Self-Averaging Identities for Random Spin Systems

Luca De Sanctis *, Silvio Franz †

November 7, 2008

Abstract

We provide a systematic treatment of self-averaging identities, whose validity is proven in integral average, for dilute spin glasses. The method is quite general, and as a special case recovers the Ghirlanda-Guerra identities, which are therefore proven, together with their extension, to be valid in dilute spin glasses. We focus on dilute spin glasses, but the results hold in all models enjoying stability with respect the perturbations we introduce; although such a stability is believed to hold for several models, we do not classify them here.

Key words and phrases: diluted spin glasses, Ghirlanda-Guerra, self-averaging.
Mathematics Subject Classification: 82B44; 82B20

1 Introduction

Despite many years of intense work, and the much awaited proof of the validity of the Parisi ansatz for the free-energy of the Sherrington-Kirkpatrick (SK) and related models, the mathematical comprehension of the thermodynamics of mean field spin glasses remains largely incomplete. We know from theoretical physics that in fully connected models all the properites of the low temperature spin glass phase can be encoded in the probability distribution of the overlap between two different copies of the system. The analysis of Parisi et al. predicts an ultrametric organization of the phases (see [12] and references therein). So far the rigorous proof (or disproof) of ultrametricity, and, more in general, the analysis of the structure of Gibbs measures at low temperature, turned out to be a very difficult task. A step in this direction was performed by Ghirlanda and Guerra in [8]. They found a simple and elegant way, based on the self-averaging of the internal energy, to prove a remarkable property of the overlaps. Given s replicas, the Gibbs measure must be such that when one adds a further replica

*ICTP, Strada Costiera 11, 34014 Trieste, Italy, <lde_sanc@ictp.it>
†LPTMS, bâtiment 100, Université Paris-Sud 11, Centre scientifique d'Orsay, 15 rue Georges Clémenceau, 91405 Orsay cedex <silvio.franz@lptms.u-psud.fr>

this is either identical to one these, or statistically independent of them; each case occurring with the same probability. More generally, various constraints on the distribution of the different overlaps have been found in the same spirit [2, 14]. Such features have found several applications [16, 4] in the rigorous analysis of spin glass models. For example, the property of non-negativity of the overlap, which in some models plays a role in turning the cavity free-energy into a rigorous lower bound, turns out to be a consequence of the Ghirlanda-Guerra self-averaging identities [16]. In the same way these identities have a role in the rigorous analysis of spin glasses close to the critical temperature [1].

In more general spin-glass systems, like finite dimensional systems or spin systems on random graphs, the statistics of the overlap is not enough to fully characterize the low temperature spin glass phase. For instance, in diluted models the statistics of the local cavity fields, or equivalently of all the multi-overlaps, is necessary to describe the low temperature thermodynamic properties. In this paper, we analyse two families of identities for the local fields and multi-overlap distributions that are a consequence of self-averaging relations. We will see that first of the two families is a consequence of the self-averaging with respect to the Gibbs measure or, equivalently, of stochastic stability, as the two phenomena turn out to be equivalent. The other family of identities is instead a consequence of self-averaging with respect to the global measure (quenched after Gibbs). The second family contains the first. The form of the identities we obtain is due only to the form of the perturbations we introduce, and does not depend on the specific form of the Hamiltonian. However, the self-averaging at the basis of the results does not necessarily holds for all Hamiltonians. So we will stick with the case of dilute spin glass (Viana-Bray model), for which the self-averaging is assured, but our method shows that the same identities we find hold whenever the self-averaging is true. We do not provide a classification of all the models exhibiting self-averaging when perturbed. A second reason to use the example of spin models on sparse random graphs (dilute spin glass models), is that we expect that our results could provide hints for progresses in the mathematical analysis of the low temperature phases. Diluted mean field spin glasses have, in recent time, attracted a lot of attention in statistical physics, due to the intrinsic interest of spin glasses where each spin interacts with a finite number of variables, but more importantly because fondamental problems in computer science, such as the random K-SAT and graph coloring, the random X-OR-SAT, tree reconstruction [11] and others, admit a formulation in terms of spin glass systems on random graphs. The cavity approach to these problems has led in many cases to results believed to be exact, albeit for the moment several rigorous proofs are still lacking.

One of the two families of the identities that we will find appeared already in [7] to discuss free energy bounds in diluted models with non-Poissonian connectivity. In [7] such a family of identities was shown to be a consequence of self-averaging of certain random polynomial function of some spins variables, and the self-averaging was deduced from the convexity of the perturbed free energy. Here we re-derive this family of identities with a different strategy, employing stochastic stability of the free energy with respect to suitable perturbations of the Hamiltonian, and we show that the stability implies self-averaging. The use of stochastic stability makes our results valid even when a measure different from the Gibbs one is considered, provided certain conditions hold (we will briefly hint at this when introducing Random Multi-Overlap Structures). Moreover, our method shows that the constraints we find are valid whenever stochastic stability or self-averaging hold, whichever turns out to be easier to study, and in the physics literature many models have been investigated from at least one of the two points of view. We also exhibit a second family of new identities, which contains the first family and follows from the self-averaging with respect to the quenched expectation.

For both families, we will start with the two-spin model, perturbed with a two-spin random field. This provides identities involving squared overlaps. The same identical method, reproduced for generic p-spin interactions, yields the same identities involving the p-th power of the overlaps. Therefore considering Hamiltonians and perturbations with all p-spin interactions we will conclude that the identities hold for all regular functions of the overlaps.

Let us stress here that the basic tool we employ is the introduction of suitable perturbations, such that the pressure (minus the free energy divided by the temperature) is convex in all the perturbing parameters. This guarantees the existence of the derivatives with respect to the perturbing parameters only almost everywhere. Therefore the relations we find are valid only when one integrates (with Lebesgue measure) back against the perturbing parameters over any given intervals. We summarize this by recalling that our relations hold in "integral average".

2 The notations

We will deal with the stereotypical dilute spin glass model, the Viana-Bray (VB), for which we are about to describe the notations we need to derive our results in the next two sections. Let α, β be non-negative real numbers (degree of connectivity and inverse temperature respectively); P_ζ be a Poisson random variable of mean ζ; $\{i_\nu\}, \{j_\nu\}$, etc. be independent identically distributed random variables, uniformly distributed over the points $\{1, \ldots, N\}$; $\{J_\nu\}, J$, etc. be independent identically distributed random variables, with symmetric distribution; $\mathcal{J}$ be the set of all the quenched random variables above. The map $\sigma : i \to \sigma_i,\ i \in \{1, \ldots, N\}$ is a spin configuration from the configuration space $\Sigma = \{-1, 1\}^N$; $\pi_\zeta(\cdot)$ is the Poisson measure of mean ζ. The Hamiltonian of the Viana-Bray model is defined as

$$H_N(\sigma, \alpha; \mathcal{J}) = -\sum_{\nu=1}^{P_{\alpha N}} J_\nu \sigma_{i_\nu} \sigma_{j_\nu} \ . \tag{1}$$

We will limit to the case $J = \pm 1$, without loss of generality [10]. We follow the usual basic definitions and notations of thermodynamics for the partition function Z_N, the pressure p_N, the free energy per site f_N and its thermodynamic limit f, so to have in general

$$Z_N(\beta,\alpha) \equiv Z(H_N;\beta,\alpha;\mathcal{J}) = \sum_{\{\sigma\}} \exp(-\beta H_N(\sigma,\alpha;\mathcal{J})) \ , \tag{2}$$

$$p_N(\beta,\alpha) = -\beta f_N(\beta,\alpha) = \frac{1}{N}\mathbb{E}\ln Z_N(\beta,\alpha) \ , \ f(\beta,\alpha) = \lim_{N\to\infty} f_N(\beta,\alpha) \ . \tag{3}$$

The Boltzmann-Gibbs average of an observable $\mathcal{O} : \Sigma \to \mathbb{R}$ is

$$\Omega(\mathcal{O}) = Z_N(\beta,\alpha;\mathcal{J})^{-1} \sum_{\{\sigma\}} \mathcal{O}(\sigma) \exp(-\beta H_N(\sigma,\alpha;\mathcal{J})) \ , \tag{4}$$

and $\langle \mathcal{O} \rangle = \mathbb{E}\Omega(\mathcal{O})$ is the global average, where $\mathbb{E}$ denotes the average with respect to the quenched variables. When dealing with more than one configuration, we need the product measure of the needed copies of Ω, which will be denoted again by Ω.

The multi-overlaps $q_{1\cdots m} : \Sigma^m \to [-1,1]$, where we use the notation $\Sigma^n = \Sigma^{(1)} \times \cdots \times \Sigma^{(n)}$, among the "replicas" $\Sigma^{(r_1)} \ni \sigma^{(r_1(}, \ldots, \Sigma^{(r_n)} \ni \sigma^{(r_n)}$ is defined by

$$q_{r_1\cdots r_n} = \frac{1}{N}\sum_{i=1}^{N} \sigma_i^{(r_1)} \cdots \sigma_i^{(r_n)} \ , \tag{5}$$

but sometimes we will just write q_n; q_1 can be identified with the magnetization m

$$m = \frac{1}{N}\sum_{i=1}^{N} \sigma_i \ .$$

Notice that

$$\mathbb{E}[\Omega(\sigma_{i_1})]^{2n} = \langle q_{1\cdots 2n} \rangle \ , \ \mathbb{E}\Omega(\sigma_{i_1}) = \mathbb{E}\Omega(m) = \langle m \rangle \ , \tag{6}$$

and that

$$\mathbb{E}[\Omega(\sigma_{i_1}\cdots\sigma_{i_p})]^{2n} = \langle q^p_{1\cdots 2n} \rangle \ , \ \mathbb{E}\Omega(\sigma_{i_1}\cdots\sigma_{i_p}) = \mathbb{E}\Omega(m^p) = \langle m^p \rangle \ , \tag{7}$$

for all integer n and p.

3 Stochastic Stability and self-averaging of the Gibbs measure

In the study of finite connectivity models it emerged that in a suitable propability space it is possible to formulate an exact variational principle for the computation of the free energy. This was obtained with the introduction of Random Multi-Overlap Structures (RaMOSt). We refer to [6] for details. The ROSt approach is based on the use of generic random weights to average the "cavity" part and the relative "internal correction" in the free energy (these are the numerator and the denominator of the trial free energy G_N introduced in (10). See [6] for details). Here we are not interested in a detailed discussion of the RaMOSt approach, but we study the effect of a perturbation to the measure of our model, which does not need to be the Gibbs measure. That is why introduce this more general weighting scheme, although the reader may keep in mind the Gibbs measure as a guiding example.

3.1 Random Multi-Overlap Structures

The proper framework for the calculation of the free energy per spin is that of the Random Multi-Overlap Structures (RaMOSt, see [6] for more details).

Definition 1 *Given a probability space* $\{\Omega, \mu(d\omega)\}$*, a* **Random Multi-Overlap Structure** $\mathcal{R}$ *is a triple* $(\tilde{\Sigma}, \{\tilde{q}_{2n}\}, \xi)$ *where*

- $\tilde{\Sigma}$ *is a discrete space;*
- $\xi : \tilde{\Sigma} \to \mathbb{R}_+$ *is a system of random weights, such that* $\sum_{\gamma \in \tilde{\Sigma}} \xi_\gamma \leq \infty$ μ*-almost surely;*
- $\tilde{q}_{2n} : \tilde{\Sigma}^{2n} \to \mathbb{R}, n \in \mathbb{N}$ *is a positive semi-definite* Multi-Overlap Kernel *(equal to 1 on the diagonal of* $\tilde{\Sigma}^{2n}$*, so that by Schwartz inequality* $|\tilde{q}| \leq 1$*).*

A RaMOSt needs to be equipped with N independent copies of a random field $\{\tilde{h}^i_\gamma(\alpha; \tilde{\mathcal{J}})\}_{i=1}^N$ and with another random field $\hat{H}_\gamma(\alpha N; \mathcal{J}')$ such that

$$\frac{d}{d\alpha}\mathbb{E}\ln \sum_{\gamma \in \tilde{\Sigma}} \xi_\gamma \exp(-\beta \tilde{h}^i_\gamma(\alpha; \tilde{\mathcal{J}})) = 2 \sum_{n>0} \frac{1}{2n} \tanh^{2n}(\beta)(1 - \langle \tilde{q}_{2n} \rangle), \quad (8)$$

$$\frac{d}{d\alpha}\mathbb{E}\ln \sum_{\gamma \in \tilde{\Sigma}} \xi_\gamma \exp(-\beta \hat{H}_\gamma(\alpha N; \mathcal{J}')) = \sum_{n>0} \frac{1}{2n} \tanh^{2n}(\beta)(1 - \langle \tilde{q}^2_{2n} \rangle). \quad (9)$$

The quenched variables in $\tilde{h}$ and $\hat{H}$ are independent one another and independent of those in the weights ξ. These two fields just introduced are employed in the definition of the trial pressure

$$G_N(\mathcal{R}; \beta, \alpha) = \frac{1}{N}\mathbb{E}\ln \frac{\sum_{\gamma,\sigma} \xi_\gamma \exp(-\beta \sum_{i=1}^N \tilde{h}^i_\gamma(\alpha; \tilde{\mathcal{J}})\sigma_i)}{\sum_\gamma \xi_\gamma \exp(-\beta \hat{H}_\gamma(\alpha N; \mathcal{J}'))} \ . \quad (10)$$

Notice that the expectation $\langle \cdot \rangle$ here is not necessarily the quenched Gibbs one: it is the generic one of the RaMOSt.

The reason why this is the proper framework for the calculation of the free energy is explained by the next [6]

Theorem 1 (Extended Variational Principle) *Taking the infimum for each* N *separately of the trial function* $G_N(\mathcal{R}; \beta, \alpha)$ *over the space of all RaMOSt's, the resulting sequence tends to the limiting pressure* $-\beta f(\beta, \alpha)$ *of the VB model as* N *tends to infinity:*

$$-\beta f(\beta, \alpha) = \lim_{N \to \infty} \inf_{\mathcal{R}} G_N(\mathcal{R}; \beta, \alpha) \ . \quad (11)$$

A RaMOSt $\mathcal{R}$ is said to be optimal if $G(\mathcal{R};\beta,\alpha) = -\beta f(\beta,\alpha) \ \ \forall\, \beta,\alpha$. Recall that we denote by Ω the measure associated to the RaMOSt weights ξ.

The Boltzmann RaMOSt [6] is optimal, and constructed by thinking of a reservoir of M spins τ

$$\Sigma = \{-1,1\}^M \ni \tau \ , \ \xi_\tau = \exp(-\beta H_M(\tau,\alpha,\mathcal{J})) \ ,$$

$$\tilde{q}_{1\cdots 2n} = \frac{1}{M}\sum_{k=1}^{M} \tau_k^{(1)} \cdots \tau_k^{(2n)}$$

with

$$\tilde{h}_\tau^i(\alpha,\tilde{\mathcal{J}}) = \sum_{\nu=1}^{\tilde{P}_{2\alpha}} \tilde{J}_\nu^i \tau_{k_\nu^i} \ , \ \hat{H}_\tau(\alpha N,\mathcal{J}') = -\sum_{\nu=1}^{\hat{P}_{\alpha N}} \hat{J}_\nu \tau_{k_\nu}\tau_{l_\nu} \ ,$$

with $\tilde{J},\hat{J}$ all independent copies of J, and with the random site variables k_ν, l_ν all mutually independent and uniformly distributed over $\{1,\dots M\}$.

Let $c_i = 2\cosh(\beta\tilde{h}_\tau^i(\alpha,\tilde{\mathcal{J}}))$. It is possible to show [6] that optimal RaMOSt's enjoy the same factorization property enjoyed by the Boltzmann RaMOSt and described in the next [6]

Theorem 2 (Factorization of optimal RaMOSt's) *The following Cesàro limit is linear in N and $\bar{\alpha}$*

$$\mathbf{C}\lim_M \mathbb{E}\ln\Omega_M\{c_1\cdots c_N \exp[-\beta\hat{H}_\tau(\bar{\alpha},\mathcal{J}')]\} = N(-\beta f + \alpha A) + \bar{\alpha}A \ , \qquad (12)$$

where

$$A = \sum_{n=1}^{\infty} \frac{1}{2n}\mathbb{E}\tanh^{2n}(\beta J)(1 - \langle q_{2n}^2\rangle) \ , \qquad (13)$$

and the equality holds in integral average, i.e. once both sides are integrated against α over any given interval.

This factorization property is called *invariance with respect to the cavity step*, or *Quasi-Stationarity*, and it is found in the hierarchical Parisi ansatz as well. When $\bar{\alpha}$ is zero, the theorem above states the factorization of the cavity fields, and it is possible to show that from this property one can deduce the family of identities we will discuss in the next subsection [3]. When one removes instead the cavity terms $c_1,\dots,c_N$ from the previous theorem, the statement becomes what is usually referred to as Stochastic Stability. We will show that the latter too implies the same family of identities. We will have in mind the case of a small perturbation of our spin system, but what we find holds for more general RaMOSt's, provided the previous theorem holds, that is for Quasi-Stationary RaMOSt's.

3.2 The first family of identities

We will now prove a lemma that expresses the stability of the Gibbs measure of our model against a macroscopic but small stochastic perturbation. In different terms, the lemma expresses the linear response of the free energy to the connectivity shift the perturbation consists of. The lemma we are about to prove will be used to show that from stochastic stability one can deduce a certain self-averaging which in turn imposes a family of constraints on the distribution of the overlaps.

Lemma 1 *Let Ω , $\langle\cdot\rangle$ be the usual infinite volume Gibbs and quenched Gibbs expectations at inverse temperature β, associated with the Hamiltonian $H_N(\sigma,\alpha;\mathcal{J})$, $N\to\infty$. Then the following equality (understood to be in the thermodynamic limit) holds*

$$\mathbb{E}\ln\Omega\exp\left(\beta'\sum_{\nu=1}^{P_{\alpha'}}J'_\nu\sigma_{i'_\nu}\sigma_{j'_\nu}\right)=\alpha'\sum_{n=1}^{\infty}\frac{1}{2n}\tanh^{2n}(\beta')(1-\langle q_{2n}^2\rangle)\ . \tag{14}$$

in integral average with respect to the degree of connectivity. The random variables $P_{\alpha'},\{J'_\nu\}$, $\{i'_\nu\},\{j'_\nu\}$ are independent copies of the analogous random variables in the Hamiltonian contained in Ω.

Notice that, in distribution

$$\beta\sum_{\nu=1}^{P_{\alpha N}}J_\nu\sigma_{i_\nu}\sigma_{j_\nu}+\beta'\sum_{\nu=1}^{P_{\alpha'}}J'_\nu\sigma_{i'_\nu}\sigma_{j'_\nu}\sim\beta\sum_{\nu=1}^{P_{\alpha N+\alpha'}}J''_\nu\sigma_{i_\nu}\sigma_{j_\nu} \tag{15}$$

where $\{J''_\nu\}$ are independent copies of J with probability $\alpha N/(\alpha N+\alpha')$ and independent copies of $J\beta'/\beta$ with probability $\alpha'/(\alpha N+\alpha')$. In the right hand side above, the quenched random variables will be collectively denoted by $\mathcal{J}''$. Notice also that the sum of Poisson random variables is a Poisson random variable with mean equal to the sum of the means, and hence we can write

$$A_t\equiv\mathbb{E}\ln\Omega\exp\left(\beta'\sum_{\nu=1}^{P_{\alpha' t}}J'_\nu\sigma_{i'_\nu}\sigma_{j'_\nu}\right)=\mathbb{E}\ln\frac{Z_N(\alpha_t;\mathcal{J}'')}{Z_N(\alpha;\mathcal{J})}\ , \tag{16}$$

where we defined, for $t\in[0,1]$,

$$\alpha_t=\alpha+\alpha'\frac{t}{N} \tag{17}$$

so that $\alpha_t\to\alpha\ \forall\ t$ as $N\to\infty$.

Proof. Let us compute the t-derivative of A_t, as defined in (16)

$$\frac{d}{dt}A_t=\mathbb{E}\sum_{m=1}^{\infty}\frac{d}{dt}\pi_{\alpha' t}(m)\ln\sum_\sigma\exp\left(\beta'\sum_{\nu=1}^{m}J'_\nu\sigma_{i'_\nu}\sigma_{j'_\nu}\right)\ . \tag{18}$$

Using the following elementary property of the Poisson measure

$$\frac{d}{dt}\pi_{t\zeta}(m) = \zeta(\pi_{t\zeta}(m-1) - \pi_{t\zeta}(m)) \tag{19}$$

we get

$$\begin{aligned}
\frac{d}{dt}A_t &= \alpha'\mathbb{E}\sum_{m=0}^{\infty}[\pi_{\alpha' t}(m-1) - \pi_{\alpha' t}(m)]\ln\sum_{\sigma}\exp(\beta'\sum_{\nu=1}^{m}J'_{\nu}\sigma_{i'_{\nu}}\sigma_{j'_{\nu}}) \\
&= \alpha'\mathbb{E}\ln\sum_{\sigma}\exp(\beta' J'\sigma_{i'_m}\sigma_{j'_m})\exp(\beta'\sum_{\nu=1}^{P_{\alpha' t}}J'_{\nu}\sigma_{i'_{\nu}}\sigma_{j'_{\nu}}) \\
&\qquad\qquad -\alpha'\mathbb{E}\ln\sum_{\sigma}\exp(\beta'\sum_{\nu=1}^{P_{\alpha' t}}J'_{\nu}\sigma_{i'_{\nu}}\sigma_{j'_{\nu}}) \\
&= \alpha'\mathbb{E}\ln\Omega_t\exp(\beta' J'\sigma_{i'_m}\sigma_{j'_m})\ ,
\end{aligned}$$

where the average Ω_t is associated with the Hamiltonian plus the t-dependent weights in the exponential in (16). Now use the following identity

$$\exp(\beta' J'\sigma_i\sigma_j) = \cosh(\beta' J') + \sigma_i\sigma_j\sinh(\beta' J')$$

to get

$$\frac{d}{dt}A_t = \alpha'\mathbb{E}\ln\Omega_t[\cosh(\beta' J')(1+\tanh(\beta' J')\sigma_{i'_m}\sigma_{j'_m})]\ . \tag{20}$$

We now expand the logarithm in power series and see that, when $N\to\infty$, as $\alpha_t\to\alpha$ the result does not depend on t, wherever the expectation Ω_t is continuous as a function of the parameter t. From the comments that preceded the current proof, formalized in (15)-(16)-(17), this is the same as assuming that Ω is regular as a function of α, because $J''\to J$ in the sense that in the large N limit J'' can only take the usual values ± 1 since the probability of being $\pm\beta'/\beta$ becomes zero. Therefore integrating against t from 0 to 1 is the same as multiplying by 1. Due to the symmetric distribution of J, the expansion of the logarithm yields the right hand side of (14), where the odd powers are missing. $\square$

Notice that the left hand side of (14) can be written, according to our notations, $\mathbb{E}\ln\Omega\exp(-\beta'\hat{H}_{\sigma}(\alpha',\mathcal{J}'))$. We want now to consider the statement of Lemma 1 in the case of two independent perturbations, assuming we are always in the thermodynamic limit. The consequent generalization obtained by using twice the fundamental theorem of calculus simply gives (in integral average)

$$\mathbb{E}\ln\Omega[\exp(-\beta'_1\hat{H}_{\sigma}(\alpha'_1;\mathcal{J}'_1) - \beta'_2\hat{H}_{\sigma}(\alpha'_2;\mathcal{J}'_2))] = (\alpha'_1+\alpha'_2)A\ , \tag{21}$$

where A again does not depend, in the thermodynamic limit, on α'_1,α'_2, and has the same form as the right hand side of (14) although the explicit for of A is not important here. In the equation above, assumed to be taken in the thermodynamic limit, Ω is the Gibbs measure associated with the unperturbed

Hamiltonian of the original model, and the same holds for the averages appearing in A, just like in the previous lemma. Clearly at this point we have

$$\frac{\partial^2}{\partial\alpha'_1\partial\alpha'_2}\mathbb{E}\ln\Omega[\exp(-\beta'_1\hat{H}_\sigma(\alpha'_1;\mathcal{J}'_1)-\beta'_2\hat{H}_\sigma(\alpha'_2;\mathcal{J}'_2))]=0\ ,$$

whenever the derivatives exist, i.e. with the possible exception of the zero measure set, by convexity. A simple computation yields

$$\begin{aligned}\frac{\partial^2}{\partial\alpha'_1\partial\alpha'_2}\mathbb{E}\ln\Omega[\exp(-\beta'_1\hat{H}_\sigma(\alpha'_1;\mathcal{J}'_1)-\beta'_2\hat{H}_\sigma(\alpha'_2;\mathcal{J}'_2))]=0&\\ =\mathbb{E}\ln\Omega'[\exp(\beta'_1J'_1\sigma_{i_1}\sigma_{j_1}+\beta'_2J'_2\sigma_{i_2}\sigma_{j_2}]&\\ -\mathbb{E}\ln\Omega'[\exp(\beta'_1J'_1\sigma_{i_1}\sigma_{j_1})]\Omega'[\exp(\beta'_2J'_2\sigma_{i_2}\sigma_{j_2})]&\end{aligned}$$

Every time a derivative with respect to a pertubing parameter is taken, the relative perturbation is added to the weights of the measure Ω, which hence replaced by a perturbed measure denoted by Ω'. If the pertubation is small (like in our case, as explained in the previous lemma) it disappears from the measure in the thermodynamic limit. Recall that this holds with the usual limitation, i.e. only in integral average, because each derivative exists only almost everywhere, and meaningful equalities are hence proven only under integration over any given interval. Hence both in the equation above and in the next calculation β'_1,β'_2 are not in the measure Ω', and we get

$$\begin{aligned}\frac{\partial^2}{\partial(\beta'_1J_1)\partial(\beta'_2J_2)}\mathbb{E}\ln\Omega'[\exp(\beta'_1J'_1\sigma_{i_1}\sigma_{j_1}+\beta'_2J'_2\sigma_{i_2}\sigma_{j_2}]&\\ =\mathbb{E}\Omega''(\sigma_{i_1}\sigma_{j_1})-\mathbb{E}\Omega''(\sigma_{i_1})\Omega''(\sigma_{j_1})=0\ ,&\quad(22)\end{aligned}$$

again in integral average with respect to all the parameters $\alpha'_1,\alpha'_2,\beta'_1,\beta'_2$. These new derivatives with respect to β'_1,β'_2 again introduce further perturbations in the weights, which is why we used the notation Ω'', but as usual in the thermodynamic limit they vanish (in integral average). The first line of (22) gives us the generator of a family of relations that we will obtain by means of an expansion in powers of β'_1,β'_2. The second line of the equation formulates the self-averaging (with respect to the Gibbs measure) implied by stochastic stability.

So we proceed starting from the next lemma and the next theorem, summarizing what we just discussed. We want here to remind ourselves of the presence of the perturbations, vanishing only in the thermodynamic limit and only with probability one in the space of all parameters, by denoting any perturbed Gibbs expectation with Ω', independently of the perturbations.

Lemma 2 *Let Ω' be the Gibbs measure including two independent perturbations of the form*

$$\hat{H}_\sigma(\alpha';\mathcal{J}')=\sum_{\nu=1}^{P_{\alpha'}}J'_\nu\sigma_{i_\nu}\sigma_{j_\nu}$$

with parameters $\alpha'_1, \alpha'_2, \beta'_1, \beta'_2$ like in (21). Then, recalling that m is the magnetization, the following self-averaging (with respect to the Gibbs measure) identity

$$\mathbb{E}\{\Omega'(m^2) - [\Omega'(m)]^2\} = 0 \tag{23}$$

holds in the thermodynamic limit in integral average with respect to the perturbing parameters $\alpha'_1, \alpha'_2, \beta'_1, \beta'_2$.

We will see again that in the first line of equation (22) the expression remains zero even without the derivative. In fact the generator of the identities we want to prove is expressed in the following

Theorem 3 *In the thermodynamic limit the following identity holds*

$$\begin{aligned} &\mathbb{E}\ln\Omega'(\exp(\beta'_1 J'_1 \sigma_{i_1}\sigma_{j_1} + \beta'_2 J'_2 \sigma_{i_2}\sigma_{j_2})) = \\ &\mathbb{E}\ln\Omega'(\exp(\beta'_1 J'_1 \sigma_{i_1}\sigma_{j_1})) + \mathbb{E}\ln\Omega'(\exp(\beta'_2 J'_2 \sigma_{i_2}\sigma_{j_2}))\ . \end{aligned} \tag{24}$$

in integral average with respect to $\alpha'_1, \alpha'_2, \beta'_1, \beta'_2$.

The relations we will derive are a simple consequence of this theorem, and fomalized in the next

Corollary 1 *In the thermodynamic limit, in integral average with respect to the perturbing parameters $\alpha'_1, \alpha'_2, \beta'_1, \beta'_2$, we have*

$$\sum_{a=0}^{\min\{r,s\}} (-)^{a+1} \frac{(2r+2s-a-1)!}{a!(2r-a)!(2s-a)!} \langle q^2_{2r} q^2_{2s}\rangle'_a = 0 \quad \forall\ r, s \in \mathbb{N}\ ,$$

where the subscript a in the global average $\langle\cdot\rangle'_a = \mathbb{E}\Omega'_a$ means that precisely a replicas are in common among those in q_{2r} and those in q_{2s}.

Proof. The following shorthand will be employed

$$t_1 = \tanh(\beta'_1 J'_1)\ ,\ t_2 = \tanh(\beta'_2 J'_2)\ ,$$

$$\Omega_1 = \Omega'(\sigma_{i_1}\sigma_{j_1})\ ,\ \Omega_2 = \Omega'(\sigma_{i_2}\sigma_{j_2})\ ,\ \Omega_{12} = \Omega'(\sigma_{i_1}\sigma_{j_1}\sigma_{i_2}\sigma_{j_2})$$

and

$$W = \Omega'(\exp(\beta'_1 J'_1 \sigma_{i_1}\sigma_{j_1} + \beta'_2 J'_2 \sigma_{i_2}\sigma_{j_2}))\ .$$

Observe that, if we let $\delta = 1, 2$,

$$\frac{\partial}{\partial \beta J'_\delta} = (1 - t^2_\delta)\frac{\partial}{\partial t_\delta}\ . \tag{25}$$

Now,

$$\ln W = \ln(1 + t_1\Omega_1 + t_2\Omega_2 + t_1 t_2 \Omega_{12}) + \ln\cosh\beta J'_1 + \ln\cosh\beta J'_2$$

and

$$\ln(1+t_1\Omega_1+t_2\Omega_2+t_1t_2\Omega_{12}) =$$
$$\sum_{n=1}^{\infty}\sum_{l=0}^{n}\sum_{m=0}^{l}\frac{(-)^{n+1}}{n}\binom{n}{l}\binom{l}{m}t_1^{n-l+m}t_2^{n-m}\Omega_1^m\Omega_2^{l-m}\Omega_{12}^{n-l}$$
$$=\sum_{n,l,m}(-)^{n+1}\frac{(n-1)!}{(n-l)!(l-m)!m!}t_1^{n-l+m}t_2^{n-m}\Omega_1^m\Omega_2^{l-m}\Omega_{12}^{n-l}\ .$$

The derivatives in (22) kill the two terms with the hyperbolic cosines, and from (25) we know that we can replace the derivatives with respect to $\beta J'_\delta$ with the derivatives with respect to t_δ, $\delta=1,2$. Notice that the logarithm just expanded is zero for $t_1=0$ and for $t_2=0$, therefore as its derivative like in (22) is zero, the logarithm itself is zero. This is why Theorem 3 holds, being (24) just the integral of the second line in (22).

Thanks to (6), if we put

$$n-l+m=r\ ,\ n-m=s\ ,\ n-l=a$$

we get

$$\sum_{r,s}\mathbb{E}[t_1^r t_2^s]\sum_{a=0}^{\min\{r,s\}}(-)^{a+1}\frac{(r+s-a-1)!}{a!(r-a)!(s-a)!}\langle q_r^2 q_s^2\rangle'_a=0$$

where, recall, $\langle\cdot\rangle_a$ means that a replicas are in common among those in q_r and those in q_s. Hence the statement of the theorem to be proven

$$\sum_{a=0}^{\min\{2r,2s\}}(-)^{a+1}\frac{(2r+2s-a-1)!}{a!(2r-a)!(2s-a)!}\langle q_{2r}^2 q_{2s}^2\rangle'_a=0\ ,$$

where only the terms with an even number of replicas in each overlap survive because of the symmetry of the variables J'_1, J'_2 in t_1, t_2. □

3.3 Generalization to smooth functions of multi-overlaps

The fact that in our formulas we always got the square power of the overlaps is due to the fact that the Hamiltonian has 2-spin interactions. Everything we did so far could then be reproduced in the case of p-spin interactions, and we would obtain the same relations just derived, except the overlaps would appear with the power p instead of 2. Clearly the perturbation needed in this case is a p-spin perturbation too. More in general, we could consider a Hamiltonian consisting of the sum (over p) of p-spin Hamiltonians for any integer p. Then we could perturb each of the p-spin Hamiltonians with its proper small p-spin perturbation, and add all these perturbations to the system. Clearly we have to make sure that all the terms in this whole Hamiltonian are weighted with sufficiently small weights so to have the necessary convergence. More explicitly, the perturbed Hamiltonian is

$$H_N(\sigma,\alpha;\mathcal{J}) = -\sum_p \left[a_p \sum_{\nu=1}^{P^{(p)}_{\alpha N}} J_\nu \sigma_{i^1_\nu} \cdots \sigma_{i^p_\nu} + b_p \lambda_p \sum_{\nu=1}^{P'^{(p)}_{\alpha'}} J'_\nu \sigma_{j^1_\nu} \cdots \sigma_{j^p_\nu} \right] ,$$

where $\sum_p |a_p|^2 = \sum_p |b_p|^2 = 1$, the notation for all the quenched variables is the usual one, and $\{\lambda_p\}$ are the independent perturbing real parameters.

It is not surprising then that we can state

Corollary 2 *The following constraints hold in integral average with respect to the set of all the perturbing parameters*

$$\sum_{a=0}^{\min\{2r,2s\}} (-)^{a+1} \frac{(2r+2s-a-1)!}{a!(2r-a)!(2s-a)!} \langle q^m_{2r} q^n_{2s} \rangle'_a = 0 \quad \forall\ r,s,m,n \in \mathbb{N}$$

in the thermodynamic limit.

Again, this corollary can be seen as a consequence of a self-averaging property, namely

$$\mathbb{E}\Omega'(\sigma_{i^1_1} \cdots \sigma_{i^m_1} \sigma_{j^1_1} \cdots \sigma_{j^n_1}) - \mathbb{E}[\Omega'(\sigma_{i^1_1} \cdots \sigma_{i^m_1}) \Omega'(\sigma_{j^1_1} \cdots \sigma_{j^n_1})] = 0 \ . \tag{26}$$

Therefore we can replace each overlap by any smooth function of the relative replicas in the statement of the corollaries.

As a last remark, notice that in [7] the strategy consisted in using the fact the second derivative of the free energy with respect to the "perturbing inverse temperatures" (β'_1, β'_2 in our case) is bounded to deduce self-averaging, and from the latter the identities. Here we used the connectivity as opposed to the inverse temperature to analyze stochastic stability, we then showed that the latter implies the self-averaging of [7]. So we obtained a comparison between self-averaging and stochastic stability (of a quite general validity), both providing a precious factorization of the Gibbs measure, and we also obtained that if any of the two is given, we know how to derive the same constraints.

4 Self-averaging of the quenched-Gibbs measure

Roughly speaking, if a convex random function does not fluctuate much, then its derivative does not fluctuate much either, with the exception of bad cases. This is well explained in Proposition 4.3 of [15] and Lemma 8.10 of [5]. We are not interested in general theorems here, in our case the convex function is the free energy density, and we only need to know that it is self-averaging (in the sense that the random free energy density does not fluctuate around its quenched expectation, in the thermodynamic limit). In the case of finite connectivity random spin systems like the VB model, a detailed proof of this can be found in [10]. The derivative of the free energy density (times $-\beta$) with respect to $-\beta$ is, in full generality, the expectation of the internal energy density $u_N = H_N/N$. Like in [9] and in section 2 of [8], we have therefore this further self-averaging (in integral average with respect to β)

$$\lim_{N\to\infty}[\langle u_N^2\rangle - \langle u_N\rangle^2] = 0$$

which implies (due to Schwartz inequality)

$$\lim_{N\to\infty}\langle u_N^{(1)}\phi_s\rangle = \lim_{N\to\infty}\langle u_N\rangle\langle\phi_s\rangle \tag{27}$$

for any bounded function ϕ_s of s replicas, and $u_N^{(1)}$ is the internal energy density in the configuration space of the replica 1. More precisely, the spin-configuration space is $\{-1,1\}^N = \Sigma$, and we consider a bounded function ϕ_s of s replicas, i.e. $\phi_s : \Sigma^s \to \mathbb{R}$. The product space Σ^s ("the space of the replicas") is equipped with the product Gibbs measure ("replica measure") Ω, but the quenched variables are the same in each factor of the product space, and this means that the measure $\langle\cdot\rangle = \mathbb{E}\Omega(\cdot)$ on the product space Σ^s is not a product measure. So $f_N^{(1)}$ is the free energy in the space which is the first factor in the product space Σ^s. Notice that Σ has the cardinality of the continuum in the thermodynamic limit $N\to\infty$.

Now we want to perturb the Hamiltonian

$$-\beta H_N(\sigma) \;\longrightarrow\; -\beta H_N(\sigma) + \beta' \sum_{\nu=1}^{P'\alpha} J'_\nu \sigma_{i'_\nu}\sigma_{j'_\nu} \ ,$$

and consider the derivative with respect to the perturbing parameter, as we did in the previous section in order to obtain an expansion in powers of β' with coefficients not depending on β' in the thermodynamic limit. As before, recall that we always need to take derivatives (with respect to the inverse temperature this time), which by convexity exist only almost everywhere, and therefore we will obtain equality only integrating back over a given interval [16].

We are going to prove, first of all, the following

Theorem 4 *For a given bounded function ϕ_s of s replicas, the following relation, holding in integral average with respect to the inverse temperature β, constrains the distribution of the 4-overlap*

$$\frac{s(s+1)(s+2)}{3!}\langle q^2_{1,s+1,s+2,s+3}\phi_s\rangle - \frac{s(s+1)}{2!}\sum_{a}^{2,s}\langle q^2_{1,a,s+1,s+2}\phi_s\rangle$$
$$+ s\sum_{a<b}^{2,s}\langle q^2_{1,a,b,s+1}\phi_s\rangle - \sum_{a<b<c}^{2,s}\langle q^2_{1,a,b,c}\phi_s\rangle = \langle q^2_{1234}\rangle\langle\phi_s\rangle \ .$$

The proof is straightforward but long, and it will be splitted into several steps.

Let us consider the right hand side of (27). Put $t = \tanh(\beta')$, $q_0 = 1$, and let us just indicate the number of replicas in the overlaps, rather than denumerating them all. The presence of the perturbation implies that the pressure and the free energy are functions, $\tilde{p}_N(\beta, \beta', \alpha)$ and $\tilde{f}_N(\beta, \beta', \alpha)$ respectively, of both β and β', and recall also that according to our notations $\tilde{p}_N(\beta, \beta', \alpha) = -\beta \tilde{f}_N(\beta, \beta', \alpha)$. Let us prove the next

Lemma 3 *The derivative of the (perturbed) pressure $\tilde{p}_N(\beta, \beta', \alpha)$ with respect to the perturbing parameter β' has the following form as a series in powers of $t = \tanh(\beta')$*

$$\partial_{\beta'} \tilde{p}_N(\beta, \beta', \alpha) = -\alpha \sum_{n=0}^{\infty} t^{2n+1} (\langle q_{2n}^2 \rangle - \langle q_{2n+2}^2 \rangle) \ . \tag{28}$$

Proof. We have

$$\begin{aligned}
\partial_{\beta'} \tilde{p}_N(\beta, \beta', \alpha) &= -\sum_{m=1}^{\infty} \pi_\alpha(m) \sum_{\nu=1}^{m} \langle J'_\nu \sigma_{i'_\nu} \sigma_{j'_\nu} \rangle_m \\
&= -\sum_{m=1}^{\infty} m \pi_\alpha(m) \langle J'_m \sigma_{i'_m} \sigma_{j'_m} \rangle_m \\
&= -\alpha \sum_{m=1}^{\infty} \pi_\alpha(m-1) \langle J'_m \sigma_{i'_m} \sigma_{j'_m} \rangle_m
\end{aligned}$$

where the sub m indicates that the variable P'_α has been fixed to m. It is easy to see that

$$\langle J'_m \sigma_{i'_m} \sigma_{j'_m} \rangle_m = \mathbb{E} \frac{\Omega(J'_m \sigma_{i'_m} \sigma_{j'_m} \exp(\beta J'_m \sigma_{i'_m} \sigma_{j'_m}))_{m-1}}{\Omega(\exp(\beta J'_m \sigma_{i'_m} \sigma_{j'_m}))_{m-1}} \ . \tag{29}$$

Hence

$$\partial_{\beta'} \tilde{p}_N(\beta, \beta', \alpha) = -\alpha \mathbb{E} J' \frac{t + w}{1 + tw} \ , \ w \equiv \Omega(\sigma_{i'_m} \sigma_{j'_m}) \ , \tag{30}$$

according to the usual notations. Now a simple expansion (that we will explicitly write in the next lemma) of $(1 + tw)^{-1}$ in powers of t yields

$$\partial_{\beta'} \tilde{p}_N(\beta, \beta', \alpha) = -\alpha \sum_{n=0}^{\infty} t^{2n+1} (\langle q_{2n}^2 \rangle - \langle q_{2n+2}^2 \rangle) \ . \tag{31}$$

So the lemma is proven and we have an expression for the right hand side of (27), if we just multiply the average of the multi-overlaps by the average of ϕ_s. □

Let us now consider the left hand side of (27), recalling that ϕ_s is a function of s replicas, that indices in the spins indicate which factor of the product space Σ^s (which replica) the spin belongs to, and that the energy density is assumed to be taken in the first replica.

Lemma 4 *Recalling that* $w \equiv \Omega(\sigma_{i'_m}\sigma_{j'_m})$, *we have*

$$
\begin{aligned}
\langle u_N^{(1)}\phi_s\rangle = -\alpha t\mathbb{E}\{\Omega[\phi_s(1+J't^{-1}\sigma_{i_1}^{(1)}\sigma_{j_1}^{(1)})\times \\
&(1+J'\sum_{a}^{2,s}\sigma_{i_1}^{(a)}\sigma_{j_1}^{(a)}t+\sum_{a<b}^{2,s}\sigma_{i_1}^{(a)}\sigma_{i_1}^{(b)}\sigma_{j_1}^{(a)}\sigma_{j_1}^{(b)}t^2+ \\
&J'\sum_{a<b<c}^{2,s}\sigma_{i_1}^{(a)}\sigma_{i_1}^{(b)}\sigma_{i_1}^{(c)}\sigma_{j_1}^{(a)}\sigma_{j_1}^{(b)}\sigma_{j_1}^{(c)}t^3+\cdots)]\times \\
&(1-J'stw+\frac{s(s+1)}{2!}t^2w^2-J'\frac{s(s+1)(s+2)}{3!}t^3w^3 \\
&+\frac{s(s+1)(s+2)(s+3)}{4!}t^4w^4-\cdots)\}\ .
\end{aligned}
$$

Proof. From the proof of the previous lemma, in particular equations (29)-(30), and by definition of replica measure, we immediately get

$$
\langle u^{(1)}\phi_s\rangle = -\alpha\mathbb{E}\frac{\Omega[J'\sigma_{i_1}^{(1)}\sigma_{j_1}^{(1)}\exp(\beta J'(\sigma_{i_1}^{(1)}\sigma_{j_1}^{(1)}+\cdots+\sigma_{i_1}^{(s)}\sigma_{j_1}^{(s)}))\phi_s]}{\Omega^s(\exp(\beta J'\sigma_{i_1}\sigma_{j_1}))}\ , \tag{32}
$$

that we rewrite as

$$
\langle u^{(1)}\phi_s\rangle = -\alpha\mathbb{E}t\frac{\Omega[(1+J't^{-1}\sigma_{i_1}^{(1)}\sigma_{j_1}^{(1)})\prod_{a=2}^{s}(1+J't\sigma_{i_1}^{(a)}\sigma_{j_1}^{(a)})\phi_s]}{(1+J'tw)^s}\ . \tag{33}
$$

Let us write explicitly the power expansion of the denominator, that we omitted in the previous lemma

$$
\begin{aligned}
\frac{1}{(1+J'tw)^s} &= 1-J'stw+\frac{s(s+1)}{2!}t^2w^2- \\
&J'\frac{s(s+1)(s+2)}{3!}t^3w^3+\frac{s(s+1)(s+2)(s+3)}{4!}t^4w^4\cdots\ .
\end{aligned}
$$

It is also clear that

$$
\begin{aligned}
\prod_{a=2}^{s}(1+J't\sigma_{i_1}^{(a)}\sigma_{j_1}^{(a)}) &= 1+J'\sum_{a}^{2,s}\sigma_{i_1}^{(a)}\sigma_{j_1}^{(a)}t+\sum_{a<b}^{2,s}\sigma_{i_1}^{(a)}\sigma_{i_1}^{(b)}\sigma_{j_1}^{(a)}\sigma_{j_1}^{(b)}t^2 \\
&+J'\sum_{a<b<c}^{2,s}\sigma_{i_1}^{(a)}\sigma_{i_1}^{(b)}\sigma_{i_1}^{(c)}\sigma_{j_1}^{(a)}\sigma_{j_1}^{(b)}\sigma_{j_1}^{(c)}t^3+\cdots\ .
\end{aligned}
$$

Gathering all the ingredients completes the proof of the lemma. □

We are now able to compare the two sides of (27), and see what the self-averaging of the internal energy density in the thermodynamic limit brings.

Equating the expressions computed in the last two lemmas gives

$$
\begin{aligned}
\sum_{n=0}^{\infty} t^{2n}(\langle q_{2n}^2\rangle - \langle q_{2n+2}^2\rangle)\langle\phi_s\rangle = \mathbb{E}\{\Omega[\phi_s(1+Jt^{-1}\sigma_{i_1}^{(1)}\sigma_{j_1}^{(1)})& \\
(1+J'\sum_{a}^{2,s}\sigma_{i_1}^{(a)}\sigma_{j_1}^{(a)}t+\sum_{a<b}^{2,s}\sigma_{i_1}^{(a)}\sigma_{i_1}^{(b)}\sigma_{j_1}^{(a)}\sigma_{j_1}^{(b)}t^2+& \\
J'\sum_{a<b<c}^{2,s}\sigma_{i_1}^{(a)}\sigma_{i_1}^{(b)}\sigma_{i_1}^{(c)}\sigma_{i_1}^{(a)}\sigma_{i_1}^{(b)}\sigma_{j_1}^{(c)}t^3+& \\
\cdots+J'^{s-1}t^{s-1}\sigma_{i_1}^{(2)}\cdots\sigma_{i_1}^{(s)}\sigma_{j_1}^{(2)}\cdots\sigma_{j_1}^{(s)})]& \\
(1-J'stw+\frac{s(s+1)}{2!}t^2w^2-J'\frac{s(s+1)(s+2)}{3!}t^3w^3& \\
+\frac{s(s+1)(s+2)(s+3)}{4!}t^4w^4-\cdots)\}\ .& \qquad (34)
\end{aligned}
$$

The equality holds for any smooth function ϕ_s (typical interesting information is obtained for $\phi_s \equiv 1$ or $\phi_{s=2n} = q_{2n}^2$), so that we get equalities between expressions involving averages of (squared) overlaps.

Let us see in detail what information we can get from the lowest orders.

Denote by $\mathbb{E}(\cdot|\mathcal{A}_s)$ the conditional expectation with respect to the sigma-algebra $\mathcal{A}_s$ generated by the overlaps of s replicas. Let us show that the usual [8] Ghirlanda-Guerra identities for the overlap hold in our quite general case too:

Proposition 1 *The Ghirlanda-Guerra relation holds*

$$
\mathbb{E}(q_{a,s+1}^2|\mathcal{A}_s) = \frac{1}{s}\langle q_{12}^2\rangle + \frac{1}{s}\sum_{b\neq a} q_{ab}^2 \qquad (35)
$$

in integral average with respect to the inverse temperature β.

Proof. In the expansion (34), where only the terms of even order survive due to the symmetry of the variables J, at the lowest order in t one gets

$$
\begin{aligned}
\langle\phi_s\rangle - \langle q_{12}^2\rangle\langle\phi_s\rangle &= \langle\phi_s\rangle - s\mathbb{E}[\omega(\sigma_{i_1}^{(1)}\sigma_{j_1}^{(1)})w\phi_s] + \sum_{a}^{2,s}\mathbb{E}[\Omega(\sigma_{i_1}^{(1)}\sigma_{i_1}^{(a)}\sigma_{j_1}^{(1)}\sigma_{j_1}^{(a)})\phi_s] \\
&= \langle\phi_s\rangle - s\langle q_{1,s+1}^2\phi_s\rangle + \sum_{a}^{2,s}\langle q_{1a}^2\phi_s\rangle\ ,
\end{aligned}
$$

which is precisely what is stated in (35), (see [16]), immediately completing the proof of the proposition. □

So the usual Ghirlanda-Guerra identities for 2-overlaps are recovered and proven to hold in dilute spin glasses too.

At the next order we get instead

$$
\begin{aligned}
\langle q_{12}^2\rangle\langle\phi_s\rangle - \langle q_{1234}^2\rangle\langle\phi_s\rangle = \sum_{a<b}^{2,s}\langle q_{ab}^2\phi_s\rangle + \frac{s(s+1)}{2!}\langle q_{s+1,s+2}^2\phi_s\rangle \\
- s\sum_{a}^{2,s}\langle q_{a,s+1}^2\phi_s\rangle - \frac{s(s+1)(s+2)}{3!}\langle q_{1,s+1,s+2,s+3}^2\phi_s\rangle \\
+ \frac{s(s+1)}{2!}\sum_{a}^{2,s}\langle q_{1,a,s+1,s+2}^2\phi_s\rangle - s\sum_{a<b}^{2,s}\langle q_{1,a,b,s+1}^2\phi_s\rangle + \sum_{a<b<c}^{2,s}\langle q_{1,a,b,c}^2\phi_s\rangle \ . \quad (36)
\end{aligned}
$$

Now consider the four 2-overlaps terms. A simple generalization of the usual Ghirlanda-Guerra relations [8] to the case when two replicas are added to a previously assigned set of other replicas, tells us that these terms cancel out. Let us check that explicitly.

Corollary 3 *Relation (35) implies*

$$
\mathbb{E}(q_{s+1,s+2}^2|\mathcal{A}_s) = \frac{2}{s+1}\langle q_{12}^2\rangle + \frac{2}{s(s+1)}\sum_{a<b}^{1,s} q_{ab}^2 \ , \qquad (37)
$$

under the same conditions, i.e. in integral average with respect to β.

Proof. Let us re-write (35) in the case of $s+1$ given replicas

$$
\mathbb{E}(q_{s+1,s+2}^2|\mathcal{A}_{s+1}) = \frac{1}{s+1}\langle q_{12}^2\rangle + \frac{1}{s+1}\sum_{b}^{1,s} q_{b,s+1}^2 \ .
$$

Now use

$$
\mathbb{E}(\mathbb{E}(\cdot|\mathcal{A}_{s+1})|\mathcal{A}_s) = \mathbb{E}(\cdot|\mathcal{A}_s) \qquad (38)
$$

to get

$$
\begin{aligned}
\mathbb{E}(q_{s+1,s+2}^2|\mathcal{A}_s) &= \frac{1}{s+1}\langle q_{12}^2\rangle + \frac{1}{s+1}\sum_{b}^{1,s}\mathbb{E}(q_{b,s+1}^2|\mathcal{A}_s) \\
&= \frac{1}{s+1}\langle q_{12}^2\rangle + \frac{1}{s+1}\left(\langle q_{12}^2\rangle + \frac{1}{s}\sum_{b}^{1,s}\sum_{c\neq b}^{1,s} q_{bc}^2\right) \ .
\end{aligned}
$$

That is

$$
\mathbb{E}(q_{s+1,s+2}^2|\mathcal{A}_s) = \frac{2}{s+1}\langle q_{12}^2\rangle + \frac{2}{s(s+1)}\sum_{a<b}^{1,s} q_{ab}^2 \ , \qquad (39)
$$

which is what we wanted to prove. □

Now with (35) and (37) in our hands, let us take the three 2-overlap terms in the right hand side of (36)

$$
\begin{aligned}
\frac{s(s+1)}{2}\langle q^2_{s+1,s+2}\phi_s\rangle &= s\langle q^2_{12}\rangle\langle\phi_s\rangle + \sum_{a<b}^{1,s}\langle q^2_{ab}\phi_s\rangle \\
-s\sum_{a}^{2,s}\langle q^2_{a,s+1}\phi_s\rangle &= -s\sum_{a}^{1,s}\langle q^2_{a,s+1}\phi_s\rangle + s\langle q^2_{1,s+1}\phi_s\rangle \\
&= -s\langle q^2_{12}\rangle\langle\phi_s\rangle - \sum_{a}^{1,s}\sum_{b\neq a}^{1,s}\langle q^2_{ab}\phi_s\rangle + \langle q^2_{12}\rangle\langle\phi_s\rangle + \sum_{a}^{2,s}\langle q^2_{1a}\phi_s\rangle \\
\sum_{a<b}^{2,s}\langle q^2_{ab}\phi_s\rangle &= \sum_{a<b}^{1,s}\langle q^2_{ab}\phi_s\rangle - \sum_{a}^{2,s}\langle q^2_{1a}\phi_s\rangle \ .
\end{aligned}
$$

The sum of these three terms cleary reduces to $\langle q^2_{12}\rangle\langle\phi_s\rangle$, which is precisely what we find in the left hand side of (36). The 2-overlap terms thus cancel out from (36). We are hence left with a new relation for 4-overlaps:

$$
\begin{aligned}
&\frac{s(s+1)(s+2)}{3!}\langle q^2_{1,s+1,s+2,s+3}\phi_s\rangle - \frac{s(s+1)}{2!}\sum_{a}^{2,s}\langle q^2_{1,a,s+1,s+2}\phi_s\rangle \\
&\qquad + s\sum_{a<b}^{2,s}\langle q^2_{1,a,b,s+1}\phi_s\rangle = \langle q^2_{1234}\rangle\langle\phi_s\rangle + \sum_{a<b<c}^{2,s}\langle q^2_{1,a,b,c}\phi_s\rangle \ ,
\end{aligned}
$$

and the proof of Theorem 4 is now complete. □

We report for sake of completeness the general expression of the generic order in the power series expansion (34). From the explicit calculation in Lemma 4 we get

$$
\begin{aligned}
&\langle q^2_{2n}\rangle\langle\phi_s\rangle - \langle q^2_{2n+2}\rangle\langle\phi_s\rangle = \\
&\qquad \sum_{m=2n-s+1}^{2n}\sum_{l=0}^{s-1}\sum_{a_1<\cdots<a_l}^{2,s}(-)^m\binom{s+m+1}{m}\times \\
&\qquad\qquad \mathbb{E}[w^m\Omega(\phi_s\sigma^{a_1}_{i_1}\cdots\sigma^{a_l}_{i_1}\sigma^{a_1}_{j_1}\cdots\sigma^{a_l}_{j_1})]\delta_{2n,m+l} \\
&\qquad + \sum_{m=2n-s+2}^{2n+1}\sum_{l=0}^{s-1}\sum_{a_1<\cdots<a_l}^{2,s}(-)^m\binom{s+m+1}{m}\times \\
&\qquad\qquad\qquad \mathbb{E}[w^m\Omega(\phi_s\sigma^1_{i_1}\sigma^1_{j_1}\sigma^{a_1}_{i_1}\cdots\sigma^{a_l}_{i_1}\sigma^{a_1}_{j_1}\cdots\sigma^{a_l}_{j_1})]\delta_{2n,m+l-1}
\end{aligned}
$$

which becomes, denoting by $x\wedge y$ the minumum between x and y,

$$
\begin{aligned}
&\langle q^2_{2n}\rangle\langle\phi_s\rangle - \langle q^2_{2n+2}\rangle\langle\phi_s\rangle = \sum_{l=0}^{2n\wedge s-1}\sum_{a_1<\cdots<a_l}^{2,s}(-)^{2n-l}\binom{2n+s-l+1}{2n-l}\times \\
&[\langle\phi_s q^2_{a_1\cdots a_l}q^2_{s+1\cdots s+2n-l}\rangle - \frac{2n-l+s+2}{2n-l+1}\langle\phi_s q^2_{1a_1\cdots a_l}q^2_{s+1\cdots s+2n-l+1}\rangle] \ . \qquad (40)
\end{aligned}
$$

In both the expressions above the term for $l=0$ is understood to be one.

The right hand side of (40), due to the presence of $1+Jt^{-1}\sigma$ in the right hand side of (34) - along with the symmetry of J, makes the expansion somewhat recursive. This means that at each order we find some terms already found in the previous order. More precisely, we claim without proving that at each $2n$-th order of the expansion, all the terms involving $2m$-overlaps with $2m \leq 2n$ cancel out thanks to a repeated use of (38) with the relations coming from the lower orders. Hence from the $2n$-th order we get new relations involving $2n+2$-overlaps only. This is what we explicitly verified only for 4-overlaps in the previous pages. More explicitly, if we re-write the difference in the right hand side of (40) as

$$\langle q_{2n}^2\rangle\langle\phi_s\rangle - \langle q_{2n+2}^2\rangle\langle\phi_s\rangle = c_{2n} - d_{2n+2} \ ,$$

we have

$$\langle q_{2n}^2\rangle\langle\phi_s\rangle = c_{2n} \ , \ \langle q_{2n+2}^2\rangle\langle\phi_s\rangle = d_{2n+2} \ , \ c_{2n} = d_{2n} \ .$$

So that the final formula becomes

$$\langle q_{2n}^2\rangle\langle\phi_s\rangle = \sum_{l=0}^{2n\wedge s-1} \sum_{a_1<\cdots<a_l}^{2,s} (-)^{2n-l} \binom{2n+s-l+1}{2n-l} \langle q_{a_1\cdots a_l}^2 q_{s+1\cdots s+2n-l}^2 \phi_s\rangle \ ,$$

and holds in integral average, as always.

4.1 generalization to smooth functions of multi-overlaps

Just like for the family of identities discussed in the previous section, we started our analysis with the most natural quantity: the energy of our model with 2-spin interactions. And so we got again some relations for the squared multi-overlaps. But we already know how to generalize these formulas to smooth functions of the overlaps. We can consider p-spin interactions, and the procedure would provide us with the same relations for the p-th power of the overlaps. Then, as already explained, we can take a convergent sum over all integer p of p-spin Hamiltonians, and consider the self-averaging of the desired one among them. The perturbed Hamiltonian is again

$$H_N(\sigma,\alpha;\mathcal{J}) = -\sum_p \left[a_p \sum_{\nu=1}^{P_{\alpha N}^{(p)}} J_\nu \sigma_{i_\nu^1}\cdots\sigma_{i_\nu^p} + b_p\lambda_p \sum_{\nu=1}^{P'^{(p)}_{\alpha'}} J'_\nu \sigma_{j_\nu^1}\cdots\sigma_{j_\nu^p} \right] \ ,$$

where $\sum_p |a_p|^2 = \sum_p |b_p|^2 = 1$, the notation for all the quenched variables is the usual one, and $\{\lambda_p\}$ are the independent perturbing real parameters. As a side remark, we just point out that (like in [8]), in the case of this second family of identities it is not necessary to consider a Hamiltonian consisting of the sum of all possible p-spin Hamiltonians: only the perturbation must be so.

Let us remind once again that all the identities we provided hold in integral average with respect to all the variables used to compute derivatives. If more that one variables is used in some derivatives, the relations hold under integral performed simultaneously against all the variables, over the Cartesian product of the chosen intervals of variation of each variable.

Concluding remarks

Notice that while we derived our identities having as reference diluted spin glasses, all that matters in the derivation are the properties of the perturbing Hamiltonian, and they are therefore generically valid (in integral average with respect to the perturbing parameters).

The Ghirlanda-Guerra identities for the overlap have been useful to prove non trivial properties of mean-field spin glasses. For instance Talagrand could prove [16] that for all models where the identities are valid, the overlap is positive with probability one. This positivity property is important as it enters in the Guerra free-energy bounds in spin systems without spin reversal symmetry, like in the case of odd-spin interactions. Unfortunately the derivation of Talagrand for the overlap does not extend immediately to the multi-overlap case. We believe however that the self-averaging identity will be useful in the mathematical analysis of diluted spin models.

Acknowledgments

LDS thanks Fabio Lucio Toninelli and Anton Bovier for useful discussions.

References

[1] A. Agostini, A. Barra, L. De Sanctis, *Positive-Overlap Transition and Critical Exponents in Mean Field Spin Glasses*, J. Stat. Mech. (2006) **P11015**.

[2] M. Aizenman, P. Contucci, *On the stability of the quenched state in mean field spin glass models*, J. Stat. Phys. **92**, 765 (1998).

[3] A. Barra, L. De Sanctis, *Stability properties and probability distribution of multi-overlaps in dilute spin glasses*, J. Stat. Mech. (2007) **P08025**.

[4] A. Bovier, *Statistical Mechanics of Disordered Systems, A mathematical perspective*, Cambridge University Press (2006).

[5] A. Bovier, V. Gayrard, *Hopfield Models as Generalized Random Mean Field Models*, Arxiv:cond-mat/9607103 (1997), in A. Bovier and P. Picco Eds., *Mathematical Aspects of Spin Glasses and Neural Networks*, Progress in Probability **41** Birkäuser, Boston-Basel-Berlin (1998).

[6] L. De Sanctis, *Random Multi-Overlap Structures and Cavity Fields in Diluted Spin Glasses*, J. Stat. Phys. **117** 785-799 (2004)

[7] S. Franz, M. Leone, F.L. Toninelli, *Replica bounds for diluted non-Poissonian spin systems*, J. Phys. A **36** 10967 (2003).

[8] S. Ghirlanda, F. Guerra, *General properties of overlap distributions in disordered spin systems. Towards Parisi ultrametricity*, J. Phys. A, **31** 9149-9155 (1998).

[9] F. Guerra, *About the overlap distribution in mean field spin glass models*, Int. Jou. Mod. Phys. B **10**, 1675-1684 (1996).

[10] F. Guerra, F.L. Toninelli, *The high temperature region of the Viana-Bray diluted spin glass model*, J. Stat. Phys. **115** (2004).

[11] M. Mézard, A. Montanari, *Reconstruction on trees and spin glass transition*, ArXiv:cond-mat/0512295.

[12] M. Mézard, G. Parisi and M. A. Virasoro, *Spin glass theory and beyond*, World Scientific, Singapore (1987).

[13] M. Mézard, G. Parisi and R. Zecchina, *Analytic and Algorithmic Solution of Random Satisfiability Problems*, Science **247** 812-815 (2002).

[14] G. Parisi, *On the probabilistic formulation of the replica approach to spin glasses*, ArXiv:cond-mat/9801081.

[15] M. Talagrand, *The Sherrington Kirkpatrick model: a challenge for mathematicians*, Probab. Rel. Fields **110**, 109-176 (1998).

[16] M. Talagrand, *Spin glasses: a challenge for mathematicians. Cavity and Mean field models*, Springer Verlag (2003).