Study of the critical properties of the Ising linear perceptron and signal detection in CDMA, by message passing technique

Juan P. Neirotti and David Saad

The Neural Computing Research Group,

Aston University, Birmingham B4 7ET, UK.

Abstract

An efficient new Bayesian inference technique is employed for studying critical properties of the Ising linear perceptron and for multiuser detection in Code Division Multiple Access (CDMA). The approach is based on a message passing technique for densely connected systems that requires some assumption about the symmetry of the solutions sought. Here we extend the previous derivation based on a replica symmetric (RS) like structure to include a more complex one step replica symmetry breaking (1RSB) ansatz. Results obtained under the RS assumption in the non-critical regime give rise to a highly efficient signal detection algorithm in the context of CDMA; while in the critical regime one observes a first order transition line that ends in a continuous phase transition point. Finite size effects are observed. While the 1RSB ansatz is not required for the original problems we investigated, it was applied to a variant of the CDMA problem that exhibits RSB behaviour, resulting in a considerable improvement in performance.

PACS numbers: 89.70.+c, 75.10.Nr, 64.60.Cn

I. INTRODUCTION

Efficient inference in large complex systems is a major challenge with significant implications in science, engineering and computing. Exact inference is computationally hard in complex systems and a range of approximation methods have been devised over the years, many of which have been originated in the physics literature. A recent review [1] highlights the links between the various approximation methods and their applications.

In the current paper, we methodologically extend a method that was introduced only recently [2] for inference in dense graphs using message passing techniques. The method has been employed previously only in the non-critical regime [3], using the most basic (RS-like) ansatz for the solution structure. In the current paper we study both critical and non-critical regimes and extend the solution structure considered to include step replica symmetry breaking (1RSB) like structures. We apply the method to two different but related problems: signal detection in Code Division Multiple Access (CDMA) and learning in the Ising linear perceptron (ILP) to demonstrate its perfromance and relevance to general inference tasks. We investigate both RS and 1RSB-like structures. The former is applied to both CDMA and ILP problems and seems to be sufficient for obtaining optimal performances; the latter is applied to a variant of the CDMA signal detection problem that exhibits RSB-like behaviour to demonstrate its efficacy for such inference tasks.

In section II we will introduce the general models studied, followed by a brief review of message passing techniques for dense systems in section IIIpassing}. The general derivation of our approach will be presented in section IV and numerical studies for both CDMA signal detection and ILP learning will be reported in section V. To demonstrate the method based on the more complex 1RSB solution structure, and to examine its efficacy to problems that require such structures, we will introduce a variant of the CDMA signal detection problem and study it numerically in section VI. We will conclude the presentation with a summary and point to future research directions. Details of the derivation will be provided in Appendices VII-VII.

II. MODELS STUDIED

We apply the method to two different but related problems: signal detection in CDMA and learning in the Ising linear perceptron (ILP).

Multiple access communication refers to the transmission of multiple messages to a single receiver. The scenario we study here, described schematically in figure 1(a), is that of K users transmitting independent messages over an additive white Gaussian noise (AWGN) channel of zero mean and variance σ_0^2. Various methods are in place for separating the messages, in particular Time, Frequency and Code Division Multiple Access [4]. The latter, is based on spreading the signal by using K individual random binary spreading codes of spreading factor N. We consider the large-system limit, in which the number of users K tends to infinity while the system load $\beta \equiv K/N$ is kept to be $\mathcal{O}(1)$. We focus on a CDMA system using binary phase shift keying (BPSK) symbols and will assume the power is completely controlled to unit energy. The received aggregated, modulated and corrupted signal is of the form:

$$y_\mu = \frac{1}{\sqrt{N}} \sum_{k=1}^{K} s_{\mu k} b_k + \sigma_0 n_\mu \qquad (1)$$

where b_k is the bit transmitted by user k, $s_{\mu k}$ is the spreading chip value, n_μ is the Gaussian noise variable drawn from $\mathcal{N}(0,1)$, and y_μ the received message. This process is reminiscent of the learning task performed by a perceptron with binary weights and linear output.

Learning in neural networks has attracted considerable theoretical interest. In particular we focus on supervised learning from examples, which relies on a training set consisting of examples of the target task [5]. We consider a perceptron, described schematically in figure 1(b), which is a network that sums a single layer of inputs $s_{\mu k}$ with synaptic weights b_k and passes the result through a transfer function y_μ

$$y_\mu = g\left(\frac{1}{\sqrt{K}} \sum_{k=1}^{K} s_{\mu k} b_k\right), \qquad (2)$$

where g is a suitable function. If $g(x) = x$ the network is termed *linear output perceptron*. If the weights $b_k \in \{\pm 1\}$ the network is called *Ising perceptron*. Learning is a search through the weight space for the perceptron that best approximates a target rule.

The similarity between the linear perceptron of Eq. (2) and the CDMA detection problem of Eq. (1) allows for a direct relation between the two problems to be established. The main difference between the problems is the regime of interest. While CDMA detection applications are of interest mainly for non-critical low load values, ILP studies focused on the critical regime. We consider both regimes in this paper, but to unify the treatment we will use the notation and scaling conventions of the CDMA system.

III. MESSAGE PASSING FOR INFERENCE IN DENSELY CONNECTED SYSTEMS

passing}

Graphical models (Bayes belief networks) provide a powerful framework for modelling statistical dependencies between variables [6–8]. They play an essential role in devising a principled probabilistic framework for inference in a broad range of applications.

Message passing techniques are typically used for inference in graphical models that can be represented by a sparse graph with a few (typically long) loops. They are aimed at obtaining (pseudo) posterior estimates for the system's variables by iteratively passing messages (locally calculated conditional probabilities) between variables. Iterative message passing of this type is guaranteed to converge to the globally correct estimate when the system is tree-like; there are no such guarantees

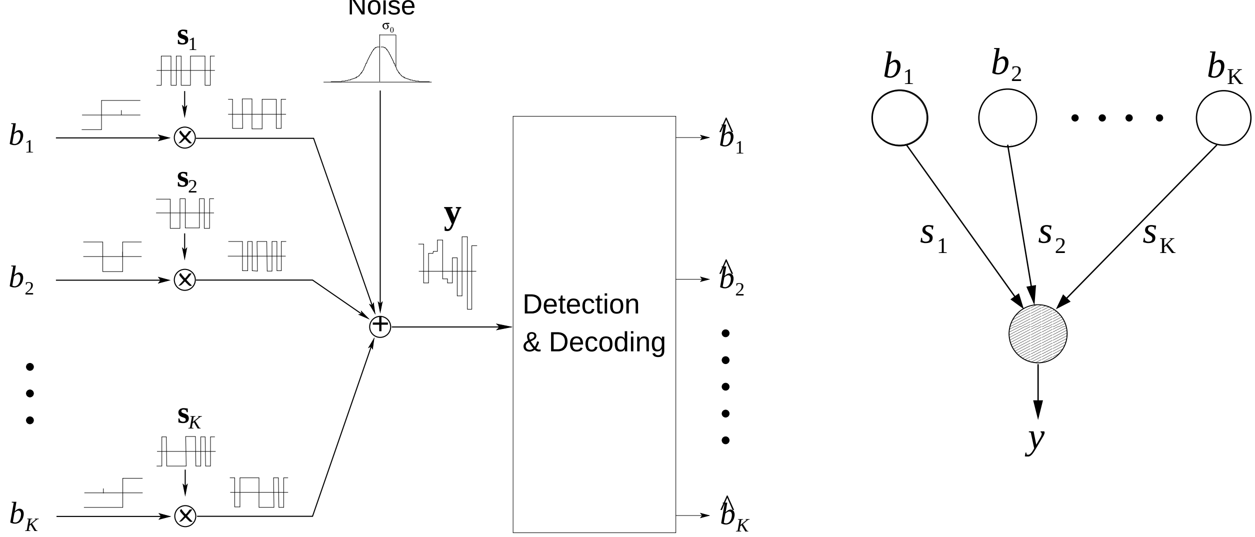

Figure 1: Schematic representation of (a) the CDMA system. (b) the ILP.

for systems with loops even in the case of large loops and a local tree-like structure (although message passing techniques have been used successfully in loopy systems, supported by some limited theory [9]). A clear link has been established between certain message passing algorithms and well known methods of statistical mechanics [1] such as the Bethe approximation [10, 11].

These inherent limitations seem to prevent the use of message passing techniques in densely connected systems due to their high connectivity, implying an exponentially growing cost, and an exponential number of loops. However, an exciting new approach has been recently suggested [2] for extending Belief Propagation (BP) techniques [6–8] to densely connected systems. In this approach, messages are grouped together, giving rise to a macroscopic random variable, drawn from a Gaussian distribution of varying mean and variance for each of the nodes. The technique has been successfully applied to CDMA signal detection problems and the results reported are competitive with those of other state of the art techniques. However, the current approach has some inherent limitations [2], presumably due to its similarity to the replica symmetric solution in equivalent Ising spin models [12, 13].

In a separate recent development [14], the replica-symmetric-equivalent BP has been extended to Survey Propagation (SP), which corresponds to one-step replica symmetry breaking in diluted systems. This new algorithm, motivated by the theoretical physics interpretation of such problems, has been highly successful in solving hard computational problems [14], far beyond other existing approaches. In addition, the algorithm facilitated theoretical studies of the corresponding physical system and contributed to our understanding of it [15]. The SP algorithm has recently been modified to handle Ising and multilayer perceptrons [16].

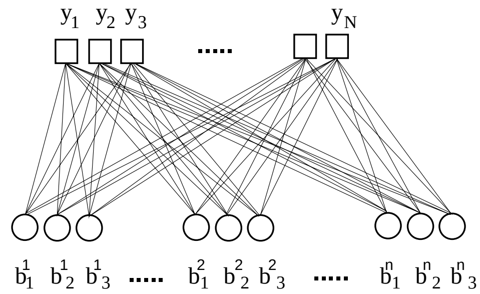

Figure 2: Replicated solutions $\mathbf{B}=(\mathbf{b}_1,\, \mathbf{b}_2, .., \mathbf{b}_K)$ given data.

IV. GENERAL FORMALISM

Recently, we presented a new approach [3], which was inspired by both, the extension of BP to densely connected graphs and the introduction of SP. The systems we consider here are characterised by multiplicity of pure states and a possible fragmentation of the space of solutions. To address the inference problem in such cases we consider an ensemble of replicated systems where averages are taken over the ensemble of potential solutions. This amounts to the presentation of a new graph, where the observables y_μ are linked to variables in all replicated systems, namely $\mathbf{B} = (\mathbf{b}^1, \mathbf{b}^2, \ldots, \mathbf{b}^n)$; where $\mathbf{b}^a = (b_1^a, b_2^a, \ldots, b_K^a)^\mathsf{T}$, as shown in Figure 2. To estimate the variables $\mathbf{B}$ given the data $\mathbf{y}^\mathsf{T} = (y_1, y_2, \ldots, y_N)$, in a Bayesian framework, we have to maximise the posterior $P(\mathbf{B}|\mathbf{y}) \propto \prod_{\mu=1}^{N} P(y_\mu|\mathbf{B})\, P(\mathbf{B})$, where we have considered independent data, and thus $P(\mathbf{y}|\mathbf{B}) = \prod_{\mu=1}^{N} P(y_\mu|\mathbf{B})$.

The likelihood so defined is of a general form; the explicit expression depends on the particular problem studied. Here, we are interested in cases where $\mathbf{b} \in \{\pm 1\}^K$ is an unbiased vector and $P(\mathbf{B}) = 2^{-Kn}$. The estimate we would like to obtain is the maximiser of the posterior marginal (MPM) $\widehat{\mathbf{b}}_k = \mathrm{argmax}_{\mathbf{b}_k \in \{\pm1\}^n} \sum_{\{\mathbf{b}_{l \neq k}\}} P(\mathbf{B}|\mathbf{y})$, which is expected to be a vector with equal entries for all replica $\widehat{b}_k^1 = \widehat{b}_k^2 = \cdots = \widehat{b}_k^n$. The number of operations required to obtain the full MPM estimator is of $\mathcal{O}\left(2^K\right)$ which is infeasible for large K values.

To obtain an approximate MPM estimate we apply a message passing technique such as BP [6–8]. In particular we are interested here in the application of BP on densely connected graphs, similar to the one presented in [2]. The latter is based on estimating a single solution and is therefore bound to fail, as have been observed, when the solution space becomes fragmented and multiple solutions emerge. This arguably corresponds to the replica symmetry breaking phenomena and occurs, for instance, when the noise level is unknown in the CDMA signal detection case.

A natural improvement is achieved by the introduction of SP [14] in problems that can be mapped onto sparsely connected graphs, which has recently been extended to include densely connected graphs [16].

Using Bayes rule one straightforwardly obtains the BP equations:

$$P^{t+1}\left(y_\mu|\mathbf{b}_k,\{y_{\nu\neq\mu}\}\right) = \sum_{\{\mathbf{b}_{l\neq k}\}} P\left(y_\mu|\mathbf{B}\right)\prod_{l\neq k} P^t\left(\mathbf{b}_l|\,\{y_{\nu\neq\mu}\}\right) \tag{3}$$

$$P^t\left(\mathbf{b}_l|\,\{y_{\nu\neq\mu}\}\right) \propto \prod_{\nu\neq\mu} P^t\left(y_\nu|\mathbf{b}_l,\{y_{\sigma\neq\nu}\}\right)P\left(\mathbf{b}_l\right) . \tag{4}$$

For calculating the posterior $P\left(\mathbf{y}|\mathbf{B}\right)$, we will suppose that the dependency of the data on the parameters is $y_\mu = \mathcal{F}\left(\sum_{l=1}^K \varepsilon_{\mu l}\mathbf{b}_l;\boldsymbol{\gamma}\right)$, where $\mathcal{F}$ is some general smooth function, $\boldsymbol{\gamma}$ are model parameters and the $\varepsilon_{\mu l}$ are small enough to ensure that $\sum_{l=1}^K \varepsilon_{\mu l} b_l^{\mathrm{a}} \sim \mathcal{O}(1)$. We define the vector $\boldsymbol{\Delta}_\mu \equiv \sum_{l=1}^K \varepsilon_{\mu l}\mathbf{b}_l = \sum_{l\neq k}\varepsilon_{\mu l}\mathbf{b}_l + \varepsilon_{\mu k}\mathbf{b}_k = \boldsymbol{\Delta}_{\mu k} + \varepsilon_{\mu k}\mathbf{b}_k$. Thus, using $y_\mu = \mathcal{F}\left(\boldsymbol{\Delta}_{\mu k} + \varepsilon_{\mu k}\mathbf{b}_k;\boldsymbol{\gamma}\right)$ we can model the likelihood such that

$$\begin{aligned}
P\left(y_\mu|\mathbf{B}\right) &= \int \mathrm{d}\boldsymbol{\Delta}_{\mu k} P\left(y_\mu,\boldsymbol{\Delta}_{\mu k}|\mathbf{B};\boldsymbol{\gamma}\right) \\
&= \int \mathrm{d}\boldsymbol{\Delta}_{\mu k} P\left(y_\mu|\boldsymbol{\Delta}_{\mu k},\mathbf{B};\boldsymbol{\gamma}\right)P\left(\boldsymbol{\Delta}_{\mu k}|\mathbf{B}\right) \\
&= \int \mathrm{d}\boldsymbol{\Delta}_{\mu k} P\left(y_\mu|\boldsymbol{\Delta}_{\mu k}+\varepsilon_{\mu k}\mathbf{b}_k;\boldsymbol{\gamma}\right)P\left(\boldsymbol{\Delta}_{\mu k}|\mathbf{B}\right) \\
&\simeq \int \mathrm{d}\boldsymbol{\Delta}_{\mu k}\left[1+\varepsilon_{\mu k}\mathbf{b}_k^{\mathsf{T}}\nabla_{\boldsymbol{\Delta}_{\mu k}}\ln P\left(y_\mu|\boldsymbol{\Delta}_{\mu k};\boldsymbol{\gamma}\right)\right]P\left(y_\mu|\boldsymbol{\Delta}_{\mu k};\boldsymbol{\gamma}\right)\,P\left(\boldsymbol{\Delta}_{\mu k}|\mathbf{B}\right) ,
\end{aligned} \tag{5}$$

where we have assumed that $P\left(y_\mu|\boldsymbol{\Delta}_{\mu k},\mathbf{B};\boldsymbol{\gamma}\right)\approx P\left(y_\mu|\boldsymbol{\Delta}_{\mu k}+\varepsilon_{\mu k}\mathbf{b}_k;\boldsymbol{\gamma}\right)$.

A. Inter-replica interaction

An explicit expression for inter-dependence between solutions is required for obtaining a closed set of update equations. We assume a dependence of the form

$$P^t\left(\mathbf{b}_k\,|\{y_{\nu\neq\mu}\}\right)\propto\exp\left\{\mathbf{h}_{\mu k}^{t\mathsf{T}}\,\mathbf{b}_k+\frac{1}{2}\mathbf{b}_k^{\mathsf{T}}\mathbf{Q}_{\mu k}^t\,\mathbf{b}_k\right\} , \tag{6}$$

where $\mathbf{h}_{\mu k}^t$ is a vector representing an external field and $\mathbf{Q}_{\mu k}^t$ the matrix of cross-replica interaction. The form of $\mathbf{Q}_{\mu k}^t$ depends upon which particular case we are considering. We assume either one of the following symmetry between replica:

$$\begin{aligned}
\left(\mathbf{h}_{\mu k}^t\right)^{\ell\mathrm{a}} &= h_{\mu k}^t \\
\left({}^{(\mathrm{RS})}\mathbf{Q}_{\mu k}^t\right)^{\mathrm{aa}'} &= \delta^{\mathrm{aa}'}\,q_{0\mu k}^t+\left(1-\delta^{\mathrm{aa}'}\right)\,q_{1\mu k}^t \\
\left({}^{(\mathrm{1RSB})}\mathbf{Q}_{\mu k}^t\right)^{\ell\mathrm{a}\,\ell'\mathrm{a}'} &= \delta^{\ell\ell'}\left({}^{(\mathrm{RS})}\mathbf{Q}_{\mu k}^t\right)^{\mathrm{aa}'}+\left(1-\delta^{\ell\ell'}\right)q_{2\mu k}^t ,
\end{aligned}$$

where ℓ is a block index that runs from 1 to L and a is a spin index that runs form 1 to n where n is the number of spins per block. We also consider that $q_{0\mu k}^t > q_{1\mu k}^t > q_{2\mu k}^t > 0$.

For both types of symmetries considered, the correlation matrix defined as:

$$\left(\boldsymbol{\Upsilon}_{\mu k}^t\right)^{\mathbf{I}\,\mathbf{I}'}\equiv\left\langle\varDelta_{\mu k}^{\mathbf{I}}\varDelta_{\mu k}^{\mathbf{I}'}\right\rangle-\left\langle\varDelta_{\mu k}^{\mathbf{I}}\right\rangle\left\langle\varDelta_{\mu k}^{\mathbf{I}'}\right\rangle$$

where $\mathbf{I}$ is an index or a pair of indexes for RS and 1RSB respectively, satisfies a self-averaging property, i.e. $\boldsymbol{\Upsilon}^t_{\mu k} \simeq \boldsymbol{\Upsilon}^t$ and preserves the symmetry of the matrix $\mathbf{Q}^t_{\mu k}$ (an explicit derivation of the entries of $\boldsymbol{\Upsilon}^t$ is presented in the Appendixes A and B):

$$\begin{aligned} \left({}^{(\mathrm{RS})}\boldsymbol{\Upsilon}^t\right)^{\mathrm{aa}'} &= \delta^{\mathrm{aa}'} X^t + \left(1-\delta^{\mathrm{aa}'}\right)\frac{1}{n} R^t \\ \left({}^{(1\mathrm{RSB})}\boldsymbol{\Upsilon}^t\right)^{\mathrm{a}\ell\,\mathrm{a}'\ell'} &= \delta^{\ell\ell'}\left[\delta^{\mathrm{aa}'} X^t + \left(1-\delta^{\mathrm{aa}'}\right)\frac{1}{n} V^t\right] + \left(1-\delta^{\ell\ell'}\right)\frac{1}{nL}\left(R^t - V^t\right) . \end{aligned}$$

The matrix $\underline{\boldsymbol{\Upsilon}}^t$ such that $\underline{\boldsymbol{\Upsilon}}^t\boldsymbol{\Upsilon}^t = \boldsymbol{\Upsilon}^t\underline{\boldsymbol{\Upsilon}}^t = \mathbb{1}$ has the same structure as $\boldsymbol{\Upsilon}^t$, $\left(\underline{\boldsymbol{\Upsilon}}^t\right)^{\mathrm{ab}} = \delta^{\mathrm{ab}}\alpha^t - \left(1-\delta^{\mathrm{ab}}\right)\frac{1}{n}\beta^t$, where $\alpha^t \simeq 1/X^t$ and $\beta^t \simeq R^t/\left[X^t\left(X^t+R^t\right)\right]$ for the RS case and $\left(\underline{\boldsymbol{\Upsilon}}^t\right)^{\mathrm{ab}} = \delta^{\mathrm{ab}}\alpha^t - \left(1-\delta^{\mathrm{ab}}\right)\frac{1}{n}\beta^t - \left(1-\delta^{\ell\ell'}\right)\frac{1}{nL}\gamma^t$, where $\alpha^t \simeq 1/X^t$, $\beta^t \simeq V^t/\left[X^t\left(X^t+V^t\right)\right]$ and $\gamma^t \simeq \left(R^t-V^t\right)/\left[\left(X^t+V^t\right)\left(X^t+R^t\right)\right]$ for the 1RSB case. Thus, the probability of $\boldsymbol{\Delta}_{\mu k}$ can be expressed as:

$$\begin{aligned} P\left(\boldsymbol{\Delta}_{\mu k}|\mathbf{B}\right) &= \sqrt{\frac{\det\left(\underline{\boldsymbol{\Upsilon}}^t\right)}{(2\pi)^n}}\exp\left\{-\frac{1}{2}\left(\boldsymbol{\Delta}_{\mu k}-\mathbf{u}^t_{\mu k}\right)^{\mathsf{T}}\underline{\boldsymbol{\Upsilon}}^t\left(\boldsymbol{\Delta}_{\mu k}-\mathbf{u}^t_{\mu k}\right)\right\} \\ &\propto \int d\vartheta\,\exp\left\{-n\frac{\left(\vartheta-u^t_{\mu k}\right)^2}{2R^t}\right\}\prod_{\mathrm{a}=1}^{n}\exp\left\{-\frac{\left(\Delta^{\mathrm{a}}_{\mu k}-\vartheta\right)^2}{2X^t}\right\} \qquad (7) \\ &\propto \int d\boldsymbol{\Theta}\,\exp\left\{-\frac{n}{2}\sum_{\ell=1}^{L}\left[\left(\vartheta^0,\vartheta^\ell\right)\mathbf{S}^t\begin{pmatrix}\vartheta^0\\ \vartheta^\ell\end{pmatrix}\right]\right\}\prod_{\ell=1}^{L}\prod_{\mathrm{a}=1}^{n}\exp\left\{-\frac{\left(\Delta^{\ell\mathrm{a}}_{\mu k}-\vartheta^{0\ell t}_{\mu k}\right)^2}{2X^t}\right\} , \qquad (8) \end{aligned}$$

where $\vartheta^{0\ell t}_{\mu k} \equiv \vartheta^0 + \vartheta^\ell + u^t_{\mu k}$, $\boldsymbol{\Theta}^{\mathsf{T}} = \left(\vartheta^0, \vartheta^1, \ldots, \vartheta^L\right)$, and the matrix $\mathbf{S}^t = \frac{1}{X^t}\begin{pmatrix}\frac{X^t\left(X^t+V^t\right)+V^t\left(X^t+R^t\right)}{X^t\left(R^t-V^t\right)} & -1\\ -1 & \frac{X^t}{V^t}\end{pmatrix}$. Expression (7) holds for the RS case while Eq. (8) holds for the 1RSB case.

The messages from nodes y_μ to nodes $\mathbf{b}_k$, in agreement with the development of the Appendix C, is represented by the expression

$$\begin{aligned} \widehat{m}^{t+1}_{\mu k} &= \varepsilon_{\mu k}\frac{\tilde{\vartheta}^t_{\mu k}-u^t_{\mu k}}{R^t} \\ &= \varepsilon_{\mu k}\frac{\tilde{\vartheta}^t_{\mu k}-u^t_{\mu k}}{R^t} + \frac{\varepsilon_{\mu k}}{2n}\frac{\mathcal{P}_2 V^t}{1-\mathcal{P}_1 V^t} , \end{aligned}$$

where $\mathcal{P}_j = \left.\frac{\partial^j\mathcal{P}}{\partial\vartheta^j}\right|_{\vartheta=\tilde{\vartheta}^t_{\mu k}}$ and $\tilde{\vartheta}^t_{\mu k}$ is given by Eq. (50) in the RS case and by Eq. (51) in the 1RSB case. The messages from nodes $\mathbf{b}_k$ to y_μ are given in any case by the expression $m^t_{\mu k} \simeq \tanh\left(\sum_{\nu\neq\mu}\widehat{m}^t_{\nu k}\right)$.

Consider the gauge field $b_k h^t_{\mu k}$ where $h^t_{\mu k} \equiv \operatorname{artanh}\left(m^t_{\mu k}\right) = \sum_{\nu\neq\mu}\operatorname{artanh}\left(\hat{m}^t_{\nu k}\right) \simeq \sum_{\nu\neq\mu}\hat{m}^t_{\nu k}$. This distribution of this field is likely to be well approximated by a Gaussian as a result of the central limit theorem. The mean and variance of the Gaussian are E^t and F^t respectively:

$$\begin{aligned} E^t &= \frac{1}{K}\sum_{k=1}^{K}\sum_{\mu=1}^{N} b_k\hat{m}^t_{\mu k} \\ F^t &= \sum_{\mu=1}^{N}\left[\frac{1}{K}\sum_{k=1}^{K}\left(b_k\hat{m}^t_{\mu k}\right)^2 - \left(\frac{1}{K}\sum_{k=1}^{K} b_k\hat{m}^t_{\mu k}\right)^2\right] \simeq \frac{1}{K}\sum_{k=1}^{K}\sum_{\mu=1}^{N}\left(\hat{m}^t_{\mu k}\right)^2 . \end{aligned}$$

We assume that this quantities are independent of the index μ by virtue of the self-averaging property. For the same reason we expect that the macroscopic variables defined as $M_\mu^t \equiv \sum_{k=1}^K b_k m_{\mu k}^t / K \simeq \sum_{k=1}^K b_k m_k^t / K = M^t$ and $N_\mu^t \equiv \sum_{k=1}^K \left(m_{\mu k}^t\right)^2 / K \simeq \sum_{k=1}^K \left(m_k^t\right)^2 / K = N^t$, where $m_k^t \simeq \tanh\left(\sum_{\nu=1}^N \widehat{m}_{\nu k}^t\right)$, are also independent of the index μ. Thus, these macroscopic variables can be evaluated as:

$$M^t = \int \mathcal{D}u \tanh\left(\sqrt{F^t}u + E^t\right) \quad N^t = \int \mathcal{D}u \tanh^2\left(\sqrt{F^t}u + E^t\right) ,$$

where $\mathcal{D}u = \exp\left(-u^2/2\right)/\sqrt{2\pi}$. Since the MPM estimator is given by $\hat{b}_k^t = \mathrm{sgn}\left(m_k^t\right) \simeq \mathrm{sgn}\left(m_{\mu k}^t\right) = \mathrm{sgn}\left(h_{\mu k}^t\right)$, the expression for the error per bit rate is:

$$P_b^t = \frac{1}{2K} \sum_{k=1}^K \left(1 - \mathrm{sgn}\left(b_k m_k^t\right)\right) , \tag{9}$$

which is minimised if the vector of true bits $\mathbf{b}$ and the vector of messages $\mathbf{m}^t$ are parallel. Therefore, the error per bit rate becomes smaller if the quantity $M^t/\sqrt{N^t} = \cos\left(\widehat{\mathbf{b}\,\mathbf{m}^t}\right)$ approaches 1. The optimisation is reached when $E^t\left(\boldsymbol{\gamma}^c\right) = F^t\left(\boldsymbol{\gamma}^c\right)$ and $\left.\frac{\partial E^t}{\partial \gamma_i} - \frac{1}{2}\frac{E^t}{F^t}\frac{\partial F^t}{\partial \gamma_i}\right|_{\gamma_i^c} = 0$ (the demonstration is presented in Appendix E).

V. CDMA AND LINEAR ISING PERCEPTRON

According to previous notation, it holds that $\varepsilon_{\mu k} = s_{\mu k}/\sqrt{N}$ for the CDMA problem and $\varepsilon_{\mu k} = s_{\mu k}/\sqrt{K}$ for the Ising perceptron. The goal is to get an accurate estimate of the vector $\mathbf{b}$ for all users given the received message vector $\mathbf{y}$ by approximating the posterior $P(\mathbf{b}|\mathbf{y})$. An expression representing the likelihood is required and is easily derived from the noise model (assuming zero mean and variance σ^2). If the arithmetic variance over replicas of the macroscopic message $\Delta_{\mu k}^{\mathrm{a}}$ is finite and independent of the sub indexes μ and k, i.e. $\varSigma^2 \equiv \frac{1}{n}\sum_{\mathrm{a}}\left(\Delta_{\mu k}^{\mathrm{a}}\right)^2 - \left(\frac{1}{n}\sum_{\mathrm{a}}\Delta_{\mu k}^{\mathrm{a}}\right)^2 < \infty \,\forall\, \mu k$, then:

$$P\left(y_\mu|\mathbf{B}\right) \simeq \sqrt{\frac{n}{2\pi\sigma^2}}\mathrm{e}^{\frac{\varSigma^2}{2\sigma^2}} \exp\left\{-\frac{\left(\mathbf{y}_\mu - \boldsymbol{\Delta}_{\mu k}\right)^{\mathsf{T}}\left(\mathbf{y}_\mu - \boldsymbol{\Delta}_{\mu k}\right)}{2\sigma^2}\right\}\left[1 + \frac{\varepsilon_{\mu k}}{\sigma^2}\mathbf{b}_k^{\mathsf{T}}\left(\mathbf{y}_\mu - \boldsymbol{\Delta}_{\mu k}\right)\right] , \tag{10}$$

where $\mathbf{y}_\mu = y_\mu \mathbf{u}$ and $\mathbf{u}^{\mathsf{T}} \equiv \overbrace{(1, 1, \cdots, 1)}^{nL}$. The function $\mathcal{P}\left(\vartheta, y_\mu\right)$ Eq. (44) obtained from this distribution is linear in ϑ, therefore $\mathcal{P}_2 = 0$. This implies that the system is replica symmetric.

To calculate correlations between replica we expand $P\left(y_\mu|\mathbf{B}\right)$ in the large N limit Eq. (10), as shown in Eq. (5). According to this model, the macroscopic variables satisfy the following relationships:

$$\begin{aligned} u_{\mu k}^t &= \frac{1}{\sqrt{\mathrm{e}_1 N}} \sum_{l \neq k} s_{\mu l} m_{\mu l}^t \\ X^t &\simeq \mathrm{e}_2\left(1 - N^t\right) , \end{aligned}$$

where $\mathrm{e}_1 = 1\,(\beta)$ for the CDMA (ILP) system and $\mathrm{e}_2 = \beta\,(1)$ for the CDMA (ILP) system. The ground state of the Hamiltonian is:

$$\tilde{\vartheta} = \frac{\sigma^2 + X^t}{\sigma^2 + X^t + R^t} u^t_{\mu k} + \frac{R^t}{\sigma^2 + X^t + R^t} y_\mu \,.$$

The message from y_μ to b^{a}_k at time $t+1$ is then given by:

$$\widehat{m}^{t+1}_{\mu k} = \varepsilon_{\mu k} \frac{y_\mu - u^t_{\mu k}}{\sigma^2 + X^t + R^t} \,. \tag{11}$$

The main difference between Eq. (11) and the equivalent equation in [2] is the dependency of the pre-factor on R^t, reflecting correlations between different solutions groups (replica). To determine this term we optimise the choice of σ^2 by applying the condition $E^t = F^t$. The imposition of this condition should lead us to a relationship between the structure of the space of solutions, represented by R^t, and the free parameter of the model σ^2. From Eq. (11) and using that $E^t = F^t \Rightarrow M^t = N^t$ we obtain:

$$E^{t+1} = \frac{\mathrm{e}_1^{-1}}{\sigma^2 + X^t + R^t} \qquad F^{t+1} = \mathrm{e}_1 \left[\sigma_0^2 + X^t\right] \left(E^{t+1}\right)^2 ,$$

which implies after a simplification that, for both cases, $R^t = \sigma_0^2 - \sigma^2$. Despite the simplicity of this result, the process from which we obtained it gives us a hint for the right way to estimate the true noise variance. For the calculation of E^t and F^t we had taken the limit $K \to \infty\,,\ N \to \infty$ with $K/N = \beta$. So the σ_0^2 that appears in the expression for F^t has been obtain from a vector of signal with an infinite number of entries y_μ. Thus

$$\lim_{N\to\infty} \frac{1}{N} \sum_{\mu=1}^{N} (y_\mu)^2 = \mathrm{e}_2 + \sigma_0^2 \,.$$

Using this expression we can finally express the message as:

$$\widehat{m}^{t+1}_{\mu k} \simeq \varepsilon_{\mu k} \frac{y_\mu - u^t_{\mu k}}{\frac{1}{N}\sum_{\mu=1}^{N} (y_\mu)^2 - \mathrm{e}_2 N^t} \,, \tag{12}$$

where no estimate on σ_0 is required.

To illustrate how the method works in a system with broken replica symetry, we will consider a CDMA signal Eq. (1) where the noise n_μ is drawn from a bi-Gaussian distribution:

$$P\left(n_\mu\right) = \frac{1-r_0}{2} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{\left(n_\mu + \varepsilon_0/\sigma_0\right)^2}{2}\right\} + \frac{1+r_0}{2} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{\left(n_\mu - \varepsilon_0/\sigma_0\right)^2}{2}\right\} , \tag{13}$$

where $r_0 \in (-1,1)$ and $\pm\varepsilon_0/\sigma_0$ are the positions of the Gaussian peaks. We also consider that $|\varepsilon_0/\sigma_0| \ll 1$. For this model the expression of the likelyhood is:

$$P\left(y_\mu | \varDelta_\mu; r, \varepsilon, \sigma^2\right) \propto \prod_{\ell=1}^{L} \prod_{\mathrm{a}=1}^{n} \left\{ \frac{1-r}{2} \exp\left[-\frac{\left(y_\mu - \varDelta^{\ell \mathrm{a}}_\mu + \varepsilon\right)^2}{2\sigma^2}\right] + \frac{1+r}{2} \exp\left[-\frac{\left(y_\mu - \varDelta^{\ell \mathrm{a}}_\mu - \varepsilon\right)^2}{2\sigma^2}\right] \right\} ,$$

where r, ε and σ^2 are estimates of the true parameters r_0, ε_0 and σ_0^2.

The function $\mathcal{P}(\vartheta, y_\mu)$ Eq. (44) has the form:

$$\mathcal{P}(\vartheta, y_\mu) = \frac{y_\mu - \vartheta}{\sigma^2 + X^t} - \frac{\varepsilon}{\sigma^2 + X^t} \tanh\left(\varepsilon \frac{y_\mu - \vartheta}{\sigma^2 + X^t} + \operatorname{arctanh}(r)\right),$$

where $X^t = \beta (1 - N^t)$.

The Eq. (50) can be expressed as:

$$\begin{aligned}
\tilde{\vartheta}^t_{\mu k} &= u^t_{\mu k} + R^t \mathcal{P}\left(\tilde{\vartheta}^t_{\mu k}, y_\mu\right) \\
y_\mu - \tilde{\vartheta}^t_{\mu k} &= y_\mu - u^t_{\mu k} - R^t \frac{y_\mu - \tilde{\vartheta}^t_{\mu k}}{\sigma^2 + X^t} + \varepsilon \frac{R^t}{\sigma^2 + X^t} \tanh\left(\varepsilon \frac{y_\mu - \tilde{\vartheta}^t_{\mu k}}{\sigma^2 + X^t} + \operatorname{arctanh}(r)\right) \\
\frac{y_\mu - \tilde{\vartheta}^t_{\mu k}}{\sigma^2 + X^t} &= \frac{y_\mu - u^t_{\mu k}}{\sigma^2 + X^t + R^t} + \frac{\varepsilon}{\sigma^2 + X^t} \frac{R^t}{\sigma^2 + X^t + R^t} \tanh\left(\varepsilon \frac{y_\mu - \tilde{\vartheta}^t_{\mu k}}{\sigma^2 + X^t} + \operatorname{arctanh}(r)\right) \\
z &= \rho_R \left(y_\mu - u^t_{\mu k}\right) + \varepsilon (\rho_0 - \rho_R) \tanh(\varepsilon z + \operatorname{arctanh(r)}) \\
&\simeq z_0 + r\, \Delta\rho_R\, \varepsilon + \left(1 - r^2\right) \Delta\rho_R\, z\, \varepsilon^2 - r\left(1 - r^2\right) \Delta\rho_R\, z^2 \varepsilon^3 - \frac{1}{3}\left(1 - r^2\right)\left(1 - 3r^2\right) \Delta\rho_R\, z^3 \varepsilon^4,
\end{aligned}$$

where $z \equiv \frac{y_\mu - \tilde{\vartheta}^t_{\mu k}}{\sigma^2 + X^t}$, $\rho_A \equiv (\sigma^2 + X^t + A)^{-1}$, $z_0 \equiv \rho_R \left(y_\mu - u^t_{\mu k}\right)$ and $\Delta\rho_R \equiv \rho_0 - \rho_R$. The solution of this equation, up to order $\mathcal{O}(\varepsilon^4)$, is:

$$z(\varepsilon) \simeq z_0 + r\, \Delta\rho_R\, \varepsilon + \left(1 - r^2\right) \Delta\rho_R \left[z_0 \varepsilon^2 + r\, \left(\Delta\rho_R - z_0^2\right) \varepsilon^3 + \left(1 - 3r^2\right) z_0 \left(\Delta\rho_R - \frac{1}{3} z_0^2\right) \varepsilon^4\right].$$

The function $\mathcal{P}$ and its two first derivatives at the ground state are:

$$\begin{aligned}
\mathcal{P}_0 &= -r\left[1 + \left(1 - r^2\right) \Delta\rho_R\, \varepsilon^2\right] \rho_R\, \varepsilon + \\
&\quad + \left[1 - \left(1 - r^2\right) \rho_R^2\, \varepsilon^2 - \left(1 - r^2\right)\left(1 - 3r^2\right) \Delta\rho_R\, \rho_R^2\, \varepsilon^4\right] \left(y_\mu - u^t_{\mu k}\right) + \\
&\quad + r\left(1 - r^2\right) \rho_R^3 \left(y_\mu - u^t_{\mu k}\right)^2 \varepsilon^3 + \frac{1}{3}\left(1 - r^2\right)\left(1 - 3r^2\right) \rho_R^4 \left(y_\mu - u^t_{\mu k}\right)^3 \varepsilon^4 \\
\mathcal{P}_1 &\simeq -\rho_0 + \mathcal{O}\left(\varepsilon^2\right) \\
\mathcal{P}_2 &= 2\rho_0^3 \left(1 - r^2\right) \left[r \varepsilon^3 + \left(1 - 3r^2\right) \rho_R \left(y_\mu - u^t_{\mu k}\right) \varepsilon^4\right],
\end{aligned}$$

therefore

$$\frac{1}{2} \frac{\mathcal{P}_2 V^t}{1 - \mathcal{P}_1 V^t} = \left(1 - r^2\right) \rho_0 \Delta\rho_V \left[r\, \varepsilon^3 + \left(1 - 3r^2\right) \rho_R \left(y_\mu - u^t_{\mu k}\right) \varepsilon^4\right],$$

where $\Delta\rho_V \equiv \rho_0 - \rho_V$, which leads to the following expression for the message:

$$\begin{aligned}
{}^{(1\mathrm{RSB})}\widehat{m}^{t+1}_{\mu k} &= \frac{s_{\mu k}}{\sqrt{N}} \left\{- \left[\rho_R + \left(1 - r^2\right)\left(\Upsilon_n - \rho_R^2\right) \varepsilon^2\right] r\, \varepsilon + \right. \\
&\quad + \rho_R \left[1 - \left(1 - r^2\right) \rho_R\, \varepsilon^2 - \left(1 - r^2\right)\left(1 - 3r^2\right)\left(\Upsilon_n - \rho_R^2\right) \varepsilon^4\right] \left(y_\mu - u^t_{\mu k}\right) + \\
&\quad \left. + r\left(1 - r^2\right) \rho_R^3\, \varepsilon^3 \left(y_\mu - u^t_{\mu k}\right)^2 + \frac{1}{3}\left(1 - r^2\right)\left(1 - 3r^2\right) \rho_R^4\, \varepsilon^4 \left(y_\mu - u^t_{\mu k}\right)^3 \right\}, \qquad (14)
\end{aligned}$$

where $\Upsilon_n \equiv \rho_0 \left(\rho_R - \frac{1}{n} \Delta\rho_V\right)$. The expresion for the message in the RS case is recovered from Eq. (14) in the limit $n \to \infty$.

For calculatino the expressions for the macroscopic variables E^{t+1} and F^{t+1} we have to perform the following sums, in the limit of $K, N \to \infty$ with $K/N = \beta < \infty$:

$$
\begin{aligned}
A_j &\equiv \lim_{K,N\to\infty} \sum_{\mu}^{N} \frac{1}{K} \sum_{k=1}^{K} \frac{s_{\mu k} b_k}{\sqrt{N}} \left(y_\mu - u_{\mu k}^t\right)^j \\
B_l &\equiv \lim_{K,N\to\infty} \frac{1}{N} \sum_{\mu}^{N} \frac{1}{K} \sum_{k=1}^{K} \left(y_\mu - u_{\mu k}^t\right)^l ,
\end{aligned}
$$

where $j = 0, \ldots, 3$ and $l = 0, \ldots, 4$. From the definition of the signal y_μ and the expression for the noise Eq. (13) we found that $A_0 = 0$, $A_1 = 1$, $A_2 = 2B_1$, $A_3 = 3B_2$, $B_0 = 1$, $B_1 = r_0 \varepsilon_0$, $B_2 = \beta\left(1 - 2M^t + N^t\right) + \sigma_0^2 + \varepsilon_0^2$, $B_3 = B_1\left(3B_2 - 2\varepsilon_0^2\right)$ and $B_4 = 3B_2^2 - 2\varepsilon_0^4$.

The expressions for the macroscopic variables are:

$$
\begin{aligned}
E^{t+1} &= \rho_R - \left(1 - r^2\right)\rho_R^2\,\varepsilon^2 + 2r\left(1 - r^2\right) B_1 \rho_R^3\,\varepsilon^3 - \left(1 - r^2\right)\left(1 - 3r^2\right)\left[\Upsilon_n - \left(1 + B_2\rho_r\right)\rho_R^2\right]\rho_R\,\varepsilon^4 \\
F^{t+1} &= B_2\rho_R^2 - 2rB_1\rho_R^2\,\varepsilon + \left[r^2 - 2\left(1 - r^2\right)B_2\rho_R\right]\rho_R^2\,\varepsilon^2 - 2r\left(1 - r^2\right)B_1\left[\Upsilon_n - \left(2 + 3B_2\rho_R\right)\rho_R^2\right]\rho_R\,\varepsilon^3 + \\
&\quad + \left(1 - r^2\right)\left[2r^2\left(\Upsilon_n - \rho_R^2\right)\rho_R + \left(1 - 3r^2\right)B_2\left(3\rho_r^2 + 2B_2\rho_R^3 - 2\Upsilon_n\right)\rho_R^2\right]\varepsilon^4 .
\end{aligned}
$$

Applying the conditions $E^t\left(\boldsymbol{\gamma}^c\right) = F^t\left(\boldsymbol{\gamma}^c\right)$ and $\left.\frac{\partial E^t}{\partial \gamma_i} - \frac{1}{2}\frac{E^t}{F^t}\frac{\partial F^t}{\partial \gamma_i}\right|_{\gamma_i^c} = 0$, where $\boldsymbol{\gamma}^{\mathsf{T}} = \left(r, \varepsilon, \sigma^2, \frac{1}{n}\right)$ we found the following conditions:

$$
\rho_R = \frac{1}{B_2} + \frac{\varepsilon^2}{B_2^2} - \frac{\varepsilon^4}{B_2^3} + \left(1 - r^2\right)^2 \frac{1 - B_2\rho_0\left(1 - \frac{1}{n}B_2\Delta\rho_V\right)}{B_2^3}\,\varepsilon^4 \tag{15}
$$

$$
r\,\varepsilon = B_1 + r\left(1 - r^2\right)\frac{1 - B_2\rho_0\left(1 - \frac{1}{n}B_2\Delta\rho_V\right)}{B_2}\,\varepsilon^3 . \tag{16}
$$

In the 1RSB case we can further simplify this expressions by a suitable choice of V^t and the number of replicas per block n. If these choices are such that:

$$
1 = B_2\rho_0\left(1 - \frac{1}{n}B_2\Delta\rho_V\right), \tag{17}
$$

this condition implies that:

$$
V^t = \frac{\left(X^t + \sigma^2\right)^2\left(\sigma_0^2 - \sigma^2\right)}{\frac{1}{n}\left(X^t + \sigma_0^2\right)^2 - \left(X^t + \sigma^2\right)\left(\sigma_0^2 - \sigma^2\right)} ,
$$

which by definition is larger than zero. This condition is satisfied if our estimate for the noise variance is smaller than the true parameter ($\sigma^2 < \sigma_0^2$). In this case the number of replicas per block has to satisfy the condition:

$$
1 \le n \le f\left(X^t; \sigma_0^2, \sigma^2\right) \equiv \frac{\left(X^t + \sigma_0^2\right)^2}{\left(X^t + \sigma^2\right)\left(\sigma_0^2 - \sigma^2\right)} .
$$

For $0 \le X^t$, the minimum value of $f\left(X^t; \sigma_0^2, \sigma^2\right)$ is reached at $X_{min} = \max\left(0, \sigma_0^2 - 2\sigma^2\right)$. It is also possible to proof that $4 \le f\left(X_{min}; \sigma_0^2, \sigma^2\right)$. Although V^t and n will not be explicitely used in the following expressions, the correct choice of the value for these parameters allows us to use Eqs. (15)

and (16) to find the final expression for the macroscopic variable E^{t+1}, where no estimates are needed for the noise parameters:

$$^{(1\mathrm{RSB})}E^{t+1} = \frac{1}{B_2 - B_1^2}.$$

For the RS case we do not have the freedom to chose a number of replicas per block, given that this case is equivalent to take $n \to \infty$. For this reason Eqs. (15) and (16) take the form:

$$\rho_R = \frac{1}{B_2} + \frac{\varepsilon^2}{B_2^2} - \frac{\varepsilon^4}{B_2^3} + \left(1-r^2\right)^2 \frac{1-B_2\rho_0}{B_2^3}\,\varepsilon^4 \tag{18}$$

$$r\,\varepsilon = B_1 + r\left(1-r^2\right)\frac{1-B_2\rho_0}{B_2}\,\varepsilon^3\,, \tag{19}$$

and thus the macroscopic variable is:

$$^{(\mathrm{RS})}E^{t+1} = {}^{(1\mathrm{RSB})}E^{t+1} + \frac{2B_1^2\left(\varepsilon^2 - B_1^2\right)}{B_2^3}\left(\frac{B_2}{X^t+\sigma^2} - 1\right),$$

which depends in both, the estimate for the noise variance σ^2 and the noise bias ε.

Given that the algorithm will deal with finite signal vectors ($N < \infty$), the quantities $B_1$and B_2 have to be approximated by the correspondent finite sums. Therefore, we have:

$$B_1 = \lim_{N,K\to\infty}\frac{1}{N}\sum_{\mu=1}^{N}\frac{1}{K}\sum_{k=1}^{K}\left(y_\mu - u_{\mu k}^t\right) \approx \frac{1}{N}\sum_{\mu=1}^{N} y_\mu \equiv \overline{B}_1$$

$$B_2 = \lim_{N,K\to\infty}\frac{1}{N}\sum_{\mu=1}^{N}\frac{1}{K}\sum_{k=1}^{K}\left(y_\mu - u_{\mu k}^t\right)^2 \approx \frac{1}{N}\sum_{\mu=1}^{N} y_\mu^2 + \beta N^t \equiv \overline{B}_2\,,$$

where we have used that $\lim_{N,K\to\infty}\frac{1}{NK}\sum_{\mu,k}u_{\mu k}^t = 0$. Observe that no information about the true noise has been used to derive these expressions. With these we can write down the messages as:

$$\begin{aligned}
^{(1\mathrm{RSB})}\widehat{m}_{\mu k}^{t+1} &= \frac{s_{\mu k}}{\sqrt{N}}\left\{-\frac{\overline{B}_1}{\overline{B}_2} + \frac{\overline{B}_1}{\overline{B}_2^2}\varepsilon^2 + \left(\frac{1}{\overline{B}_2} + \frac{\overline{B}_1^2}{\overline{B}_2^2} - \frac{3\varepsilon^2 - 2\overline{B}_1^2}{\overline{B}_2^3}\varepsilon^2\right)\left(y_\mu - u_{\mu k}^t\right) + \right.\\
&\left. + \frac{\overline{B}_1\left(\varepsilon^2 - \overline{B}_1^2\right)}{\overline{B}_2^3}\left(y_\mu - u_{\mu k}^t\right)^2 + \frac{1}{3}\frac{\left(\varepsilon^2 - \overline{B}_1^2\right)\left(\varepsilon^2 - 3\overline{B}_1^2\right)}{\overline{B}_2^4}\left(y_\mu - u_{\mu k}^t\right)^3\right\}\\
^{(\mathrm{RS})}\widehat{m}_{\mu k}^{t+1} &= {}^{(1\mathrm{RSB})}\widehat{m}_{\mu k}^{t+1} + \frac{s_{\mu k}}{\sqrt{N}}\left(1 - \frac{\overline{B}_2}{X^t+\sigma^2}\right)\frac{\varepsilon^2 - \overline{B}_1^2}{\overline{B}_2^2}\left[\overline{B}_1 + 2\frac{\varepsilon^2 - 2\overline{B}_1^2}{\overline{B}_2^2}\left(y_\mu - u_{\mu k}^t\right)\right].
\end{aligned}$$

A. Steady state and critical analysis

The steady state equations for the macroscopic variables N^t and E^t are obtained by taken the limit $t \to \infty$. Let us define $\overline{N} \equiv \lim_{t\to\infty} N^t$ and $\overline{E} \equiv \lim_{t\to\infty} E^t$. In this asymptotic regime the following relationships hold:

$$\overline{N}\left(\sigma_0^2,\beta\right) = \int \mathcal{D}u\,\tanh^2\left(\sqrt{\overline{E}\left(\sigma_0^2,\beta\right)}u + \overline{E}\left(\sigma_0^2,\beta\right)\right) \qquad \overline{E}\left(\sigma_0^2,\beta\right) = \frac{\mathrm{e}_1^{-1}}{\sigma_0^2 + \mathrm{e}_2\left(1 - \overline{N}\left(\sigma_0^2,\beta\right)\right)} \tag{20}$$

and from thees expressions we can calculate the error per bit rate:

$$\overline{P}_b\left(\sigma_0^2,\beta\right) = \frac{1}{2}\left[1 + \mathrm{erf}\left(\sqrt{\frac{\overline{E}\left(\sigma_0^2,\beta\right)}{2}}\right)\right]. \tag{21}$$

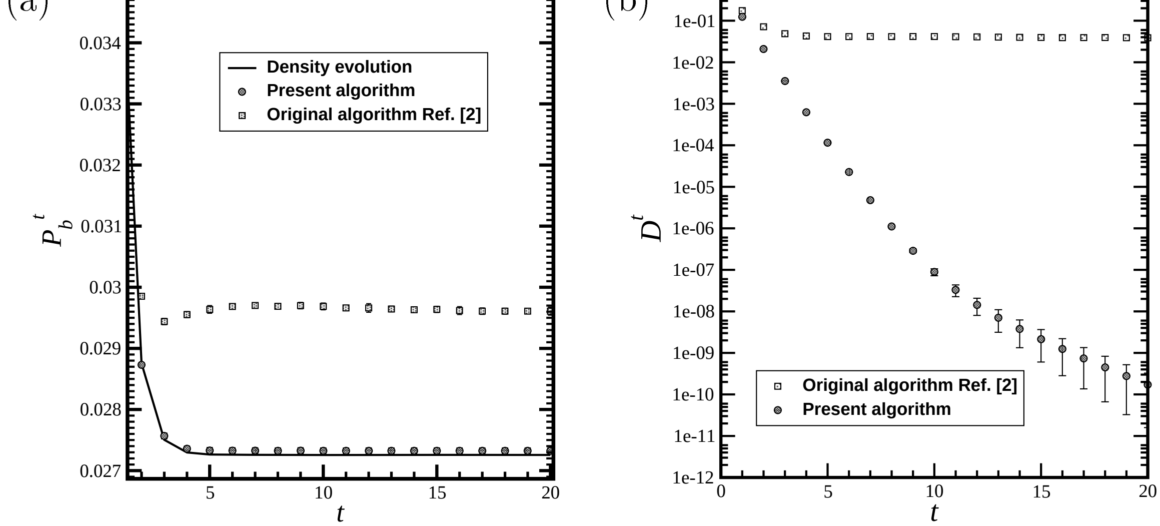

Figure 3: (a) Error probability of the inferred solution evolving in time. The system load $\beta = 0.25$, true noise level $\sigma_0^2 = 0.25$ and estimated noise $\sigma^2 = 0.01$. Squares represent results of the original algorithm [2], solid line the dynamics obtained from our equations; circles represent results obtained from the suggested practical algorithm. Variances are smaller than the symbol size. (b) D^t, a measure of convergence in the obtained solutions, as a function of time; symbols are as in the main figure.

B. Results

The inference algorithm requires an iterative update of Eqs.(49,12) and converges to a reliable estimate of the signal, with no need for an accurate prior information of the noise level. The computational complexity of the algorithm is of $\mathcal{O}(K^2)$.

To test the performance of our algorithm we carried out a set of experiments of CDMA signal detection problem under typical conditions. Error probability of the inferred signals has been calculated for a system load of $\beta = 0.25$, where the true noise level is $\sigma_0^2 = 0.25$ and the estimated noise is $\sigma^2 = 0.01$, as shown in Figure 3 (a). The solid line represents the expected theoretical results (density evolution), knowing the exact values of σ_0^2 and σ^2, while circles represent simulation results obtained via the suggested *practical* algorithm, where no such knowledge is assumed. The results presented are based on 10^5 trials per point and a system size $N = 2000$ and are superior to those obtained using the original algorithm [2].

Another performance measure one should consider is

$$D^t \equiv \frac{1}{K} \left(\mathbf{m}^t - \mathbf{m}^{t-1}\right) \cdot \left(\mathbf{m}^t - \mathbf{m}^{t-1}\right) ,$$

that provides an indication to the stability of the solutions obtained. In Figure 3 (b) we see that results obtained from our algorithm show convergence to a reliable solution in stark contrast to the original algorithm [2]. The physical interpretation of the difference between the two results is assumed to be related to a replica symmetry breaking phenomenon.

For the ILP, the $K > N$ regime is highly interesting as the system develops a critical behaviour for a range of σ_0^2 values. We carried out a set of experiments for this system (the CDMA scaling was kept for consistency) based on density evolution. In Figure 4 (a) we present curves of $\overline{P}_b$, defined in Eq. (21), as a function of the inverse load β^{-1} for different values of σ_0^2. Three different regimes have been observed: For $\sigma_0^2 < 0.15$ the curves exhibit a discontinuity at a value of β that varies with σ_0^2 (first order phase transition-like behaviour). At $\sigma_0^2 = 0.15$ the curve becomes continuous but its slope diverges (second order phase transition-like behaviour). The $\overline{P}_b$ curves show analytical behaviour for noise values above 0.15. In Figure 4 (b) we present a phase diagram of the CDMA system. It shows the dependency of the critical load β_C^{-1} as a function of the noise parameter. The first order line ends in a second order transition point marked by a circle.

Another indication for the critical behaviour is the number of steps required for the recursive update of Eq. 20 to convergence. In Figure 5 (a) we present the number of iterations needed to reach a steady state as a function of β^{-1} when the noise parameter is set to $\sigma_0^2 = 0.1$. The number of iterations diverge when the critical value of β is reached.

Finally, we wish to explore the efficiency of the algorithm as a function of the system size. In Figure 5 (b) we present the result of iterating Eqs. (49) and (12) for system sizes of K=200, 400, 800, 1600 and 3200. The curves represent mean values over 1000 experiments. There is a strong dependency of the error per bit rate on the size of the system, which is expected to converge to the asymptotic limit (infinite system size) represented by the solid line.

To test the performance of the 1RSB algorithm we carried out a set of experiments of CDMA signal detection problem with bi-Gaussian noise. Error probability of the inferred signals has been calculated for a system load of β=0.25, where the true noise level is σ_0^2=0.25, bias of $\varepsilon_0 = 0.06$ and weight r_0 =0.6. The estimated noise parameters are σ^2=0.01 and $\varepsilon = 0.2$, as shown in Figure 6 (a). The circles represent simulation results obtained via the 1RSB algorithm. The results presented are based on 10^5 trials per point and a system size $N = 1000$ and are superior to those obtained using the RS algorithm. In Figure 6 (b) we see that results obtained from the 1RSB algorithm and from the RS algorithm show convergence to a reliable solution.

VI. CDMA SIGNAL DETECTION WITH DUAL-PEAKED GAUSSIAN NOISE

VII. CONCLUSIONS

In summary, we present a new algorithm for using belief propagation in densely connected systems that enables one to obtain reliable solutions even when the solution space is fragmented. It represents an extension to existing algorithms of that type which is reminiscent to the extension of BP to SP. The algorithm has been tested on the signal detection problem in CDMA and has provided superior results to other existing algorithms [2, 17]. Further research is required to fully determine the

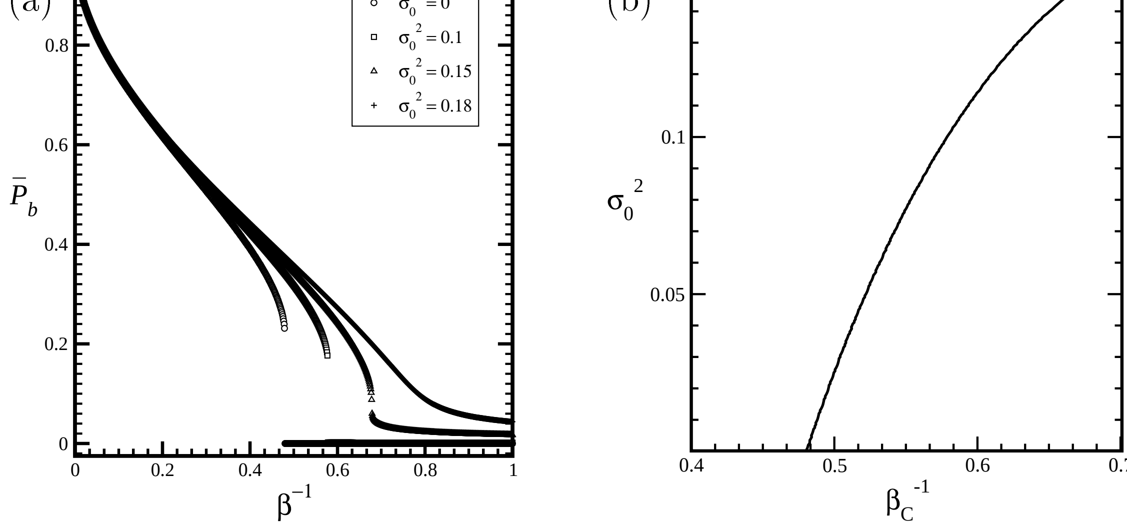

Figure 4: (a) $\overline{P}_b$ at the steady state, Eq. (21), as a function of β^{-1} for different values of the noise parameter. For values of σ_0^2 below 0.15 the curves show discontinuity at certain β values, which becomes continuous but non-analytic at $\sigma_0^2 = 0.15$ around $\beta^{-1} \simeq 0.68$. For noise variance values above $\sigma_0^2 = 0.15$ the curves become analytical. (b) Position of the non analyticity of the error rate curve β_C^{-1} as a function of the noise parameter σ_0^2. This first order phase transition-like curve ends in a second order phase transition-like point marked by $\circ$.

potential of the new algorithm.

Acknowledgments

Support from the EU FP-6 EVERGROW IP is gratefully acknowledged.

Appendix A: The Replica Symmetric (RS) Ansatz

With this setting, the interaction term in Eq. (6) is now:

$$\mathbf{b}_k^{\mathsf{T}} \mathbf{Q}_{\mu k}^t \mathbf{b}_k = n \left(q_{0\mu k}^t - q_{1\mu k}^t \right) + q_{1\mu k}^t \left(\sum_{\mathrm{a}=1}^{n} b_k^{\mathrm{a}} \right)^2 ,$$

An expression for equation (6) immediately follows

$$\begin{aligned} P^t \left(\mathbf{b}_k | \left\{ y_{\nu \neq \mu} \right\} \right) &= [\mathcal{Z}_{\mu k}^t]^{-1} \exp \left\{ h_{\mu k}^t \sum_{\mathrm{a}=1}^{n} b_k^{\mathrm{a}} + \frac{1}{2} q_{1\mu k}^t \left(\sum_{\mathrm{a}=1}^{n} b_k^{\mathrm{a}} \right)^2 \right\} \\ &= [\mathcal{Z}_{\mu k}^t]^{-1} \int_{-\infty}^{\infty} \mathrm{d}x \, \exp \left\{ -\frac{x^2}{2 q_{1\mu k}^t} + \left(x + h_{\mu k}^t \right) \sum_{\mathrm{a}=1}^{n} b_k^{\mathrm{a}} \right\} \end{aligned}$$

where $\mathcal{Z}_{\mu k}^t$ is a normalisation constant. The diagonal elements $q_{0\mu k}^t$ fix the zero of the free energy, so we can take them equal to zero without loss of generality.

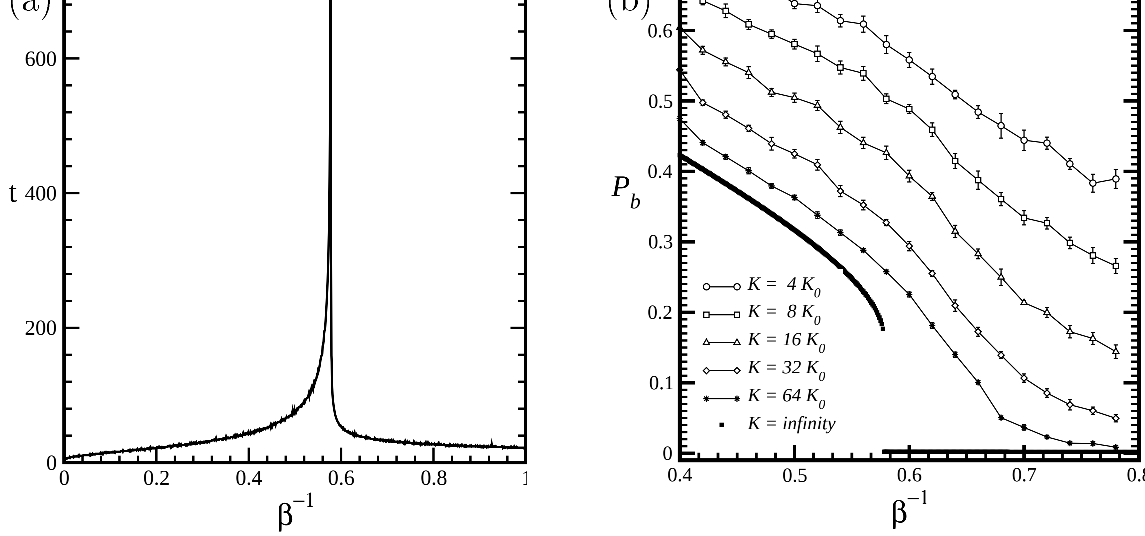

Figure 5: (a) Number of iterations of Eq. (20) required for convergence as a function of β, for $\sigma_0^2 = 0.10$. the error rate curve exhibits a discontinuity. (b) Finite size effects observed in the error rate curve when the Eqs. (49) and (12) are iterated over the number of steps needed to reach the steady state. The noise level used is $\sigma_0^2 = 0.10$ with $K_0 = 50$. The curves are mean values over 1000 experiments. The curve obtained from the iteration of the steady state equations is presented as a reference.

We expect the free energy obtained from the well behaved distribution P^t to be self-averaging, thus

$$\lim_{n\to\infty} \frac{1}{n} \log\left(\overline{\mathcal{Z}_{\mu k}^t}\right) = \lim_{n\to\infty} \frac{1}{n} \log\left(\mathcal{Z}_{\mu k}^t\left(\hat{h}, \hat{q}_1\right)\right),$$

where $\hat{h}$ and $\hat{q}_1$ are the mean value of the parameters when extracted for some suitable distributions and the over-line represents the mean value of the partition function over such distributions.

In the following calculation we will drop the upper-index t and the sub-indices μ and k to simplify the notation. To obtain the scaling behaviour of the various parameters we calculate $\mathcal{Z}\left(h, q_1\right)$ explicitly, assuming the parameter q_1 is taken from a normal distribution $\mathcal{N}\left(\hat{q}_1, \sigma_q^2\right)$. The partition function takes the form :

$$\mathcal{Z}\left(h, q_1\right) = \int_{-\infty}^{\infty} \frac{\mathrm{d}x}{\sqrt{2\pi q_1}} \exp\left(-\frac{(x-h)^2}{2q_1} + n\ln\left(2\cosh(x)\right)\right). \tag{22}$$

Thus, the mean value of the partition function over the set of parameters is:

$$\overline{\mathcal{Z}\left(h, q_1\right)} = \int \mathcal{D}_{q_1}\, \mathcal{Z}\left(h, q_1\right),$$

where $\mathcal{D}_{q_1} = \mathrm{d}q_1 \mathcal{N}\left(\hat{q}_1, \sigma_{q_1}^2\right)$. The normalisation can be expressed as:

$$\begin{aligned}
\overline{\mathcal{Z}\left(h, q_1\right)} &= \sum_{\mathrm{a}=0}^{n} \binom{n}{\mathrm{a}} \exp\left\{n\left[h\left(1-\frac{2\mathrm{a}}{n}\right) + \frac{\hat{q}_1}{2}\left(1-\frac{2\mathrm{a}}{n}\right)^2 n + \frac{\sigma_{q_1}^2}{8}\left(1-\frac{2\mathrm{a}}{n}\right)^4 n^3\right]\right\} \\
&= \mathcal{A}(n)\,(n+1)\binom{n}{n/2} \exp\left\{n\left[|h| + n\frac{\hat{q}_1}{2} + n^3\frac{\sigma_{q_1}^2}{8}\right]\right\} \\
&\simeq \sqrt{\frac{2}{\pi}}\mathcal{A}(n) \exp\left\{n\left[\ln(2) + |h| + n\frac{\hat{q}_1}{2} + n^3\frac{\sigma_{q_1}^2}{8}\right]\right\},
\end{aligned}$$

where $\mathcal{A}(n) \sim \mathcal{O}(1)$. Thus, $h \sim \mathcal{O}(1)$, $\hat{q}_1 \sim \mathcal{O}(n^{-1})$ and $\sigma^2_{q_1} \sim \mathcal{O}(n^{-3})$. From now on we will take the off-diagonal elements of the RS matrix $\mathbf{Q}^t_{\mu k}$ equal to $g^t_{1\mu k}/n$, where $g^t_{1\mu k} \sim \mathcal{O}(1)$.

The form of the marginalised posterior at time t is then:

$$P^t(\mathbf{b}_k | \{y_{\nu\neq\mu}\}) = \frac{\displaystyle\int_{-\infty}^{\infty} \mathrm{d}x \, \exp\left\{ -n \frac{\left(x - h^t_{\mu k}\right)^2}{2 g^t_{1\mu k}} + x \sum_{\mathrm{a}=1}^{n} b^{\mathrm{a}}_k \right\}}{\displaystyle\int_{-\infty}^{\infty} \mathrm{d}x \, \exp\left\{ -n \Phi\left(x; h^t_{\mu k}, g^t_{1\mu k}\right)\right\}}, \tag{23}$$

where

$$\Phi\left(x; h^t_{\mu k}, g^t_{1\mu k}\right) = \frac{\left(x - h^t_{\mu k}\right)^2}{2 g^t_{1\mu k}} - \ln\left(2\cosh(x)\right).$$

The function $\Phi(x; h, g_1)$ presents one or two minima according to the following table:

h	g_1	Number of minima
$h \in \mathbb{R}$	$0 < g_1 \leq 1$	one min.
$\|h\| = h_c$	$g_1 > 1$	one min. and one hump
$\|h\| < h_c$	$g_1 > 1$	two min.

where $h_c = \sqrt{g_1(g_1-1)} - \cosh^{-1}\left(\sqrt{g_1}\right)$. g_1 plays the role of the inverse of the Temperature. Below the critical value $g_{1c} = 1$ a spontaneous magnetisation appears.
This is the result from the analysis of the equation:

$$\frac{\partial \Phi(x; h, g_1)}{\partial x} = \frac{x - h}{g_1} - \tanh(x) = 0. \tag{24}$$

The case with 2 maxima is presented in Figure 7.

We define the mean values from the distribution Eq. (23). If the field h is not zero, as it is shown in Figure 7, $[\exp(-\Phi)]^n$develops one dominant maximum when $n \to \infty$. For large enough n, only the contribution from this maximum is relevant to the integrals in Eq. (23). But if this is the case, it is possible to proof that the algorithm obtained from this assumption is the same as the one presented in [2]. Thus, the field has to be small enough in order to have a new regime where the two extrema contribute. At the same time, it is important to note that although small, a non zero field will favour the solution of Eq. (24) satisfying $\mathrm{sgn}(x) = \mathrm{sgn}(h)$. To analyse which is the appropriate behaviour of the field, we will explore the solutions of Eq. (24) in the regime $0 \lesssim |h| \ll 1$. With this aim, suppose that the solutions for the Eq. (24) at zero field are $x_0 = \pm g_1 |m|$ where $|m| \equiv |\tanh(x_0)|$ and $\mathrm{sgn}(m) = \mathrm{sgn}(h)$. If the field is small enough, we can expand the solutions of Eq. (24) as: $x_{\pm h} = \pm g_1 m + \xi(m, g_1)h$ where $\xi(m, g_1)h$ is expected to be small and satisfies $\mathrm{sgn}(\xi(m, g_1)h) = \mathrm{sgn}(h)$. Observe that if the field is positive (negative), both roots are displaced to the right (left) with respect to the zero field solutions. Using this expression for the roots in Eq. (24) and disregarding terms of $\mathcal{O}\left(h^2\right)$ we find that

$$\xi(m,g_1)=\frac{1}{1-g_1\left(1-m^2\right)}\,. \tag{25}$$

The expression for the exponent Φ near the roots and in the $0\lesssim |h| \ll 1$ regime is then $\Phi\left(x_{\pm h};h\rightarrow 0,g_1\right) \simeq \Phi\left(x_0;0,g_1\right)\mp mh = \Phi_0 \mp mh$, and, by the definition of the m, the product mh is positive defined.

Let us define $\beta_{\pm h}\left(m,g_1\right)\equiv\left(1-m^2\right)\left[1\mp 2\xi\left(m,g_1\right)mh\right]$. We expect that, for n large enough, the following approximation to be valid:

$$\begin{aligned}\exp\left\{-n\Phi\left(x;h\rightarrow 0,g_1\right)\right\} \simeq\ & \mathrm{e}^{-n\Phi_0}\left\{\mathrm{e}^{nmh}\exp\left\{-\frac{n}{2}\left[g_1^{-1}-\beta_h\left(m,g_1\right)\right]\left(x-x_h\right)^2\right\}\right.\\ &\left.+\mathrm{e}^{-nmh}\exp\left\{-\frac{n}{2}\left[g_1^{-1}-\beta_{-h}\left(m,g_1\right)\right]\left(x-x_{-h}\right)^2\right\}\right\}. \end{aligned} \tag{26}$$

Using Eq. (26) we can calculate the expression for the normalisation Eq. (22):

$$\begin{aligned}\mathcal{Z}\left(h\rightarrow 0,g_1\right) \simeq\ & \mathrm{e}^{-n\left(\Phi_0-mh\right)}\int \mathrm{d}x\,\exp\left\{-\frac{n}{2}\left[g_1^{-1}-\beta_h\left(m,g_1\right)\right]\left(x-x_h\right)^2\right\}\\ &+\mathrm{e}^{-n\left(\Phi_0+mh\right)}\int \mathrm{d}x\,\exp\left\{-\frac{n}{2}\left[g_1^{-1}-\beta_{-h}\left(m,g_1\right)\right]\left(x-x_{-h}\right)^2\right\}\\ \simeq\ & \sqrt{\frac{2\pi\xi\left(m,g_1\right)}{n\left(1-m^2\right)}}\mathrm{e}^{-n\Phi_0}\left\{\,\mathrm{e}^{nmh}\left(1-g_1\left(1-m^2\right)\xi^2\left(m,g_1\right)mh\right)\right.\\ &\left.+\ \mathrm{e}^{-nmh}\left(1+g_1\left(1-m^2\right)\xi^2\left(m,g_1\right)mh\right)\right\}\,. \end{aligned} \tag{27}$$

The mean value of a given function $f(x)$ with respect to the probability defined in Eq. (23) is then:

$$\begin{aligned}\left\langle f(x)|h\rightarrow 0,g_1\right\rangle \simeq\ & \mathcal{Z}^{-1}\mathrm{e}^{-n\left(\Phi_0-mh\right)}\int\mathrm{d}x\,\exp\left\{-\frac{n}{2}\left[g_1^{-1}-\left(1-m^2\right)\left(1-2\xi\left(m,g_1\right)mh\right)\right]\left(x-x_h\right)^2\right\}\\ &\left[f\left(x_h\right)+\left(x-x_h\right)f'\left(x_h\right)+\frac{1}{2}\left(x-x_h\right)^2 f''\left(x_h\right)\right]\\ &+\mathcal{Z}^{-1}\mathrm{e}^{-n\left(\Phi_0+mh\right)}\int\mathrm{d}x\,\exp\left\{-\frac{n}{2}\left[g_1^{-1}-\left(1-m^2\right)\left(1+2\xi\left(m,g_1\right)mh\right)\right]\left(x-x_{-h}\right)^2\right\}\\ &\left[f\left(x_{-h}\right)+\left(x-x_{-h}\right)f'\left(x_{-h}\right)+\frac{1}{2}\left(x-x_{-h}\right)^2 f''\left(x_{-h}\right)\right], \end{aligned}$$

which implies, considering that the integrals of the linear terms are zero and keeping only the leading terms in the expansions, that the expectation values have the form:

$$\begin{aligned}\left\langle f(x)|h\rightarrow 0,g_1\right\rangle \simeq\ & \left[1-\mathrm{e}^{-2nmh}\left(1+2\xi^2\left(m,g_1\right)mh\right)\right]\left\{f\left(x_h\right)+\frac{g_1}{2n}\xi\left(m,g_1\right)f''\left(x_h\right)\right\}\\ &+\mathrm{e}^{-2nmh}\left(1+2\xi^2\left(m,g_1\right)mh\right)f\left(x_{-h}\right), \end{aligned}$$

considering the expansion of $f\left(x_{\pm h}\right)\simeq f\left(\pm g_1 m+\xi\left(m,g_1\right)h\right)\simeq f\left(\pm g_1 m\right)+\xi\left(m,g_1\right)f'\left(\pm g_1 m\right)h$ and disregarding terms of $\mathcal{O}\left(h\mathrm{e}^{-2nmh}\right)$, we can write:

$$\left\langle f(x)|h\rightarrow 0,g_1\right\rangle \simeq f\left(mg_1\right)+\frac{g_1}{2n}\xi\left(m,g_1\right)f''\left(mg_1\right)-\mathrm{e}^{-2nmh}\left[f\left(mg_1\right)-f\left(-mg_1\right)\right]+f'\left(mg_1\right)\xi\left(m,g_1\right)h. \tag{28}$$

The one and two spin mean values are then:

$$\begin{aligned}\left\langle b_k^{\rm a}|h_{\mu k}^t\rightarrow 0,g_{\mu k}^t\right\rangle &= \sum_{\{\mathbf{b}_k\}} P^t\left(\mathbf{b}_k|\left\{y_{\nu\neq\mu}\right\}\right)b_k^{\rm a}=\left\langle \tanh(x)|h_{\mu k}^t\rightarrow 0,g_{1\mu k}^t\right\rangle\\ &\simeq \left[1-\frac{g_{1\mu k}^t}{n}\left[1-\left(m_{\mu k}^t\right)^2\right]\xi\left(m_{\mu k}^t,g_{1\mu k}^t\right)-2{\rm e}^{-2nm_{\mu k}^t h_{\mu k}^t}\right]m_{\mu k}^t\\ &\quad +\xi\left(m_{\mu k}^t,g_{1\mu k}^t\right)\left[1-\left(m_{\mu k}^t\right)^2\right]h_{\mu k}^t\end{aligned}$$

and

$$\left\langle b_k^{\rm a}b_k^{\rm b}|h_{\mu k}^t\rightarrow 0,g_{1\mu k}^t\right\rangle = P^t\left(\mathbf{b}_k|\left\{y_{\nu\neq\mu}\right\}\right)b_k^{\rm a}b_k^{\rm b}=\delta^{\rm ab}+\left(1-\delta^{\rm ab}\right)\left\langle \tanh^2(x)|h_{\mu k}^t\rightarrow 0,g_{1\mu k}^t\right\rangle\,,$$

where

$$\left\langle \tanh^2(x)|h_{\mu k}^t\rightarrow 0,g_{1\mu k}^t\right\rangle=\left(m_{\mu k}^t\right)^2+\xi\left(m_{\mu k}^t,g_{1\mu k}^t\right)\left[1-\left(m_{\mu k}^t\right)^2\right]\left\{\frac{g_{1\mu k}^t}{n}\left[1-3\left(m_{\mu k}^t\right)^2\right]+2m_{\mu k}^t h_{\mu k}^t\right\}\,,$$

and

$$\left\langle b_k^{\rm a}b_l^{\rm b}|h_{\mu k}^t\rightarrow 0,g_{1\mu k}^t\right\rangle=\left\langle b_k^{\rm a}|h_{\mu k}^t\rightarrow 0,g_{1\mu k}^t\right\rangle\left\langle b_l^{\rm b}|h_{\mu k}^t\rightarrow 0,g_{1\mu k}^t\right\rangle.$$

Thus, the leading terms for the covariance matrix are:

$$\begin{aligned}\left(\boldsymbol{\Psi}_{\mu kl}^t\right)^{\rm ab} &\equiv \left\langle b_k^{\rm a}b_l^{\rm b}|h_{\mu k}^t\rightarrow 0,g_{1\mu k}^t;h_{\mu l}^t\rightarrow 0,g_{1\mu l}^t\right\rangle-\left\langle b_k^{\rm a}|h_{\mu k}^t\rightarrow 0,g_{1\mu k}^t\right\rangle\left\langle b_l^{\rm b}|h_{\mu l}^t\rightarrow 0,g_{1\mu l}^t\right\rangle=\delta_{kl}\left(\boldsymbol{\Psi}_{\mu k}^t\right)^{\rm ab}\\ \left(\boldsymbol{\Psi}_{\mu k}^t\right)^{\rm ab} &\simeq \delta^{\rm ab}\left[1-\left(m_{\mu k}^t\right)^2\right]+\left(1-\delta^{\rm ab}\right)\left\{\frac{g_{1\mu k}^t}{n}\xi\left(m_{\mu k}^t,g_{1\mu k}^t\right)\left[1-\left(m_{\mu k}^t\right)^2\right]^2+4{\rm e}^{-2nm_{\mu k}^t h_{\mu k}^t}\left(m_{\mu k}^t\right)^2\right\}.\end{aligned}$$

If we ask the non-diagonal elements of this covariance matrix to have the same scaling as the inter-replica interaction matrix, the field has to behave in such a way that the exponential term contributes at most in $\mathcal{O}\left(n^{-1}\right)$. Thus we suppose the field to behave as $m_{\mu k}^t h_{\mu k}^t<\frac{1}{n}\ln\left|\frac{2n}{n_{\mu k}^t}\right|$, where the $n_{\mu k}^t$ are suitable constants. With this asymptotic behaviour, the expression for the entries in the covariance matrix is

$$\left(\boldsymbol{\Psi}_{\mu k}^t\right)^{\rm ab}\simeq\delta^{\rm ab}\left[1-\left(m_{\mu k}^t\right)^2\right]+\left(1-\delta^{\rm ab}\right)\frac{g_{1\mu k}^t\xi\left(m_{\mu k}^t,g_{1\mu k}^t\right)}{n}\left[1-\left(m_{\mu k}^t\right)^2\right]^2.$$

If the $\varepsilon_{\mu k}$ and $b_k^{\rm a}$ are unbiased variables, the variable $\mathit{\Delta}_{\mu k}^{\rm a}=\sum_{l\neq k}\varepsilon_{\mu l}b_l^{\rm a}$, by virtue of the central limit theorem, obeys a normal distribution, with mean value and covariance matrix given by:

$$\begin{aligned}\left(\mathbf{u}_{\mu k}^t\right)^{\rm a}\equiv\left\langle\mathit{\Delta}_{\mu k}^{\rm a}\right\rangle &= \sum_{\left\{\mathbf{b}_{l\neq k}\right\}}\prod_{l\neq k}P^t\left(\mathbf{b}_l|\left\{y_{\nu\neq\mu}\right\}\right)\sum_{l\neq k}\varepsilon_{\mu l}b_l^{\rm a}=\sum_{l\neq k}\varepsilon_{\mu l}m_{\mu l}^t\\ \left(\boldsymbol{\Upsilon}_{\mu k}^t\right)^{\rm ab}\equiv\left\langle\mathit{\Delta}_{\mu k}^{\rm a}\mathit{\Delta}_{\mu k}^{\rm b}\right\rangle-\left\langle\mathit{\Delta}_k^{\rm a}\right\rangle\left\langle\mathit{\Delta}_k^{\rm b}\right\rangle &= \sum_{\left\{\mathbf{b}_{l\neq k}\right\}}\prod_{l\neq k}P^t\left(\mathbf{b}_l|\left\{y_{\nu\neq\mu}\right\}\right)\sum_{\substack{l\neq k\\ j\neq k}}\varepsilon_{\mu l}\varepsilon_{\mu j}b_l^{\rm a}b_j^{\rm b}-\left(\sum_{l\neq k}\varepsilon_{\mu l}m_{\mu l}^t\right)^2\\ &= \sum_{l\neq k}\varepsilon_{\mu l}^2\left(\boldsymbol{\Psi}_{\mu lj}^t\right)^{\rm ab}=\delta^{\rm ab}X_{\mu k}+\left(1-\delta^{\rm ab}\right)\frac{1}{n}R_{\mu k}^t,\end{aligned}$$

where

$$X_{\mu k}^t \equiv \sum_{l\neq k} \varepsilon_{\mu l}^2 \left[1-\left(m_{\mu l}^t\right)^2\right] \tag{29}$$

$$R_{\mu k}^t \equiv \sum_{l\neq k} \varepsilon_{\mu l}^2\, g_{1\mu l}^t\, \xi\left(m_{\mu l}^t, g_{1\mu l}^t\right) \left[1-\left(m_{\mu l}^t\right)^2\right]^2 ,$$

are macroscopic variables of $\mathcal{O}(1)$. In particular, $R_{\mu k}^t$ is a free variable that can be used to optimise a given performance measure. This variables have the property of being self-averaging, therefore we can drop the sub-indexes μ and k.

Appendix B: The One Step Replica Symmetry Breaking (1RSB) Ansatz

The system has nL spins. Both, the number of blocks L is considered to be large. As before we are interested in the regime where $nL \to \infty$.

With this setting, the interaction term in Eq. (6) is now:

$$\mathbf{b}_k^\top \mathbf{Q}_{\mu k}^t \mathbf{b}_k = -q_{1\mu k}^t nL + \left(q_{1\mu k}^t - q_{2\mu k}^t\right) \sum_{\ell=1}^{L} \left(\sum_{\mathrm{a}=1}^{n} b_k^{\ell \mathrm{a}}\right)^2 + q_{2\mu k}^t \left(\sum_{\ell=1}^{L}\sum_{\mathrm{a}=1}^{n} b_k^{\ell \mathrm{a}}\right)^2 ,$$

thus we have now $L+1$ squared sums in the exponent that can be replaced by integrals:

$$P^t\left(\mathbf{b}_k \middle| \{y_{\nu\neq\mu}\}\right) = \left[\mathcal{Z}_{\mu k}^t\right]^{-1} \int \mathrm{d}\boldsymbol{x}\, \exp\left\{-\frac{x_0^2}{2q_{2\mu k}^t} - \sum_{\ell=1}^{L} \frac{x_\ell^2}{2\Delta q_{\mu k}^t} + \sum_{\ell=1}^{L} \left(x_0 + x_\ell + h_{\mu k}^t\right) \sum_{\mathrm{a}=1}^{n} b_k^{\ell \mathrm{a}}\right\} ,$$

where $\Delta q_{\mu k}^t \equiv q_{1\mu k}^t - q_{2\mu k}^t > 0$ and $\boldsymbol{x}^\top = (x_0, x_1, \ldots, x_L)$. We expect the free energy obtained from the well behaved distribution P^t to be self-averaging, thus:

$$\lim_{n\to\infty} \lim_{L\to\infty} \frac{1}{nL} \log\left(\overline{\mathcal{Z}_{\mu k}^t}\right) = \lim_{n\to\infty} \lim_{L\to\infty} \frac{1}{nL} \log\left(\mathcal{Z}_{\mu k}^t\left(h_{\mu k}^t, q_{1\mu k}^t, q_{2\mu k}^t\right)\right) ,$$

which is satisfied if the entries behave like $q_{2\mu k}^t \sim g_{2\mu k}^t/nL$ and $\Delta q_{\mu k}^t \sim g_{1\mu k}^t/n$, where $g_{1\mu k}^t$ and $g_{2\mu k}^t \sim \mathcal{O}(1)$. Using this new scaled parameters, the expression for the normalisation is $\mathcal{Z}_{\mu k}^t = \int \mathrm{d}\mathbf{x}\, \exp\left\{-nL\Phi\left(\mathbf{x}; h_{\mu k}^t, g_{1\mu k}^t, g_{2\mu k}^t\right)\right\}$ where

$$\Phi\left(\mathbf{x}; h_{\mu k}^t, g_{1\mu k}^t, g_{2\mu k}^t\right) \equiv \frac{x_0^2}{2g_{1\mu k}^t} + \frac{1}{L}\sum_{\ell=1}^{L} \frac{x_\ell^2}{2g_{2\mu k}^t} - \frac{1}{L}\sum_{\ell=1}^{L} \log\left[2\cosh\left(x_0 + x_\ell + h_{\mu k}^t\right)\right] .$$

As before we will drop the indexes μ, k, and t in order to simplify the notation. The critical points of the function $\Phi\left(\mathbf{x}; h, g_1, g_2\right)$ satisfy the following set of equations:

$$\frac{\partial \Phi}{\partial x_0} = \frac{x_0}{g_1} - \frac{1}{L}\sum_{\ell=1}^{L} \tanh\left(x_0 + x_\ell + h\right) = 0$$

$$\frac{\partial \Phi}{\partial x_\ell} = \frac{x_\ell}{g_2} - \tanh\left(x_0 + x_\ell + h\right) = 0 ,$$

which are satisfied for the following values:

$$\begin{aligned}\overline{x}_0 &= \frac{g_1}{g_2}\frac{1}{L}\sum_{\ell=1}^{L}\overline{x}_\ell = \frac{g_1}{g_2}\overline{x} \\ \frac{\overline{x}_\ell}{g_2} &= \tanh\left(\overline{x}_\ell + \frac{g_1}{g_2}\overline{x} + h\right),\end{aligned} \tag{30}$$

where $\overline{x} \equiv \frac{1}{L}\sum_{\ell=1}^{L}\overline{x}_\ell$. The second equation in the set Eq. (30), has the same form for all $\ell = 1, \ldots, L$ and in the small field regime it has at most three different solutions. From the three possible solutions, one is a local maximum. From the other two, the one that has the same sign as h is the dominant one. Thus we can expect, for all ℓ, $\overline{x}_\ell = \overline{x}$. This reduces the set of $L+1$ equations to only one:

$$\frac{\overline{x}}{g_2} = \tanh\left(\frac{G}{g_2}\overline{x} + h\right),$$

where $G \equiv g_1 + g_2$. With the substitution $u = (G/g_2)\,\overline{x}$ the equation has the same form as Eq. (24), i.e. $u = G\tanh(u+h)$. If again we consider that the field h is small, the solutions can be expressed as an expansion of the zero field solutions $u_{\pm h} \simeq \pm Gm + \xi(m,G)h$, where $\xi(m,G)$ is given by Eq. (25), and $\mathrm{sgn}(m) = \mathrm{sgn}(h)$. Using these expansions the critical values are given by: $\overline{x}_{0,\pm h} \simeq g_2\left[\pm m + G^{-1}\xi(m,G)h\right]$ and $\overline{x}_{\ell,\pm h} \simeq g_1\left[\pm m + G^{-1}\xi(m,G)h\right]$ for all $\ell = 1, \ldots, L$.

As in the RS case, the expansion of Φ around the critical points in the small field regime is $\Phi\left(\overline{\mathbf{x}}_{\pm h}; h \to 0, g_1, g_2\right) \simeq \Phi\left(\overline{\mathbf{x}}_0; 0, g_1, g_2\right) \mp mh = \Phi_0 \mp mh$. So the dominant solution is the one that shares the sign with the field.

For a system sufficiently large with nL spins, we expect the following expansion to be valid:

$$\begin{aligned}\exp\left\{-nL\Phi\left(\mathbf{x}; h \to 0, g_1, g_2\right)\right\} \simeq\ & \mathrm{e}^{-nL\Phi_0}\left\{\mathrm{e}^{nLmh}\exp\left[-\frac{nL}{2}\left(\mathbf{x}-\overline{\mathbf{x}}_h\right)^{\mathsf{T}}\mathbf{H}_{\Phi,h}\left(\mathbf{x}-\overline{\mathbf{x}}_h\right)\right]\right. \\ & \left.+\mathrm{e}^{-nLmh}\exp\left[-\frac{nL}{2}\left(\mathbf{x}-\overline{\mathbf{x}}_{-h}\right)^{\mathsf{T}}\mathbf{H}_{\Phi,-h}\left(\mathbf{x}-\overline{\mathbf{x}}_{-h}\right)\right]\right\},\end{aligned} \tag{31}$$

where $\mathbf{H}_{\Phi,\pm h}$ is the Hessian of Φ in $\overline{\mathbf{x}}_{\pm h}$.

Let us define the quantity $\beta_{\pm h} \equiv (1-m^2)\left[1 \mp 2\xi(m,G)\,mh\right]$. The entries of the Hessian are:

$$\begin{aligned}\left.\frac{\partial^2\Phi}{\partial x_0^2}\right|_{\overline{\mathbf{x}}_{\pm h}} &\simeq g_1^{-1} - \beta_{\pm h} \equiv \alpha_{\pm h} \\ \left.\frac{\partial^2\Phi}{\partial x_\ell^2}\right|_{\overline{\mathbf{x}}_{\pm h}} &\simeq \frac{1}{L}\left(g_2^{-1} - \beta_{\pm h}\right) \equiv \frac{1}{L}\gamma_{\pm h} \\ \left.\frac{\partial^2\Phi}{\partial x_0 \partial x_\ell}\right|_{\overline{\mathbf{x}}_{\pm h}} &\simeq -\frac{1}{L}\beta_{\pm h} \\ \left.\frac{\partial^2\Phi}{\partial x_\ell \partial x_{\ell'}}\right|_{\overline{\mathbf{x}}_{\pm h}} &= 0\end{aligned}$$

thus, the correspondent characteristic equation is:

$$\det\left(\mathbf{H}_{\Phi,\pm h} - \lambda\mathbb{1}\right) = \left(\frac{1}{L}\gamma_{\pm h} - \lambda\right)^{L-1}\left\{\left(\alpha_{\pm h} - \lambda\right)\left(\frac{1}{L}\gamma_{\pm h} - \lambda\right) - \frac{1}{L}\beta_{\pm h}^2\right\} = 0\,.$$

The solutions for this equation, disregarding terms of $\mathcal{O}\left(L^{-2}\right)$ and $\mathcal{O}\left(hL^{-1}\right)$, are:

$$\begin{aligned} \lambda_{0,\pm h} &\simeq \alpha_0 + \frac{1}{L}\frac{\beta_0^2}{\alpha_0} \pm 2\xi(m,G)\left(1-m^2\right)\, mh \\ \lambda_{1,\pm h} \simeq \lambda_1 &= \frac{1}{L}\left(\gamma_0 - \frac{\beta_0^2}{\alpha_0}\right) \\ \lambda_{\ell,\pm h} \simeq \lambda_\ell &= \frac{1}{L}\gamma_0 \quad \forall\, \ell = 2,\ldots,L\,. \end{aligned} \tag{32}$$

The correspondent eigenvectors, up to order L^{-1}, are:

$$\begin{aligned} \mathbf{u}_{0,\pm h} &= \left(1, \overbrace{-\frac{1}{L}\frac{\beta_{\pm h}}{\alpha_{\pm h}}, -\frac{1}{L}\frac{\beta_{\pm h}}{\alpha_{\pm h}}, \ldots, -\frac{1}{L}\frac{\beta_{\pm h}}{\alpha_{\pm h}}}^{L \text{ times}}\right)^{\mathsf{T}} \\ \mathbf{u}_{1,\pm h} &= \frac{1}{\sqrt{L}}\left(\frac{\beta_{\pm h}}{\alpha_{\pm h}}, \overbrace{1,1,\ldots,1}^{L \text{ times}}\right)^{\mathsf{T}} \\ \mathbf{u}_{\ell,\pm h} &= \frac{1}{\sqrt{\ell\,(\ell-1)}}\left(0, \overbrace{1,1,\ldots,1}^{\ell-1 \text{ times}}, -\left(\ell-1\right), \overbrace{0,0,\ldots,0}^{L-\ell \text{ times}}\right)^{\mathsf{T}} \qquad \forall\, \ell = 2,\ldots,L\,. \end{aligned} \tag{33}$$

This vectors satisfy the normalisation condition $\mathbf{u}_{\ell,\pm h}^{\mathsf{T}}\mathbf{u}_{\ell',\pm h} = \delta^{\ell\ell'}\left[1+\mathcal{O}\left(L^{-1}\right)\right] \quad \forall \ell,\ell' = 0,1,\ldots,L$. The linear transformation from the canonical basis to the basis of eigenvectors is then represented by the matrix with entries equal to:

$$\begin{aligned} \left(\mathbf{U}_{\pm h}\right)_{ij} \simeq \left(\mathbf{U}_0\right)_{ij} &= \delta_{0i}\delta_{0j} + \frac{1}{\sqrt{j(j-1)}}\left[\sum_{k=1}^{j-1}\delta_{ki} - \left(1-\delta_{0j}\right)\left(1-\delta_{1j}\right)\delta_{ij}(j-1)\right] \\ &\quad + \frac{1}{\sqrt{L}}\,\delta_{1j}\left[\delta_{0i}\,\frac{\beta_0}{\alpha_0} + \left(1-\delta_{0i}\right)\right] - \frac{1}{L}\,\delta_{0j}\left(1-\delta_{0i}\right)\frac{\beta_0}{\alpha_0}\,, \end{aligned} \tag{34}$$

disregarding terms of $\mathcal{O}\left(hL^{-1/2}\right)$. Because this transformation is a rigid rotation, the following properties are satisfied: $\left|\det\left(\mathbf{U}_{\pm}\right)\right| = 1$ and $\mathbf{U}_{\pm h}^{\mathsf{T}}\mathbf{U}_{\pm h} = \mathbf{U}_{\pm h}\mathbf{U}_{\pm h}^{\mathsf{T}} = \mathbb{1}$.

In Eq. (31), the second order terms can be re-written using the diagonal representation of the Hessian. Therefore, keeping only terms of order $\mathcal{O}\left(L^{-1}\right)$ we have that: $\left(\mathbf{x}-\mathbf{x}_{\pm h}\right)^{\mathsf{T}}\mathbf{H}_{\Phi,\pm h}\left(\mathbf{x}-\mathbf{x}_{\pm h}\right) = \left(\mathbf{x}-\mathbf{x}_{\pm h}\right)^{\mathsf{T}}\mathbf{U}_0\mathbf{U}_0^{\mathsf{T}}\mathbf{H}_{\Phi,\pm h}\mathbf{U}_0\mathbf{U}_0^{\mathsf{T}}\left(\mathbf{x}-\mathbf{x}_{\pm h}\right) = \left(\mathbf{y}-\mathbf{y}_{\pm h}\right)^{\mathsf{T}}\mathbf{H}'_{\Phi,\pm h}\left(\mathbf{y}-\mathbf{y}_{\pm h}\right)$, where $\mathbf{y} \equiv \mathbf{U}_0^{\mathsf{T}}\mathbf{x}$ and $\mathbf{H}'_{\Phi,\pm h} \equiv \mathbf{U}_0^{\mathsf{T}}\mathbf{H}_{\Phi,\pm h}\mathbf{U}_0$ is the diagonal representation of the Hessian, i.e. $\left(\mathbf{H}'_{\Phi,\pm h}\right)_{ij} = \delta_{ij}\lambda_{i,\pm h}$. Using this information and Eq. (31) we can write down the following expression for the normalisation:

$$\begin{aligned} \mathcal{Z}\left(h\to 0, g_1, g_2\right) &\simeq \mathrm{e}^{-nL(\Phi_0 - mh)}\int \mathrm{d}\mathbf{y}\,\exp\left[-\frac{nL}{2}\left(\mathbf{y}-\overline{\mathbf{y}}_h\right)^{\mathsf{T}}\mathbf{H}'_{\Phi,h}\left(\mathbf{y}-\overline{\mathbf{y}}_h\right)\right] \\ &\quad + \mathrm{e}^{-nL(\Phi_0 + mh)}\int \mathrm{d}\mathbf{y}\,\exp\left[-\frac{nL}{2}\left(\mathbf{y}-\overline{\mathbf{y}}_{-h}\right)^{\mathsf{T}}\mathbf{H}'_{\Phi,-h}\left(\mathbf{y}-\overline{\mathbf{y}}_{-h}\right)\right] \\ &\simeq \mathrm{e}^{-nL\Phi_0}\left(\frac{2\pi}{nL}\right)^{\frac{L+1}{2}}\left[\mathrm{e}^{nLmh}\prod_{\ell=0}^{L}\lambda_{\ell,h}^{-\frac{1}{2}} + \mathrm{e}^{-nLmh}\prod_{\ell=0}^{L}\lambda_{\ell,-h}^{-\frac{1}{2}}\right]. \end{aligned}$$

For a small field, the product of the eigenvalues can be approached by:

$$\prod_{\ell=0}^{L}\lambda_{\ell,\pm h}^{-\frac{1}{2}} \simeq \left[1 \mp g_1\xi\left(m,g_1\right)\xi(m,G)\,\left(1-m^2\right)mh\right]\prod_{\ell=0}^{L}\lambda_{\ell,0}^{-\frac{1}{2}}\,.$$

Thus

$$\mathcal{Z}\left(h\to 0,g_1,g_2\right)\simeq \mathrm{e}^{-nL\Phi_0}\left(\frac{2\pi}{nL}\right)^{\frac{L+1}{2}}\prod_{\ell=0}^{L}\lambda_{\ell,0}^{-\frac{1}{2}}\left\{\mathrm{e}^{nLmh}\left[1-g_1\xi\left(m,g_1\right)\xi(m,G)\left(1-m^2\right)mh\right]\right.$$
$$\left.+\,\mathrm{e}^{-nLmh}\left[1+g_1\xi\left(m,g_1\right)\xi(m,G)\left(1-m^2\right)mh\right]\right\}\,.$$

The mean value of a given function $f(\mathbf{x})$ is then:

$$\begin{aligned}\langle f(\mathbf{x})|h\to 0,g_1,g_2\rangle \simeq\ & \mathcal{Z}^{-1}\mathrm{e}^{-nL(\Phi_0-mh)}\int \mathrm{d}\mathbf{y}\,\exp\left\{-\frac{nL}{2}\left(\mathbf{y}-\overline{\mathbf{y}}_h\right)^{\mathsf{T}}\mathbf{H}'_{\Phi,h}\left(\mathbf{y}-\overline{\mathbf{y}}_h\right)\right\}\\ &\left[f\left(\mathbf{x}_h\right)+\frac{1}{2}\left(\mathbf{y}-\overline{\mathbf{y}}_h\right)^{\mathsf{T}}\mathbf{H}'_{f,h}\left(\mathbf{y}-\overline{\mathbf{y}}_h\right)\right]\\ &+\mathcal{Z}^{-1}\mathrm{e}^{-nL(\Phi_0+mh)}\int \mathrm{d}\mathbf{y}\,\exp\left\{-\frac{nL}{2}\left(\mathbf{y}-\overline{\mathbf{y}}_{-h}\right)^{\mathsf{T}}\mathbf{H}'_{\Phi,-h}\left(\mathbf{y}-\overline{\mathbf{y}}_{-h}\right)\right\}\\ &\left[f\left(\mathbf{x}_{-h}\right)+\frac{1}{2}\left(\mathbf{y}-\overline{\mathbf{y}}_{-h}\right)^{\mathsf{T}}\mathbf{H}'_{f,-h}\left(\mathbf{y}-\overline{\mathbf{y}}_{-h}\right)\right]\,,\end{aligned}$$

where $\mathbf{H}'_{f,\pm h}$ is the Hessian of the function f in the basis of eigenvectors of $\mathbf{H}_{\Phi\pm h}$, evaluated in the critical points. The linear terms in the expansion of $f\left(\mathbf{x}\right)$ do not contribute to the expectation value. The Gaussian integral of the cross products of the type $\left(y_i-\overline{y}_{i,\pm h}\right)\left(y_j-\overline{y}_{j,\pm h}\right)$ with $i\neq j$ are zero, thus the Gaussian integral of the second term in the expansion of $f\left(\mathbf{x}\right)$ is:

$$\begin{aligned}I_\pm &= \frac{1}{2}\mathcal{Z}^{-1}\mathrm{e}^{-nL(\Phi_0\mp mh)}\int \mathrm{d}\mathbf{y}\,\exp\left[-\frac{nL}{2}\left(\mathbf{y}-\overline{\mathbf{y}}_{\pm h}\right)^{\mathsf{T}}\mathbf{H}'_{\Phi,\pm h}\left(\mathbf{y}-\overline{\mathbf{y}}_{\pm h}\right)\right]\left(\mathbf{y}-\overline{\mathbf{y}}_{\pm h}\right)^{\mathsf{T}}\mathbf{H}'_{f,\pm h}\left(\mathbf{y}-\overline{\mathbf{y}}_{\pm h}\right)\\ I_+ &\simeq \frac{1}{2}\left\{1-\mathrm{e}^{-2nLmh}\left[1+g_1\xi\left(m,g_1\right)\xi(m,G)\left(1-m^2\right)mh\right]\right\}\frac{1}{nL}\sum_{\ell=0}^{L}\lambda_{\ell,h}^{-1}\left(\mathbf{H}'_{f,h}\right)_{\ell\ell}\\ I_- &\simeq \frac{1}{2}\mathrm{e}^{-2nLmh}\left[1+g_1\xi\left(m,g_1\right)\xi(m,G)\left(1-m^2\right)mh\right]\frac{1}{nL}\sum_{\ell=0}^{L}\lambda_{\ell,-h}^{-1}\left(\mathbf{H}'_{f,-h}\right)_{\ell\ell}\,.\end{aligned}$$

Using the expansion $f\left(\mathbf{x}_\pm\right)\simeq f\left(\pm\mathbf{x}_0+h\,\boldsymbol{\xi}(m,G)\right)\simeq f\left(\pm\mathbf{x}_0\right)+h\,\boldsymbol{\xi}^{\mathsf{T}}(m,G)\nabla f\left(\pm\mathbf{x}_0\right)=f\left(\pm\mathbf{x}_0\right)+\delta f\left(\pm\mathbf{x}_0\right)h$ where $\mathbf{x}_0^{\mathsf{T}}=m\left(g_1,\overbrace{g_2,g_2,\ldots,g_2}^{L\text{ times}}\right)$ and $\boldsymbol{\xi}^{\mathsf{T}}(m,G)=G^{-1}\xi(m,G)\left(g_1,\overbrace{g_2,g_2,\ldots,g_2}^{L\text{ times}}\right)$, the diagonal entries of the transformed Hessian are:

$$\begin{aligned}\left(\mathbf{H}'_{f,\pm h}\right)_{\ell\ell} &= \sum_{i,j=0}^{L}\left(\mathbf{U}_{\pm h}\right)_{\ell i}\left(\mathbf{U}_{\pm h}\right)_{\ell j}\left(\mathbf{H}_{f,\pm h}\right)_{ij}\\ &= \sum_{i,j=0}^{L}\left(\mathbf{U}_{\pm h}\right)_{\ell i}\left(\mathbf{U}_{\pm h}\right)_{\ell j}\left.\frac{\partial^2 f}{\partial x_i\partial x_j}\right|_{\mathbf{x}_\pm}\\ &\simeq \sum_{i,j=0}^{L}\left(\mathbf{U}_0\right)_{\ell i}\left(\mathbf{U}_0\right)_{\ell j}\left(\left.\frac{\partial^2 f}{\partial x_i\partial x_j}\right|_{\pm\mathbf{x}_0}+h\,\boldsymbol{\xi}^{\mathsf{T}}(m,G)\nabla\left.\frac{\partial^2 f}{\partial x_i\partial x_j}\right|_{\pm\mathbf{x}_0}\right)\\ &\simeq \left(\mathbf{H}'_f\big|_{\pm\mathbf{x}_0}\right)_{\ell\ell}+\left(\delta\mathbf{H}'_f\big|_{\pm\mathbf{x}_0}\right)_{\ell\ell}h\,,\end{aligned}$$

and in this manner, the last factor in the integrals becomes:

$$\frac{1}{nL}\sum_{\ell=0}^{L}\lambda_{\ell,-h}^{-1}\left(\mathbf{H}'_{f,\pm h}\right)_{\ell\ell}\simeq\frac{1}{nL}\frac{1}{\alpha_0}\left(\mathbf{H}'_f\big|_{\pm\mathbf{x}_0}\right)_{00}+\frac{1}{n}\frac{\alpha_0}{\left(\alpha_0\gamma_0-\beta_0^2\right)}\left(\mathbf{H}'_f\big|_{\pm\mathbf{x}_0}\right)_{11}+\frac{1}{n}\frac{1}{\gamma_0}\sum_{\ell=2}^{L}\left(\mathbf{H}'_f\big|_{\pm\mathbf{x}_0}\right)_{\ell\ell}\,,$$

disregarding terms of $\mathcal{O}\left(\frac{h}{n}+\frac{h}{L}\right)$. Thus:

$$\langle f(\mathbf{x})|h\to 0, g_1, g_2\rangle \simeq f(\mathbf{x}_0)+\frac{1}{2}\frac{1}{nL}\sum_{\ell=0}^{L}\lambda_{\ell,0}^{-1}\left(\mathbf{H}'_f\big|_{\mathbf{x}_0}\right)_{\ell\ell}-\mathrm{e}^{-2nLmh}[f(\mathbf{x}_0)-f(-\mathbf{x}_0)]+\delta f(\mathbf{x}_0)h, \quad (35)$$

where we have disregarded terms of $\mathcal{O}\left(\frac{h}{n}+\frac{h}{L}\right)$,$\mathcal{O}\left(\frac{1}{n}\mathrm{e}^{-2nLmh}\right)$ and $\mathcal{O}\left(h\mathrm{e}^{-2nLmh}\right)$. By simple inspection, Eq. (35) is equivalent to the RS mean value Eq. (28).

The one spin mean value is then:

$$\left\langle b_k^{\ell\mathrm{a}}|h_{\mu k}^t\to 0, g_{1\mu k}^t, g_{2\mu k}^t\right\rangle = \sum_{\{\mathbf{b}_k\}}P^t\left(\mathbf{b}_k|\left\{y_{\nu\neq\mu}\right\}\right)b_k^{\ell\mathrm{a}} = \left\langle\tanh\left(x_0+x_\ell+h_{\mu k}^t\right)|h_{\mu k}^t\to 0, g_{1\mu k}^t, g_{2\mu k}^t\right\rangle .$$

The expansion for $f(\mathbf{x})=\tanh\left(x_0+x_\ell+h_{\mu k}^t\right)$ is

$$f(\mathbf{x})\simeq m_{\mu k}^t+\left[1-\left(m_{\mu k}^t\right)^2\right]\left(1,\overbrace{0,0,\ldots,0}^{\ell-1\text{ times}},1,\overbrace{0,0,\ldots,0}^{L-\ell\text{ times}}\right)^{\mathsf{T}}\boldsymbol{\xi}\left(m_{\mu k}^t, G_{\mu k}^t\right)h_{\mu k}^t\,,$$

therefore

$$\begin{aligned}\left\langle b_k^{\ell\mathrm{a}}|h_{\mu k}^t\to 0, g_{1\mu k}^t, g_{2\mu k}^t\right\rangle &\simeq \left(1-2\mathrm{e}^{-2nLm_{\mu k}^t h_{\mu k}^t}\right)m_{\mu k}^t+\xi\left(m_{\mu k}^t, G_{\mu k}^t\right)\left[1-\left(m_{\mu k}^t\right)^2\right]h_{\mu k}^t\\ &\quad -m_{\mu k}^t\left[1-\left(m_{\mu k}^t\right)^2\right]\frac{1}{nL}\sum_{k=0}^{L}\lambda_{k,0}^{-1}\left(\mathbf{M}'_{0\ell}\right)_{kk}\,,\end{aligned}$$

where $\left(\mathbf{M}_{0\ell}\right)_{ij}=\delta_{0i}\delta_{0j}+\delta_{0i}\delta_{\ell j}+\delta_{\ell i}\delta_{0j}+\delta_{\ell i}\delta_{\ell j}$ is a matrix such that $\mathbf{H}_{\tanh(x_0+x_\ell)}\big|_{\mathbf{x}_0} = -2m_{\mu k}^t\left[1-\left(m_{\mu k}^t\right)^2\right]\mathbf{M}_{0\ell}$. In the basis of eigenvalues of $\mathbf{H}_{\varPhi}$, the expression for the diagonal elements of this matrix is:

$$\begin{aligned}\left(\mathbf{M}'_{0\ell}\right)_{kk} &= \sum_{i,j=0}^{L}\left(\mathbf{U}_{\pm h}\right)_{ik}\left(\mathbf{U}_{\pm h}\right)_{jk}\left(\delta_{0i}\delta_{0j}+\delta_{0i}\delta_{\ell j}+\delta_{\ell i}\delta_{0j}+\delta_{\ell i}\delta_{\ell j}\right)\\ &= \left(\left(\mathbf{U}_{\pm h}\right)_{0k}+\left(\mathbf{U}_{\pm h}\right)_{\ell k}\right)^2\\ \left(\mathbf{M}'_{0\ell}\right)_{00} &\simeq 1-\frac{2}{L}\frac{\beta_0}{\alpha_0}\\ \left(\mathbf{M}'_{0\ell}\right)_{11} &\simeq \frac{1}{L}\left(\frac{\alpha_0+\beta_0}{\alpha_0}\right)^2\\ \left(\mathbf{M}'_{0\ell}\right)_{kk} &= \delta_{k\ell}\frac{\ell-1}{\ell}+\varTheta\left(k-\ell+1\right)\frac{1}{k\left(k-1\right)}\quad\forall\,\ell=2,\ldots,L\,,\end{aligned} \quad (36)$$

where $\varTheta(l-k)=1$ if $l>k$ and 0 otherwise. The sum of the inverse of the eigenvalues times the diagonal elements Eq. (36) is:

$$\begin{aligned}
\frac{1}{nL}\sum_{k=0}^{L}\lambda_{k,0}^{-1}\left(\mathbf{M}'_{0\ell}\right)_{kk} &\simeq \frac{1}{n}\frac{1}{\gamma_0}\sum_{k=2}^{L}\left[\delta_{k\ell}\frac{\ell-1}{\ell}+\Theta\left(k-\ell+1\right)\frac{1}{k\left(k-1\right)}\right]+\frac{1}{nL}\frac{1}{\alpha_0}\left[1+\frac{\left(\alpha_0+\beta_0\right)^2}{\alpha_0\gamma_0-\beta_0^2}\right] \\
&= \frac{1}{n}\frac{1}{\gamma_0}\left[\sum_{k=2}^{L}\frac{1}{k\left(k-1\right)}\right]+\frac{1}{nL}\frac{1}{\alpha_0}\left[1+\frac{\left(\alpha_0+\beta_0\right)^2}{\alpha_0\gamma-\beta_0^2}\right] \\
&= \frac{1}{n}\frac{1}{\gamma_0}+\frac{1}{nL}\frac{1}{\alpha_0}\left[1+\frac{\left(\alpha_0+\beta_0\right)^2}{\alpha_0\gamma_0-\beta_0^2}-\frac{\alpha_0}{\gamma_0}\right] \\
&= \frac{1}{n}g_{2\mu k}^t\xi\left(m_{\mu k}^t,g_{2\mu k}^t\right)+\frac{1}{nL}\left[G_{\mu k}^t\xi\left(m_{\mu k}^t,G_{\mu k}^t\right)-g_{2\mu k}^t\xi\left(m_{\mu k}^t,g_{2\mu k}^t\right)\right]
\end{aligned}$$

where we have used that $\sum_{k=2}^{L}\left[k\left(k-1\right)\right]^{-1} = (L-1)/L$, $\gamma_0^{-1} = g_{2\mu k}^t\xi\left(m_{\mu k}^t,g_{2\mu k}^t\right)$ and $\frac{1}{\alpha_0}\left[1+\frac{\left(\alpha_0+\beta_0\right)^2}{\alpha_0\gamma_0-\beta_0^2}-\frac{\alpha_0}{\gamma_0}\right] = G_{\mu k}^t\xi\left(m_{\mu k}^t,G_{\mu k}^t\right)-g_{2\mu k}^t\xi\left(m_{\mu k}^t,g_{2\mu k}^t\right)$. The final expression for the expectation value of one spin is:

$$\begin{aligned}
\left\langle b_k^{\ell\mathrm{a}}|h_{\mu k}^t\to 0,g_{1\mu k}^t,g_{2\mu k}^t\right\rangle &\simeq \left(1-2\mathrm{e}^{-2nLm_{\mu k}^t h_{\mu k}^t}\right)m_{\mu k}^t-\frac{g_{2\mu k}^t\xi\left(m_{\mu k}^t,g_{2\mu k}^t\right)}{n}\left[1-\left(m_{\mu k}^t\right)^2\right]m_{\mu k}^t \\
&-\frac{G_{\mu k}^t\xi\left(m_{\mu k}^t,G_{\mu k}^t\right)-g_{2\mu k}^t\xi\left(m_{\mu k}^t,g_{2\mu k}^t\right)}{nL}\left[1-\left(m_{\mu k}^t\right)^2\right]m_{\mu k}^t \\
&+\xi\left(m_{\mu k}^t,G_{\mu k}^t\right)\left[1-\left(m_{\mu k}^t\right)^2\right]h_{\mu k}^t\,. \qquad (37)
\end{aligned}$$

To calculate $\left\langle b_k^{\ell\mathrm{a}}b_k^{\ell\mathrm{a}'}|h_{\mu k}^t\to 0,g_{\mu k}^t,\Delta g_{\mu k}^t\right\rangle$ which is an off-diagonal element in the same block, we can apply the Eq. (35) with $f(\mathbf{x})=\tanh^2\left(x_0+x_\ell+h_{\mu k}^t\right)$, thus the Hessian matrix is $\mathbf{H}_{\tanh^2(x_0+x_\ell)}\big|_{\mathbf{x}_0} = 2\left[1-\left(m_{\mu k}^t\right)^2\right]\left[1-3\left(m_{\mu k}^t\right)^2\right]\mathbf{M}_{0\ell}$, thus:

$$\begin{aligned}
\left\langle b_k^{\ell\mathrm{a}}b_k^{\ell\mathrm{a}'}|h_{\mu k}^t\to 0,g_{1\mu k}^t,g_{2\mu k}^t\right\rangle &\simeq \left(m_{\mu k}^t\right)^2+\frac{g_{2\mu k}^t\xi\left(m_{\mu k}^t,g_{2\mu k}^t\right)}{n}\left[1-\left(m_{\mu k}^t\right)^2\right]\left[1-3\left(m_{\mu k}^t\right)^2\right] \\
&+\frac{G_{\mu k}^t\xi\left(m_{\mu k}^t,G_{\mu k}^t\right)-g_{2\mu k}^t\xi\left(m_{\mu k}^t,g_{2\mu k}^t\right)}{nL}\left[1-\left(m_{\mu k}^t\right)^2\right]\left[1-3\left(m_{\mu k}^t\right)^2\right] \\
&+2\xi\left(m_{\mu k}^t,G_{\mu k}^t\right)\left[1-\left(m_{\mu k}^t\right)^2\right]m_{\mu k}^t h_{\mu k}^t. \qquad (38)
\end{aligned}$$

Finally, to calculate the expectation value for the product of two spins belonging to different blocks, $\left\langle b_k^{\ell\mathrm{a}}b_k^{\ell'\mathrm{a}}|h_{\mu k}^t\to 0,g_{1\mu k}^t,g_{2\mu k}^t\right\rangle$ we set $f(\mathbf{x}) = \tanh\left(x_0+x_\ell+h_{\mu k}^t\right)\tanh\left(x_0+x_{\ell'}+h_{\mu k}^t\right)$, thus the Hessian matrix is:

$$\begin{aligned}
\left(\mathbf{H}_{\tanh(x_0+x_\ell)\tanh(x_0+x_{\ell'})}\big|_{\mathbf{x}_0}\right)_{ij} &= \mathcal{M}_0\left(m_{\mu k}^t\right)\left(2\delta_{i0}\delta_{j0}+\delta_{i0}\delta_{j\ell}+\delta_{i\ell}\delta_{j0}+\delta_{i0}\delta_{j\ell'}+\delta_{i\ell'}\delta_{j0}\right) \\
&+\mathcal{M}_1\left(m_{\mu k}^t\right)\left(\delta_{i\ell}\delta_{j\ell'}+\delta_{i\ell'}\delta_{j\ell}\right)-2\mathcal{M}_2\left(m_{\mu k}^t\right)\left(\delta_{i\ell}\delta_{j\ell}+\delta_{i\ell'}\delta_{j\ell'}\right)\,,
\end{aligned}$$

where $\mathcal{M}_0\left(m_{\mu k}^t\right) \equiv \left[1-\left(m_{\mu k}^t\right)^2\right]\left[1-3\left(m_{\mu k}^t\right)^2\right]$, $\mathcal{M}_1\left(m_{\mu k}^t\right) \equiv \left[1-\left(m_{\mu k}^t\right)^2\right]^2$ and $\mathcal{M}_2\left(m_{\mu k}^t\right) \equiv \left(m_{\mu k}^t\right)^2\left[1-\left(m_{\mu k}^t\right)^2\right]$. The diagonal elements $\mathcal{K}_{0\ell\ell';k}\equiv\left(\mathbf{H}'_{\tanh(x_0+x_\ell)\tanh(x_0+x_{\ell'})}\big|_{\mathbf{x}_0}\right)_{kk}$ in the basis of eigenvectors of $\mathbf{H}_{\Phi}$ are:

$$
\begin{aligned}
\mathcal{K}_{0\ell\ell';0} &\simeq 2\mathcal{M}_0\left(m^t_{\mu k}\right) \\
\mathcal{K}_{0\ell\ell';1} &\simeq \frac{2\mathcal{M}_0\left(m^t_{\mu k}\right)}{L}\left[\frac{\beta_0}{\alpha_0}\left(\frac{\beta_0+2\alpha_0}{\alpha_0}\right)+1\right] \\
\mathcal{K}_{0\ell\ell';j} &= -2\delta_{j\ell}\mathcal{M}_2\left(m^t_{\mu k}\right)\frac{\ell-1}{\ell}-\left[\Theta\left(j-\ell\right)+\Theta\left(\ell'-j\right)\right]\frac{\mathcal{M}_2\left(m^t_{\mu k}\right)}{j(j-1)} \\
&\quad -2\delta_{j\ell'}\left[\frac{\mathcal{M}_1\left(m^t_{\mu k}\right)}{\ell'}+\frac{\mathcal{M}_2\left(m^t_{\mu k}\right)}{\ell'}\left(\ell'-1+\frac{1}{\ell'-1}\right)\right]+2\Theta\left(j-\ell'\right)\frac{\mathcal{M}_0\left(m^t_{\mu k}\right)}{j(j-1)}\,,
\end{aligned}
$$

thus, the sum of the diagonal elements is:

$$
\begin{aligned}
\frac{1}{2}\frac{1}{nL}\sum_{k=0}^{L}\lambda^{-1}_{k,-h}\mathcal{K}_{0\ell\ell';k} &\simeq \frac{1}{nL}\frac{\mathcal{M}_0\left(m^t_{\mu k}\right)}{\alpha_0}\left[1+\frac{\left(\beta_0+\alpha_0\right)^2}{\left(\alpha_0\gamma_0-\beta_0^2\right)}\right]-\frac{1}{n}\frac{1}{\gamma_0}\left\{\mathcal{M}_2\left(m^t_{\mu k}\right)\left[\frac{\ell-1}{\ell}+\sum_{j=\ell+1}^{\ell'-1}\frac{1}{j(j-1)}\right]\right. \\
&\quad \left.+\frac{\mathcal{M}_1\left(m^t_{\mu k}\right)}{\ell'}+\frac{\mathcal{M}_2\left(m^t_{\mu k}\right)}{\ell'}\left(\ell'-1+\frac{1}{\ell'-1}\right)-\mathcal{M}_0\left(m^t_{\mu k}\right)\sum_{j=\ell'+1}^{L}\frac{1}{j(j-1)}\right\} \\
&= -\frac{g^t_{2\mu k}\xi\left(m^t_{\mu k},g^t_{2\mu k}\right)}{n}\left[\mathcal{M}_1\left(m^t_{\mu k}\right)-\mathcal{M}_0\left(m^t_{\mu k}\right)\right] \\
&\quad +\frac{G^t_{\mu k}\xi\left(m^t_{\mu k},G^t_{\mu k}\right)-g^t_{2\mu k}\xi\left(m^t_{\mu k},g^t_{2\mu k}\right)}{nL}\mathcal{M}_0\left(m^t_{\mu k}\right)\,.
\end{aligned}
$$

In this manner

$$
\begin{aligned}
\left\langle b^{\ell \mathrm{a}}_k b^{\ell'\mathrm{a}}_k|h^t_{\mu k}\to 0,g^t_{\mu k},\Delta g^t_{\mu k}\right\rangle &\simeq \left(m^t_{\mu k}\right)^2-2\frac{g^t_{2\mu k}\xi\left(m^t_{\mu k},g^t_{2\mu k}\right)}{n}\left(m^t_{\mu k}\right)^2\left[1-\left(m^t_{\mu k}\right)^2\right] \\
&\quad +\frac{G^t_{\mu k}\xi\left(m^t_{\mu k},G^t_{\mu k}\right)-g^t_{2\mu k}\xi\left(m^t_{\mu k},g^t_{2\mu k}\right)}{nL}\left[1-3\left(m^t_{\mu k}\right)^2\right]\left[1-\left(m^t_{\mu k}\right)^2\right] \\
&\quad +2\xi\left(m^t_{\mu k},G^t_{\mu k}\right)\left[1-\left(m^t_{\mu k}\right)^2\right]m^t_{\mu k}h^t_{\mu k}\,. \qquad (39)
\end{aligned}
$$

Remembering that $\left\langle b^{\ell \mathrm{a}}_k b^{\ell \mathrm{a}}_k|h^t_{\mu k},g^t_{1\mu k},g^t_{2\mu k}\right\rangle=1$ and using Eqs. (37), (38) and (39), the covariance matrix entries can be then calculated:

$$
\begin{aligned}
\left(\mathbf{\Psi}^t_{\mu kl}\right)^{\ell \mathrm{a}\,\ell'\mathrm{a}'} &= \left\langle b^{\ell \mathrm{a}}_k b^{\ell'\mathrm{a}'}_l|h^t_{\mu k}\to 0,g^t_{1\mu k},g^t_{2\mu k};h^t_{\mu l}\to 0,g^t_{1\mu l},g^t_{2\mu l}\right\rangle \\
&\quad -\left\langle b^{\ell \mathrm{a}}_k|h^t_{\mu k}\to 0,g^t_{1\mu k},g^t_{2\mu k}\right\rangle\left\langle b^{\ell'\mathrm{a}'}_l|h^t_{\mu l}\to 0,g^t_{1\mu l},g^t_{2\mu l}\right\rangle=\delta_{kl}\left(\mathbf{\Psi}^t_{\mu kk}\right)^{\ell \mathrm{a}\,\ell'\mathrm{a}'} \\
\left(\mathbf{\Psi}^t_{\mu kk}\right)^{\ell \mathrm{a}\,\ell'\mathrm{a}'} &\simeq \delta^{\ell\ell'}\delta^{\mathrm{a}\mathrm{a}'}\left[1-\left(m^t_{\mu k}\right)^2\right]+\delta^{\ell\ell'}\left(1-\delta^{\mathrm{a}\mathrm{a}'}\right)\frac{g^t_{2\mu k}\xi\left(m^t_{\mu k},g^t_{2\mu k}\right)}{n}\left[1-\left(m^t_{\mu k}\right)^2\right]^2 \\
&\quad +\left(1-\delta^{\ell\ell'}\right)\frac{G^t_{\mu k}\xi\left(m^t_{\mu k},G^t_{\mu k}\right)-g^t_{2\mu k}\xi\left(m^t_{\mu k},g^t_{2\mu k}\right)}{nL}\left[1-\left(m^t_{\mu k}\right)^2\right]^2\,,
\end{aligned}
$$

where we have kept only the dominant terms at each entry, disregarding terms of order $\mathcal{O}\left(\mathrm{e}^{-2nLm^t_{\mu k}h^t_{\mu k}}+\frac{h}{n}+\frac{h}{L}\right)$.

If the $\varepsilon_{\mu k}$ and b^{a}_k are unbiased variables, the variable $\Delta^{\mathrm{a}}_{\mu k}=\sum_{l\neq k}\varepsilon_{\mu l}b^{\mathrm{a}}_l$, by virtue of the central limit theorem, obeys a normal distribution, with mean value and covariance matrix given by:

$$\begin{aligned}
\left(\mathbf{u}^t_{\mu k}\right)^{\ell \mathrm{a}} &\equiv \left\langle \Delta^{\ell \mathrm{a}}_{\mu k}\right\rangle = \sum_{\left\{\mathbf{b}_{l\neq k}\right\}} \prod_{l\neq k} P^t\left(\mathbf{b}_l \middle| \left\{y_{\nu\neq\mu}\right\}\right) \sum_{l\neq k} \varepsilon_{\mu l} b^{\ell \mathrm{a}}_l = \sum_{l\neq k} \varepsilon_{\mu l} m^t_{\mu l} \\
\left(\boldsymbol{\Upsilon}^t_{\mu k}\right)^{\ell \mathrm{a}\, \ell' \mathrm{a}'} &\equiv \left\langle \Delta^{\ell \mathrm{a}}_{\mu k} \Delta^{\ell' \mathrm{a}'}_{\mu k}\right\rangle - \left\langle \Delta^{\ell \mathrm{a}}_{\mu k}\right\rangle \left\langle \Delta^{\ell' \mathrm{a}'}_{\mu k}\right\rangle \\
&= \sum_{\left\{\mathbf{b}_{l\neq k}\right\}} \prod_{l\neq k} P^t\left(\mathbf{b}_l \middle| \left\{y_{\nu\neq\mu}\right\}\right) \sum_{\substack{l\neq k\\ j\neq k}} \varepsilon_{\mu l}\varepsilon_{\mu j} b^{\ell \mathrm{a}}_l b^{\ell' \mathrm{a}'}_j - \left(\sum_{l\neq k} \varepsilon_{\mu l} m^t_{\mu l}\right)^2 \\
&= \sum_{l\neq k} \varepsilon^2_{\mu l} \left(\boldsymbol{\Psi}^t_{\mu l j}\right)^{\ell \mathrm{a}\, \ell' \mathrm{a}'} = \delta^{\ell\ell'}\delta^{\mathrm{a}\mathrm{a}'} X^t_{\mu k} + \delta^{\ell\ell'}\left(1-\delta^{\mathrm{a}\mathrm{a}'}\right) \frac{1}{n} V^t_{\mu k} + \left(1-\delta^{\ell\ell'}\right) \frac{1}{nL}\left(R^t_{\mu k} - V^t_{\mu k}\right),
\end{aligned}$$

where $X^t_{\mu k}$ is given by Eqs. (29) and

$$\begin{aligned}
V^t_{\mu k} &\equiv \sum_{l\neq k} \varepsilon^2_{\mu l} g^t_{2\mu l} \xi\left(m^t_{\mu l}, g^t_{2\mu l}\right)\left[1-\left(m^t_{\mu l}\right)^2\right]^2 \\
R^t_{\mu k} &\equiv \sum_{l\neq k} \varepsilon^2_{\mu l} G^t_{\mu l} \xi\left(m^t_{\mu l}, G^t_{\mu l}\right)\left[1-\left(m^t_{\mu l}\right)^2\right]^2
\end{aligned}$$

are macroscopic variables of $\mathcal{O}(1)$. In particular, $R^t_{\mu k}$ and $V^t_{\mu k}$ are free variables that can be used to optimise a given performance measure. This variables have the property of being self averaging, therefore we can drop the sub-indexes μ and k. The eigenvalues and eigenvectors of the matrix $\boldsymbol{\Upsilon}^t$ satisfy the following relation:

$$X^t v^{\ell \mathrm{a}} - \frac{1}{n} V^t \sum_{\mathrm{b}\neq \mathrm{a}} v^{\ell \mathrm{b}} + \frac{1}{nL}\left(R^t - V^t\right) \sum_{\ell'\neq \ell} \sum_{\mathrm{b}=1}^{n} v^{\ell \mathrm{b}} = \lambda v^{\ell \mathrm{a}}.$$

If $\sum_{\mathrm{b}=1}^{n} v^{\ell \mathrm{b}}$ is independent of ℓ, then for two different entries of the eigenvector we have that:

$$\left(X^t - \frac{1}{n} V^t\right)\left(v^{\ell \mathrm{a}} - v^{\ell' \mathrm{a}'}\right) = \lambda\left(v^{\ell \mathrm{a}} - v^{\ell' \mathrm{a}'}\right),$$

which implies that either $v^{\ell \mathrm{a}} = v^{\ell' \mathrm{a}'}$ $\forall \ell$, ℓ', a, and a$'$ or $\lambda_1 = X^t - \frac{1}{n} V^t$. The first option leads to the eigenvalue $\lambda_0 \simeq X^t + R^t + \mathcal{O}\left(n^{-1}\right)$ and $\mathbf{v}^{\mathsf{T}}_0 = \mathbf{v}^{\mathsf{T}}_{1,1} = \left(\overbrace{1,\ldots,1}^{nL}\right)/\sqrt{nL}$. λ_1 is $(nL-1)$-fold degenerated, and their eigenvectors are:

$$\mathbf{v}^{\mathsf{T}}_{1,\mathrm{a}>1} = \frac{1}{\sqrt{L}}\left(\overbrace{\mathbf{w}^{\mathsf{T}}_{\mathrm{a}},\ldots,\mathbf{w}^{\mathsf{T}}_{\mathrm{a}}}^{L}\right) \tag{40}$$

$$\mathbf{v}^{\mathsf{T}}_{\ell>1,\mathrm{a}} = \frac{1}{\sqrt{\ell\left(\ell-1\right)}}\left(\overbrace{\mathbf{w}^{\mathsf{T}}_{\mathrm{a}},\ldots,\mathbf{w}^{\mathsf{T}}_{\mathrm{a}}}^{\ell-1}, -\left(\ell-1\right)\mathbf{w}^{\mathsf{T}}_{\mathrm{a}}, \overbrace{\mathbf{0}^{\mathsf{T}},\ldots,\mathbf{0}^{\mathsf{T}}}^{L-\ell}\right) \tag{41}$$

where $\mathbf{0}^{\mathsf{T}} = \left(\overbrace{0,\ldots,0}^{n}\right)$ and

$$\begin{aligned}
\mathbf{w}^{\mathsf{T}}_1 &= \frac{1}{\sqrt{n}}\left(\overbrace{1,\ldots,1}^{n}\right) \\
\mathbf{w}^{\mathsf{T}}_{\mathrm{a}>1} &= \frac{1}{\sqrt{\mathrm{a}\left(\mathrm{a}-1\right)}}\left(\overbrace{1,1,\ldots,1}^{\mathrm{a}-1}, -\left(\mathrm{a}-1\right), \overbrace{0,0,\ldots,0}^{n-\mathrm{a}}\right)
\end{aligned}$$

are the eigenvectors of the RS matrix. Observe that the number of eigenvectors that satisfy Eq. (40) is $n-1$ (there are n different $\mathbf{w}_{\rm a}$), and the number of eigenvectors that satisfy Eq. (41) is $(L-1)\,n$ which gives us a total of $nL-1$ eigenvectors to the λ_1 eigenvalue.

Appendix C: The messages

From Eqs. (3) and (4) and with the application of Eq. (7) or Eq. (8) in Eq. (5) we can express the message from nodes y_μ to nodes $b_k^{\rm a}$ at time $t+1$ as:

$$\begin{aligned}\widehat{m}_{\mu k}^{t+1} &= \frac{\sum_{\{\mathbf{B}\}} b_k^{\rm a'} \prod_{\rm a=1}^{n} P\left(y_\mu|\mathbf{b}^{\rm a}\right) P\left(\mathbf{b}^{\rm a}\right) \prod_{l\neq k} P\left(\mathbf{b}_l|\left\{y_{\nu\neq\mu}\right\}\right)}{\sum_{\{\mathbf{B}\}} \prod_{\rm a=1}^{n} P\left(y_\mu|\mathbf{b}^{\rm a}\right) P\left(\mathbf{b}^{\rm a}\right) \prod_{l\neq k} P\left(\mathbf{b}_l|\left\{y_{\nu\neq\mu}\right\}\right)} \\ &= \frac{\int \mathrm{d}\boldsymbol{\Delta}_{\mu k} \sum_{\{\mathbf{b}_k\}} b_k^{\rm a'}\, P\left(\boldsymbol{\Delta}_{\mu k}|\mathbf{B}\right) P\left(y_\mu|\boldsymbol{\Delta}_{\mu k};\boldsymbol{\gamma}\right)\left[1+\varepsilon_{\mu k}\mathbf{b}_k^{\mathsf{T}}\nabla_{\boldsymbol{\Delta}_{\mu k}} \ln P\left(y_\mu|\boldsymbol{\Delta}_{\mu k};\boldsymbol{\gamma}\right)\right]}{\int \mathrm{d}\boldsymbol{\Delta}_{\mu k} \sum_{\{\mathbf{b}_k\}} P\left(\boldsymbol{\Delta}_{\mu k}|\mathbf{B}\right) P\left(y_\mu|\boldsymbol{\Delta}_{\mu k};\boldsymbol{\gamma}\right)\left[1+\varepsilon_{\mu k}\mathbf{b}_k^{\mathsf{T}}\nabla_{\boldsymbol{\Delta}_{\mu k}} \ln P\left(y_\mu|\boldsymbol{\Delta}_{\mu k};\boldsymbol{\gamma}\right)\right]}\,.\end{aligned}$$

If $P\left(y_\mu|\boldsymbol{\Delta}_{\mu k};\boldsymbol{\gamma}\right)=\prod_{\rm a=1}^{n} P\left(y_\mu|\Delta_{\mu k}^{\rm a};\boldsymbol{\gamma}\right)$ and desregarding $\mathcal{O}\left(\varepsilon_{\mu k}^2\right)$, the traces on $\mathbf{b}_k$ can be written as

$$\begin{aligned}\sum_{\{\mathbf{b}_k\}} P\left(y_\mu|\boldsymbol{\Delta}_{\mu k};\boldsymbol{\gamma}\right)\left[1+\varepsilon_{\mu k}\mathbf{b}_k^{\mathsf{T}}\nabla_{\boldsymbol{\Delta}_{\mu k}} \ln P\left(y_\mu|\boldsymbol{\Delta}_{\mu k};\boldsymbol{\gamma}\right)\right] &= 2^n P\left(y_\mu|\boldsymbol{\Delta}_{\mu k};\boldsymbol{\gamma}\right) \\ \sum_{\{\mathbf{b}_k\}} b_k^{\rm a'}\, P\left(y_\mu|\boldsymbol{\Delta}_{\mu k};\boldsymbol{\gamma}\right)\left[1+\varepsilon_{\mu k}\mathbf{b}_k^{\mathsf{T}}\nabla_{\boldsymbol{\Delta}_{\mu k}} \ln P\left(y_\mu|\boldsymbol{\Delta}_{\mu k};\boldsymbol{\gamma}\right)\right] &= 2^n \varepsilon_{\mu k} P\left(y_\mu|\boldsymbol{\Delta}_{\mu k};\boldsymbol{\gamma}\right) \frac{\partial}{\partial \Delta_{\mu k}^{\rm a'}} \ln P\left(y_\mu|\Delta_{\mu k}^{\tilde{\rm a}};\boldsymbol{\gamma}\right)\,,\end{aligned}$$

thus

$$\begin{aligned}{}^{(\mathrm{RS})}\widehat{m}_{\mu k}^{t+1} &\simeq \varepsilon_{\mu k} \frac{\int \mathrm{d}\boldsymbol{\Delta}_{\mu k} P\left(\boldsymbol{\Delta}_{\mu k}|\mathbf{B}\right) \prod_{\rm a=1}^{n} P\left(y_\mu|\Delta_{\mu k}^{\rm a};\boldsymbol{\gamma}\right) \frac{\partial}{\partial \Delta_{\mu k}^{\rm a'}} \ln P\left(y_\mu|\Delta_{\mu k}^{\rm a'};\boldsymbol{\gamma}\right)}{\int \mathrm{d}\boldsymbol{\Delta}_{\mu k} P\left(\boldsymbol{\Delta}_{\mu k}|\mathbf{B}\right) \prod_{\rm a=1}^{n} P\left(y_\mu|\Delta_{\mu k}^{\rm a};\boldsymbol{\gamma}\right)} \qquad (42) \\ &= \frac{\varepsilon_{\mu k}}{{}^{(\mathrm{RS})}\mathcal{N}_{\mu k}^{t}} \int \mathrm{d}\vartheta\, \exp\left\{-n\,\frac{\left(\vartheta-u_{\mu k}^{t}\right)^2}{2R^t}\right\} \\ &\quad \times\left[\int \mathrm{d}\Delta \exp\left\{-\frac{\left(\Delta-\vartheta\right)^2}{2X^t}+\ln P\left(y_\mu|\Delta;\boldsymbol{\gamma}\right)\right\}\right]^{n-1} \\ &\quad \times \int \mathrm{d}\Delta \exp\left\{-\frac{\left(\Delta-\vartheta\right)^2}{2X^t}\right\} \frac{\partial}{\partial \Delta} P\left(y_\mu|\Delta;\boldsymbol{\gamma}\right)\,,\end{aligned}$$

and

$$
\begin{aligned}
{}^{(1\mathrm{RSB})}\widehat{m}^{t+1}_{\mu k} &\simeq \varepsilon_{\mu k}\frac{\int \mathrm{d}\boldsymbol{\Delta}_{\mu k} P\left(\boldsymbol{\Delta}_{\mu k}|\mathbf{B}\right)\prod_{\ell=1}^{L}\prod_{\mathrm{a}=1}^{n} P\left(y_\mu|\Delta^{\ell\mathrm{a}}_{\mu k};\boldsymbol{\gamma}\right)\frac{\partial}{\partial\Delta^{\ell'\mathrm{a}'}_{\mu k}}\ln P\left(y_\mu|\Delta^{\ell'\mathrm{a}'}_{\mu k};\boldsymbol{\gamma}\right)}{\int \mathrm{d}\boldsymbol{\Delta}_{\mu k} P\left(\boldsymbol{\Delta}_{\mu k}|\mathbf{B}\right)\prod_{\mathrm{a}=1}^{n} P\left(y_\mu|\Delta^{\mathrm{a}}_{\mu k};\boldsymbol{\gamma}\right)} \\
&= \frac{\varepsilon_{\mu k}}{{}^{(1\mathrm{RSB})}\mathscr{N}^{t}_{\mu k}}\int \mathrm{d}\boldsymbol{\Theta}\,\exp\left\{-\frac{n}{2}\sum_{\ell=1}^{L}\left(\vartheta^0,\vartheta^\ell\right)\mathbf{S}^t\begin{pmatrix}\vartheta^0\\ \vartheta^\ell\end{pmatrix}\right\} \\
&\quad\times\prod_{\ell\neq\ell'}\left[\left[\int \mathrm{d}\Delta\exp\left\{-\frac{\left(\Delta-\vartheta^{0\ell t}_{\mu k}\right)^2}{2X^t}+\ln P\left(y_\mu|\Delta;\boldsymbol{\gamma}\right)\right\}\right]\right]^n \\
&\quad\times\left[\int \mathrm{d}\Delta\exp\left\{-\frac{\left(\Delta-\vartheta^{0\ell t}_{\mu k}\right)^2}{2X^t}+\ln P\left(y_\mu|\Delta;\boldsymbol{\gamma}\right)\right\}\right]^{n-1} \\
&\quad\times\int \mathrm{d}\Delta\exp\left\{-\frac{\left(\Delta-\vartheta^{0\ell t}_{\mu k}\right)^2}{2X^t}\right\}\frac{\partial}{\partial\Delta}P\left(y_\mu|\Delta;\boldsymbol{\gamma}\right)\,,
\end{aligned}
$$

where ${}^{(\mathrm{RS})}\mathscr{N}^t_{\mu k}$ and ${}^{(1\mathrm{RSB})}\mathscr{N}^t_{\mu k}$ are suitable normalisation constants, $\boldsymbol{\Theta}^{\mathsf{T}}=\left(\vartheta^0,\ldots,\vartheta^\ell,\ldots,\vartheta^L\right)$ and $\vartheta^{0\ell t}_{\mu k}\equiv\vartheta^0+\vartheta^\ell+u^t_{\mu k}$. Let us define the following variable $z=\left(\Delta-\vartheta\right)/\sqrt{2X^t}$, where ϑ stands for the RS or 1RSB $\left(\vartheta^{0\ell t}_{\mu k}\right)$ deffinition, according to the case. We can define:

$$
\begin{aligned}
\mathcal{G}\left(\vartheta,y_\mu\right) &\equiv \int \mathrm{d}\Delta\exp\left\{-\frac{\left(\Delta-\vartheta\right)^2}{2X^t}\right\}P\left(y_\mu|\Delta;\boldsymbol{\gamma}\right) && (43)\\
\mathcal{P}\left(\vartheta,y_\mu\right) &\equiv \left[\mathcal{G}\left(\vartheta,y_\mu\right)\right]^{-1}\int \mathrm{d}\Delta\exp\left\{-\frac{\left(\Delta-\vartheta\right)^2}{2X^t}\right\}\frac{\partial}{\partial\Delta}P\left(y_\mu|\Delta;\boldsymbol{\gamma}\right) \\
&= \left[\mathcal{G}\left(\vartheta,y_\mu\right)\right]^{-1}\int \mathrm{d}\Delta\exp\left\{-\frac{\left(\Delta-\vartheta\right)^2}{2X^t}\right\}\frac{\Delta-\vartheta}{X^t}P\left(y_\mu|\Delta;\boldsymbol{\gamma}\right) \\
&= \frac{\partial}{\partial\vartheta}\ln\mathcal{G}\left(\vartheta,y_\mu\right) && (44)\\
{}^{(\mathrm{RS})}\mathcal{H}\left(\vartheta,y_\mu\right) &\equiv \frac{\left(\vartheta-u^t_{\mu k}\right)^2}{2R^t_{\mu k}}-\ln\mathcal{G}\left(\vartheta,y_\mu\right) && (45)\\
{}^{(1\mathrm{RSB})}\mathcal{H}\left(\vartheta^0,\vartheta^\ell,y_\mu\right) &\equiv \frac{1}{2}\left(\vartheta^0,\vartheta^\ell\right)\mathbf{S}^t\begin{pmatrix}\vartheta^0\\ \vartheta^\ell\end{pmatrix}-\ln\mathcal{G}\left(\vartheta^{0\ell t}_{\mu k},y_\mu\right)\,, && (46)
\end{aligned}
$$

where $\mathbf{S}^t=\frac{1}{X^t}\begin{pmatrix}\frac{X^t\left(X^t+V^t\right)+V^t\left(X^t+R^t\right)}{X^t\left(R^t-V^t\right)} & -1\\ -1 & \frac{X^t}{V^t}\end{pmatrix}$. Thus the expression for the RS message is:

$$
{}^{(\mathrm{RS})}\widehat{m}^{t+1}_{\mu k} = \varepsilon_{\mu k}\frac{\int \mathrm{d}\vartheta\,\exp\left\{-n^{(\mathrm{RS})}\mathcal{H}\left(\vartheta,y_\mu\right)\right\}\mathcal{P}\left(\vartheta,y_\mu\right)}{\int \mathrm{d}\vartheta\,\exp\left\{-n^{(\mathrm{RS})}\mathcal{H}\left(\vartheta,y_\mu\right)\right\}}\,.
$$

In the large n limit, only the solutions $\tilde{\vartheta}^t_{\mu k}$ of $\frac{\partial}{\partial\vartheta}{}^{(\mathrm{RS})}\mathcal{H}=0$ corresponding to the lowest minimum contributes to the integral. The ground states of the Hamiltonians are calculated in Sec. VII. Using this information the final expression for the message is:

$$
{}^{(\mathrm{RS})}\widehat{m}^{t+1}_{\mu k}=\varepsilon_{\mu k}\frac{\tilde{\vartheta}^t_{\mu k}-u^t_{\mu k}}{R^t}\,, \tag{47}
$$

where $\tilde{\vartheta}^t_{\mu k}$ is given by Eq. (50).

The 1RSB case is a little more delicate. The effective Hamiltonian that appears as the argument of the exponential is a sum over L Hamiltonians ${}^{(1\text{RSB})}\mathcal{H}\left(\vartheta^0, \vartheta^\ell, y_\mu\right)$. Therefore, the Taylor expansion of the total Hamiltonian near the ground state Eq. (51) is:

$$\sum_{\ell=1}^{L} {}^{(1\text{RSB})}\mathcal{H}\left(\vartheta^0, \vartheta^\ell, y_\mu\right) \simeq LE_0 + \frac{L}{2} h_0 \left(\Delta\vartheta^0\right)^2 + h_1 \Delta\vartheta^0 \sum_{\ell=1}^{L} \Delta\vartheta^\ell + \frac{1}{2} h_2 \sum_{\ell=1}^{L} \left(\Delta\vartheta^\ell\right)^2 + \mathcal{O}\left(\Delta\vartheta^3\right),$$

where $E_0 = {}^{(1\text{RSB})}\mathcal{H}\left(\tilde{\vartheta}^{0t}_{\mu k}, \tilde{\vartheta}^{\ell t}_{\mu k}, y_\mu\right)$ is the energy of the ground state, $\Delta\vartheta^i = \vartheta^i - \tilde{\vartheta}^{it}_{\mu k}$ $i = 0, \ell$ and the entries h_0, h_1 and h_2 satisfy the equation

$$\begin{pmatrix} h_0 & h_1 \\ h_1 & h_2 \end{pmatrix} = \mathbf{S}^t - \left.\frac{\partial \mathcal{P}}{\partial \vartheta}\right|_{\vartheta = \tilde{\vartheta}^t_{\mu k}} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix},$$

where $\tilde{\vartheta}^{0\ell t}_{\mu k}$ is the solution of Eq. (51). If $\boldsymbol{\Theta}^\mathsf{T} = \left(\vartheta^0, \vartheta^1, \ldots, \vartheta^L\right)$ and $\left(\mathbf{H}_\mathcal{H}\right)_{ij} = \delta_{jk}\left[\delta_{j0} h_0 + \left(1 - \delta_{j0}\right) L^{-1} h_2\right] + \left(\delta_{j0} + \delta_{k0}\right)\left(1 - \delta_{jk}\right) L^{-1} h_1$ is the Hessian of $\sum_{\ell=1}^{L} {}^{(1\text{RSB})}\mathcal{H}\left(\vartheta^0, \vartheta^\ell, y_\mu\right)$, then

$$\sum_{\ell=1}^{L} {}^{(1\text{RSB})}\mathcal{H}\left(\vartheta^0, \vartheta^\ell, y_\mu\right) \simeq LE_0 + \frac{L}{2} \Delta\boldsymbol{\Theta}^\mathsf{T} \mathbf{H}_\mathcal{H} \Delta\boldsymbol{\Theta}.$$

The matrix $\mathbf{H}_\mathcal{H}$ has the same structure as $\mathbf{H}_\Phi$, therefore, the eigenvalues and eigenvectors of $\mathbf{H}_\mathcal{H}$ can be obtained adapting Eqs. (32) and (33) by the substitutions $\alpha_0 = h_0$, $-\beta_0 = h_1$ and $\gamma_0 = h_2$. Expanding $\mathcal{P}\left(\vartheta, y_\mu\right)$ at the ground state $\tilde{\vartheta}^{0\ell t}_{\mu k}$ we obtained $\mathcal{P}\left(\vartheta^{0\ell' t}_{\mu k}, y_\mu\right) \simeq \mathcal{P}_0 + \mathcal{P}_1\left(\Delta\vartheta^0 + \Delta\vartheta^{\ell'}\right) + \frac{1}{2}\mathcal{P}_2\left(\Delta\vartheta^0 + \Delta\vartheta^{\ell'}\right)^2$ where $\mathcal{P}_j \equiv \left.\frac{\partial^j \mathcal{P}}{\partial \vartheta^j}\right|_{\vartheta = \tilde{\vartheta}^t_{\mu k}}$. Thus

$${}^{(1\text{RSB})}\widehat{m}^{t+1}_{\mu k} = \varepsilon_{\mu k} \frac{\displaystyle\int \mathrm{d}\boldsymbol{\Theta} \exp\left\{-\frac{nL}{2}\Delta\boldsymbol{\Theta}^\mathsf{T}\mathbf{H}_\mathcal{H}\Delta\boldsymbol{\Theta}\right\}\left[\mathcal{P}_0 + \mathcal{P}_1\left(\Delta\vartheta^0 + \Delta\vartheta^{\ell'}\right) + \frac{1}{2}\mathcal{P}_2\left(\Delta\vartheta^0 + \Delta\vartheta^{\ell'}\right)^2\right]}{\displaystyle\int \mathrm{d}\boldsymbol{\Theta} \exp\left\{-\frac{nL}{2}\Delta\boldsymbol{\Theta}^\mathsf{T}\mathbf{H}_\mathcal{H}\Delta\boldsymbol{\Theta}\right\}}$$

where the term proportional to $\mathcal{P}_1$ vanishes for parity reasons. In the basis of eigenvectors of $\mathbf{H}_\mathcal{H}$, i.e. $\boldsymbol{\Gamma} = \mathbf{U}^\mathsf{T}\boldsymbol{\Theta} = \left(\gamma_0, \gamma_1, \ldots, \gamma_L\right)^\mathsf{T}$ where $\mathbf{U}$ is adapted from Eq. (34), the message has the form:

$${}^{(1\text{RSB})}\widehat{m}^{t+1}_{\mu k} \simeq \varepsilon_{\mu k} \frac{\displaystyle\int \mathrm{d}\boldsymbol{\Gamma} \exp\left\{-\frac{nL}{2}\sum_{\ell=0}^{L}\lambda_\ell\left(\Delta\gamma^\ell\right)^2\right\}\left(\mathcal{P}_0 + \tfrac{1}{2}\mathcal{P}_2 \Delta\boldsymbol{\Gamma}^\mathsf{T}\mathbf{M}'_{0\ell'}\Delta\boldsymbol{\Gamma}\right)}{\displaystyle\int \mathrm{d}\boldsymbol{\Gamma} \exp\left\{-\frac{nL}{2}\sum_{\ell=0}^{L}\lambda_\ell\left(\Delta\gamma^\ell\right)^2\right\}},$$

where λ_ℓ are the eigenvalues of $\mathbf{H}_\mathcal{H}$ and $\mathbf{M}'_{0\ell'}$ is adapted from Eq. (36).

The expression for the message is reduced to

$$\begin{aligned} {}^{(1\text{RSB})}\widehat{m}^{t+1}_{\mu k} &\simeq \varepsilon_{\mu k}\left[\mathcal{P}_0 + \frac{1}{n}\frac{\mathcal{P}_2}{2h_2} + \mathcal{O}\left(\frac{1}{nL}\right)\right] \\ &\simeq \varepsilon_{\mu k}\frac{\tilde{\vartheta}^t_{\mu k} - u^t_{\mu k}}{R^t} + \frac{\varepsilon_{\mu k}}{2n}\frac{\mathcal{P}_2 V^t}{1 - \mathcal{P}_1 V^t}. \end{aligned} \tag{48}$$

The expression for the messages from **b**-nodes to **y**-nodes is:

$$
\begin{aligned}
m_{\mu k}^{t} &= \sum_{\{\mathbf{b}_k\}} b_k^{\mathrm{a}'} P^{t}\left(\mathbf{b}_k | \{y_{\nu\neq\mu}\}\right) \\
&= \frac{\sum_{\{\mathbf{b}_k\}} b_k^{\mathrm{a}'} \prod_{\nu\neq\mu} \sum_{\{\mathbf{b}_{l\neq k}\}} P\left(y_\nu|\mathbf{B}\right) \prod_{l\neq k} P^{t-1}\left(\mathbf{b}_l | \{y_{\sigma\neq\nu}\}\right)}{\sum_{\{\mathbf{b}_k\}} \prod_{\nu\neq\mu} \sum_{\{\mathbf{b}_{l\neq k}\}} P\left(y_\nu|\mathbf{B}\right) \prod_{l\neq k} P^{t-1}\left(\mathbf{b}_l | \{y_{\sigma\neq\nu}\}\right)} ,
\end{aligned}
$$

which can be approximated by

$$
\begin{aligned}
m_{\mu k}^{t} &\simeq \frac{\sum_{\{\mathbf{b}_k\}} b_k^{\mathrm{a}'} \prod_{\nu\neq\mu} \int \mathrm{d}\boldsymbol{\Delta}_{\nu k} P\left(\boldsymbol{\Delta}_{\nu k}|\mathbf{B}\right) \prod_{\mathrm{a}=1}^{n} P\left(y_\nu|\Delta_{\nu k}^{\mathrm{a}};\boldsymbol{\gamma}\right)\left[1+\varepsilon_{\nu k} b_k^{\mathrm{a}} \frac{\partial}{\partial \Delta_{\nu k}^{\mathrm{a}}} \ln P\left(y_\nu|\boldsymbol{\Delta}_{\nu k};\boldsymbol{\gamma}\right)\right]}{\sum_{\{\mathbf{b}_k\}} \prod_{\nu\neq\mu} \int \mathrm{d}\boldsymbol{\Delta}_{\nu k} P\left(\boldsymbol{\Delta}_{\nu k}|\mathbf{B}\right) \prod_{\mathrm{a}=1}^{n} P\left(y_\nu|\Delta_{\nu k}^{\mathrm{a}};\boldsymbol{\gamma}\right)\left[1+\varepsilon_{\nu k} b_k^{\mathrm{a}} \frac{\partial}{\partial \Delta_{\nu k}^{\mathrm{a}}} \ln P\left(y_\nu|\boldsymbol{\Delta}_{\nu k};\boldsymbol{\gamma}\right)\right]} \\
&= \frac{\sum_{b_k^{\mathrm{a}'}=\pm 1} b_k^{\mathrm{a}'} \prod_{\nu\neq\mu} \int \mathrm{d}\boldsymbol{\Delta}_{\nu k} P\left(\boldsymbol{\Delta}_{\nu k}|\mathbf{B}\right) P\left(y_\nu|\boldsymbol{\Delta}_{\nu k};\boldsymbol{\gamma}\right)\left[1+\varepsilon_{\nu k} b_k^{\mathrm{a}'} \frac{\partial}{\partial \Delta_{\nu k}^{\mathrm{a}'}} \ln P\left(y_\nu|\boldsymbol{\Delta}_{\nu k};\boldsymbol{\gamma}\right)\right]}{\sum_{b_k^{\mathrm{a}'}=\pm 1} \prod_{\nu\neq\mu} \int \mathrm{d}\boldsymbol{\Delta}_{\nu k} P\left(\boldsymbol{\Delta}_{\nu k}|\mathbf{B}\right) P\left(y_\nu|\boldsymbol{\Delta}_{\nu k};\boldsymbol{\gamma}\right)\left[1+\varepsilon_{\nu k} b_k^{\mathrm{a}'} \frac{\partial}{\partial \Delta_{\nu k}^{\mathrm{a}'}} \ln P\left(y_\nu|\boldsymbol{\Delta}_{\nu k};\boldsymbol{\gamma}\right)\right]} \\
&= \frac{\sum_{b_k^{\mathrm{a}}=\pm 1} b_k^{\mathrm{a}} \prod_{\nu\neq\mu} \frac{1+\widehat{m}_{\nu k}^{t} b_k^{\mathrm{a}}}{\mathcal{N}_{\nu k}^{t}}}{\sum_{b_k^{\mathrm{a}}=\pm 1} \prod_{\nu\neq\mu} \frac{1+\widehat{m}_{\nu k}^{t} b_k^{\mathrm{a}}}{\mathcal{N}_{\nu k}^{t}}} = \frac{\prod_{\nu\neq\mu} \frac{1+\widehat{m}_{\nu k}^{t}}{\mathcal{N}_{\nu k}^{t}} - \prod_{\nu\neq\mu} \frac{1-\widehat{m}_{\nu k}^{t}}{\mathcal{N}_{\nu k}^{t}}}{\prod_{\nu\neq\mu} \frac{1+\widehat{m}_{\nu k}^{t}}{\mathcal{N}_{\nu k}^{t}} + \prod_{\nu\neq\mu} \frac{1-\widehat{m}_{\nu k}^{t}}{\mathcal{N}_{\nu k}^{t}}} = \tanh\left(\sum_{\nu\neq\mu} \operatorname{arctanh}\left(\widehat{m}_{\nu k}^{t}\right)\right) ,
\end{aligned}
$$

but being $\widehat{m}_{\nu k}^{t} \sim \mathcal{O}\left(\varepsilon_{\nu k}\right)$ we have that

$$
m_{\mu k}^{t} \simeq \tanh\left(\sum_{\nu\neq\mu} \widehat{m}_{\nu k}^{t}\right) . \tag{49}
$$

Appendix D: The ground state of $\mathcal{H}$

For the RS case the equation to be solved is:

$$
\begin{aligned}
\frac{\partial}{\partial\vartheta}{}^{(\mathrm{RS})}\mathcal{H}\left(\vartheta, y_\mu\right) &= \frac{\vartheta - u_{\mu k}^{t}}{R_{\mu k}^{t}} - \frac{\partial}{\partial\vartheta} \ln \mathcal{G}\left(\vartheta, y_\mu\right) \\
&= \frac{\vartheta - u_{\mu k}^{t}}{R_{\mu k}^{t}} - \mathcal{P}\left(\vartheta, y_\mu\right) ,
\end{aligned}
$$

thus, the equation to be satisfied is:

$$
\tilde{\vartheta}_{\mu k}^{t} = u_{\mu k}^{t} + R^{t} \mathcal{P}\left(\tilde{\vartheta}_{\mu k}^{t}, y_\mu\right) . \tag{50}
$$

For the 1RSB case we have that $\frac{\partial}{\partial\vartheta^{0}}{}^{(\mathrm{1RSB})}\mathcal{H} = \frac{\partial}{\partial\vartheta^{\ell}}{}^{(\mathrm{1RSB})}\mathcal{H} = 0$, therefore:

$$
\begin{aligned}
0 &= \frac{X^{t}\left(X^{t}+V^{t}\right)+V^{t}\left(X^{t}+R^{t}\right)}{X^{t}\left(R^{t}-V^{t}\right)} \tilde{\vartheta}_{\mu k}^{0t} - \tilde{\vartheta}_{\mu k}^{\ell t} - X^{t} \mathcal{P}\left(\tilde{\vartheta}_{\mu k}^{0\ell t}, y_\mu\right) \\
0 &= -\tilde{\vartheta}_{\mu k}^{0t} + \frac{X^{t}}{V^{t}} \tilde{\vartheta}_{\mu k}^{\ell t} - X^{t} \mathcal{P}\left(\tilde{\vartheta}_{\mu k}^{0\ell t}, y_\mu\right) ,
\end{aligned}
$$

which is equivalent to:

$$\tilde{\vartheta}^{0\ell t}_{\mu k} = u^t_{\mu k} + R^t \mathcal{P}\left(\tilde{\vartheta}^{0\ell t}_{\mu k}, y_\mu\right), \tag{51}$$

where $\vartheta^{0\ell t}_{\mu k} = \vartheta^0 + \vartheta^\ell + u^t_{\mu k}$. Observed that Eq. (51) is equal to Eq. (50) and that the ground state $\tilde{\vartheta}^t_{\mu k}$ is independent of the indices 0 and ℓ.

Appendix E: The optimisation condition

Our goal is to make our algorithm to return a better estimate of the message at each iteration. Applying a method equivalent to the EM algorithm [18] we expect to find a suitable set of parameters $\boldsymbol{\gamma}^c$ that increase as much as possible the drop in the error per bit rate. Our error function has the form:

$$\mathscr{E}^t(\boldsymbol{\gamma}) \equiv \lambda^2 P_b^t - M^t/\sqrt{N^t}, \tag{52}$$

where λ^2 is a positive constant.

Observe that

$$M^t - N^t = \frac{1}{\sqrt{2\pi F^t}} \int \mathrm{d}z \, \exp\left[-\frac{z^2 + (E^t)^2 + \ln\cosh(z)}{2F^t}\right] \tanh(z) \sinh\left(\frac{E^t - F^t}{F^t} z\right),$$

and observe that $\mathrm{sgn}\left[\tanh(z)\sinh\left(\frac{E^t - F^t}{F^t}z\right)\right] = \mathrm{sgn}\left(E^t - F^t\right) \ \forall z$. Therefore $\mathrm{sgn}\left(E^t - F^t\right) = \mathrm{sgn}\left(M^t - N^t\right)$.

The second term of the RHS of Eq. (52) is an implicit function of the parameters $\boldsymbol{\gamma}$ through E^t and F^t, therefore:

$$\begin{aligned}
\mathrm{d}\left(\frac{M^t}{\sqrt{N^t}}\right) &= \frac{\partial}{\partial E^t}\left(\frac{M^t}{\sqrt{N^t}}\right)\mathrm{d}E^t + \frac{\partial}{\partial F^t}\left(\frac{M^t}{\sqrt{N^t}}\right)\mathrm{d}F^t \\
\frac{\partial}{\partial \gamma_i}\left(\frac{M^t}{\sqrt{N^t}}\right) &= \frac{\partial}{\partial E^t}\left(\frac{M^t}{\sqrt{N^t}}\right)\frac{\partial E^t}{\partial \gamma_i} + \frac{\partial}{\partial F^t}\left(\frac{M^t}{\sqrt{N^t}}\right)\frac{\partial F^t}{\partial \gamma_i},
\end{aligned}$$

where the partial derivatives are:

$$\begin{aligned}
\frac{\partial}{\partial E^t}\left(\frac{M^t}{\sqrt{N^t}}\right) &= \left(N^t\right)^{-\frac{3}{2}} \int \mathcal{D}z \left[1 - \tanh^2\left(\sqrt{F^t}z + E^t\right)\right]\left[N^t - M^t \tanh\left(\sqrt{F^t}z + E^t\right)\right] \\
\frac{\partial}{\partial F^t}\left(\frac{M^t}{\sqrt{N^t}}\right) &= \left(N^t\right)^{-\frac{3}{2}} \int \mathcal{D}z \frac{z}{2\sqrt{F^t}} \left[1 - \tanh^2\left(\sqrt{F^t}z + E^t\right)\right]\left[N^t - M^t \tanh\left(\sqrt{F^t}z + E^t\right)\right].
\end{aligned}$$

By the definition of the field $b_k h^t_{\mu k}$ we have that $\mathrm{sgn}\left(b_k h^t_{\mu k}\right) = \mathrm{sgn}\left(b_k m^t_{\mu k}\right) = \mathrm{sgn}\left(b_k m^t_k\right)$:

$$\begin{aligned}
P_b^t &\simeq \frac{1}{2K}\sum_{k=1}^{K}\left(1 - \mathrm{sgn}\left(b_k h^t_{\mu k}\right)\right) \\
&\sim \int_{-\infty}^{\infty} \frac{\mathrm{d}u}{\sqrt{2\pi F^t}} \exp\left\{-\frac{\left(u - E^t\right)^2}{2F^t}\right\} \frac{1}{2}\left(1 - \mathrm{sgn}(u)\right) \\
&= \int_{-\infty}^{-E^t/\sqrt{F^t}} \mathcal{D}u,
\end{aligned}$$

and we suppose that E^t and F^t are both explicit functions of the parameters $\boldsymbol{\gamma}$. Thus:

$$\frac{\partial P_b^t}{\partial \gamma_i} = -\frac{1}{\sqrt{2\pi F^t}} \exp\left[-\frac{(E^t)^2}{2F^t}\right] \left\{\frac{\partial E^t}{\partial \gamma_i} - \frac{1}{2}\frac{E^t}{F^t}\frac{\partial F^t}{\partial \gamma_i}\right\} .$$

By differentiation Eq. (52) we obtain:

$$\begin{aligned}
\frac{\partial}{\partial \gamma_i}\mathscr{E}^t &= -\frac{\lambda^2}{\sqrt{2\pi F^t}} \exp\left[-\frac{(E^t)^2}{2F^t}\right] \left(\frac{\partial E^t}{\partial \gamma_i} - \frac{1}{2}\frac{E^t}{F^t}\frac{\partial F^t}{\partial \gamma_i}\right) \\
&\quad - \left(N^t\right)^{-\frac{3}{2}} \int \mathcal{D}z \, \frac{N^t - M^t \tanh\left(\sqrt{F^t}z + E^t\right)}{\cosh^2\left(\sqrt{F^t}z + E^t\right)} \left(\frac{\partial E^t}{\partial \gamma_i} + \frac{z}{2\sqrt{F^t}}\frac{\partial F^t}{\partial \gamma_i}\right) \\
&= -\left(F^t N^t\right)^{-\frac{3}{2}} \int \frac{\mathrm{d}u}{\sqrt{2\pi}} \exp\left[-\frac{(u-E^t)^2}{2F^t}\right] u \, \frac{N^t - M^t \tanh(u)}{\cosh^2(u)} \\
&\quad - \left(\frac{\partial E^t}{\partial \gamma_i} - \frac{1}{2}\frac{E^t}{F^t}\frac{\partial F^t}{\partial \gamma_i}\right) \\
&\quad \times \left\{\frac{\lambda^2}{\sqrt{2\pi F^t}} \exp\left[\frac{(E^t)^2}{2F^t}\right] + \int \frac{\mathrm{d}u}{\sqrt{2\pi\left(F^t N^t\right)^3}} \exp\left[-\frac{(u-E^t)^2}{2F^t}\right] \frac{N^t - M^t \tanh(u)}{\cosh^2(u)}\right\} \qquad (53)
\end{aligned}$$

The first term of the RHS of Eq. (53) is independent of the index i and it is zero if and only if the integrand is an odd function. This is true if $\tanh(u) = \frac{N^t}{M^t} \tanh\left(\frac{uE^t}{F^t}\right)$ $\forall u \in \mathbb{R}$. This condition is only satisfied if $E^t(\boldsymbol{\gamma}^c) = F^t(\boldsymbol{\gamma}^c)$ which automatically makes $M^t = N^t$. By the application of this condition, the sum between curly brackets in the second term at the RHS of Eq. (53) is always positive. Therefore it must be $\left.\frac{\partial E^t}{\partial \gamma_i} - \frac{1}{2}\frac{E^t}{F^t}\frac{\partial F^t}{\partial \gamma_i}\right|_{\gamma_i^c} = 0$ in order to make $\frac{\partial}{\partial \gamma_i}\mathscr{E}^t = 0$.

The conditions $E^t(\boldsymbol{\gamma}^c) = F^t(\boldsymbol{\gamma}^c)$ and $\left.\frac{\partial E^t}{\partial \gamma_i} - \frac{1}{2}\frac{E^t}{F^t}\frac{\partial F^t}{\partial \gamma_i}\right|_{\gamma_i^c} = 0$ imply that:

$$\begin{aligned}
\ln E^t &= e_0 + \mathbf{e}_1^\mathsf{T}(\boldsymbol{\gamma} - \boldsymbol{\gamma}^c) + \frac{1}{2}(\boldsymbol{\gamma} - \boldsymbol{\gamma}^c)^\mathsf{T}\mathbf{E}_2(\boldsymbol{\gamma} - \boldsymbol{\gamma}^c) + \ldots \\
\ln F^t &= e_0 + 2\mathbf{e}_1^\mathsf{T}(\boldsymbol{\gamma} - \boldsymbol{\gamma}^c) + \frac{1}{2}(\boldsymbol{\gamma} - \boldsymbol{\gamma}^c)^\mathsf{T}\mathbf{F}_2(\boldsymbol{\gamma} - \boldsymbol{\gamma}^c) + \ldots ,
\end{aligned}$$

therefore, if the critical point is a minimum, then $E^t/\sqrt{F^t}$ has a pick at the criticality and the expansion $E^t/\sqrt{F^t} = \exp\left\{\frac{1}{2}e_0 - \frac{1}{2}(\boldsymbol{\gamma} - \boldsymbol{\gamma}^c)^\mathsf{T}\left(\frac{1}{2}\mathbf{F}_2 - \mathbf{E}_2\right)(\boldsymbol{\gamma} - \boldsymbol{\gamma}^c) + \ldots\right\}$ has a second term that satisfy the conditions: $\det\left(\frac{1}{2}\mathbf{F}_2 - \mathbf{E}_2\right) > 0$ and $\left(\frac{1}{2}\mathbf{F}_2 - \mathbf{E}_2\right)_{ii} > 0$.

---

[1] M. Opper and D. Saad, *Advanced Mean Field Methods: Theory and Practice*, MIT Press, Cambridge, MA 2001

[2] Y. Kabashima, J. Phys. A **36**, 11111 (2003).

[3] J.P. Neirotti and D. Saad, Europhys. Lett. **71**, 866 (2005).

[4] S. Verdú, *Multiuser Detection*, Cambridge University Press UK (1998).

[5] H. S. Seung, H. Sompolinsky and N. Tishby, Phys. Rev. A **45**, 6056 (1992).

[6] J. Pearl, *Probabilistic Reasoning in Intelligent Systems*, Morgan Kaufmann Publishers, San Francisco, CA (1988).

[7] F.V. Jensen, *An Introduction to Bayesian Networks*, UCL Press, London (1996).

[8] D.J.C. MacKay, *Information Theory, Inference and Learning Algorithms*, Cambridge University Press (2003).

[9] Y. Weiss, *Neural Computation* **12**, 1 (2000).

[10] Y. Kabashima, D. Saad, Europhys. Lett. **44**, 668 (1998).

[11] J.S. Yedidia, W.T. Freeman and Y. Weiss, in *Advances in Neural Information Processing Systems* **13**, 698 (2000).

[12] M. Mézard, G. Parisi and M.A. Virasoro, *Spin Glass Theory and Beyond*, World Scientific, Singapore (1987).

[13] H. Nishimori, *Statistical Physics of Spin Glasses and Information Processing*, Oxford University Press, UK, (2001).

[14] M. Mézard, G. Parisi and R. Zecchina, Science **297**, 812 (2002).

[15] M. Mézard and R. Zecchina, Phys. Rev. E **66**, 056126 (2002).

[16] A. Braustein and R. Zecchina, Phys. Rev. Lett., **96** 030201 (2006)

[17] Y. Kabashima, cs.IT/0506062 (2005).

[18] A.P. Dempster, N. M. Laird and D.B. Rubin, J. Roy. Stat. Soc. B **39**, 1 (1977).

[19] }J.P. Neirotti and D. Saad, in preparation (2005).

[20] Comparison to the recently presented signal detection algorithm of Ref. [17] cannot be carried out due to absence of numerical data.

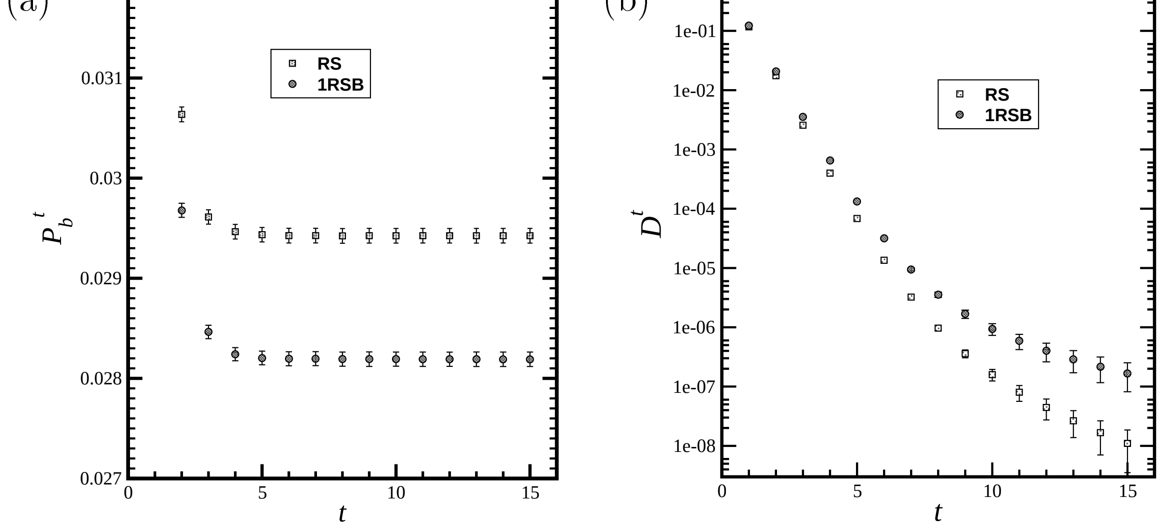

Figure 6: (a) Error probability of the inferred solution evolving in time, for the bi_gaussian noise case. The system load $\beta = 0.25$, true noise level $\sigma_0^2 = 0.25$ and estimated noise $\sigma^2 = 0.01$. Squares represent results of the RS algorithm and circles represent results obtained from the 1RSB algorithm. (b) D^t, a measure of convergence in the obtained solutions, as a function of time; symbols are as in the main figure.

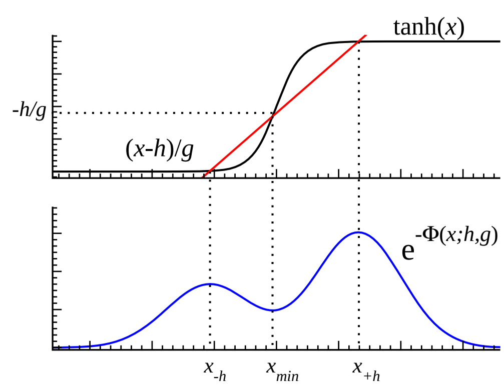

Figure 7: Solutions for the mean field Equation with two maxima and one minimum for a positive value of the field h.