D Edwards-Anderson spin glasses? Commun. Math. Phys. **224**, 205–218 (2001)
19. Newman, C.M.: Topics in Disordered Systems. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel (1997)
20. Newman, C.M., Stein, D.L.: Metastate approach to thermodynamic chaos. Phys. Rev. E **55**, 5194–5211 (1997)
21. Newman, C.M., Stein, D.L.: The state(s) of replica symmetry breaking: mean field theories vs. short-ranged spin glasses J. Stat. Phys. **106**(1–2), 213–244 (2002)