Journal of Statistical Physics, Vol. 124, No. 6, September 2006 (© 2006)
DOI: 10.1007/s10955-006-9162-3

Reconstruction on Trees and Spin Glass Transition

Marc Mézard[1] and Andrea Montanari[2]

Received December 14, 2005; accepted June 30, 2006
Published Online: September 15, 2006

Consider an information source generating a symbol at the root of a tree network whose links correspond to noisy communication channels, and broadcasting it through the network. We study the problem of reconstructing the transmitted symbol from the information received at the leaves. In the large system limit, reconstruction is possible when the channel noise is smaller than a threshold.

We show that this threshold coincides with the dynamical (replica symmetry breaking) glass transition for an associated statistical physics problem. Motivated by this correspondence, we derive a variational principle which implies new rigorous bounds on the reconstruction threshold. Finally, we apply a standard numerical procedure used in statistical physics, to predict the reconstruction thresholds in various channels. In particular, we prove a bound on the reconstruction problem for the antiferromagnetic "Potts" channels, which implies, in the noiseless limit, new results on random proper colorings of infinite regular trees.

This relation to the reconstruction problem also offers interesting perspective for putting on a clean mathematical basis the theory of glasses on random graphs.

KEY WORDS: reconstruction, spin glasses, reconstruction threshold, phase transition

PACS: 02.50.−r (Probability theory, stochastic processes, and statistics), 64.70.Pf (Glass transitions), 89.75.Hc (Networks and genealogical trees)

1. INTRODUCTION

Consider the following broadcast problem.[(18)] An information source at the root of a tree network produces a letter taken from a q-ary alphabet $x \in \{1, \dots, q\}$ (we shall sometimes refer to a letter from this alphabet as to a 'color'). The symbol

[1] Laboratoire de Physique Théorique et Modèles Statistiques, Université de Paris-Sud, bâtiment 100, 91405 Orsay Cedex, France.
[2] Laboratoire de Physique Théorique de l'Ecole Normale Supérieure, 24 rue Lhomond 75231, Paris Cedex 05, France.

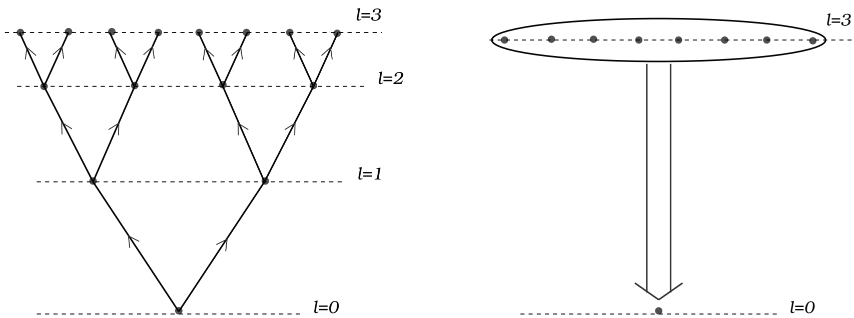

Fig. 1. *Left*: Broadcast on a tree. The signal is sent from the root. Each edge is a noisy communication channel broadcasting upwards. *Right*: The reconstruction problem asks to find what signal was sent from the root, given the signals received on the leaves.

is propagated along the edges of the tree. For simplicity we start with a regular k-ary tree $\mathbb{T}_k$, cf. Fig. 1, in which every vertex has exactly k descendants (every vertex has degree $k+1$ except the root which has degree k), a more general setting is described in Ref. 9 and in Sec. 7. Each edge of the tree is an instance of the same noisy communication channel: If the letter x is transmitted through the channel, $y \in \{1, \ldots, q\}$ is received with probability $\pi(y|x)$ (with $\pi(y|x) \geq 0$ and $\sum_y \pi(y|x) = 1$). The problem of reconstruction is the following: consider all the symbols received at the vertices of the ℓth generation. Does this configuration contain a non-vanishing information on the letter transmitted by the root, in the large ℓ limit?

Beyond its fundamental interest in probability, this problem is relevant to genetics (propagation of genes from an ancestor),[24] to statistical physics (models on Bethe lattices), and information theory (the problem being equivalent to computing the information capacity of the tree network).[8]

An important general bound was obtained by Kesten and Stigum (KS).[13,12] Consider the matrix π with entries $\pi(y|x)$, $x, y \in \{1, \ldots, q\}$ and let $\lambda_2(\pi)$ be its eigenvalue with the second largest absolute value. Then, if $k|\lambda_2(\pi)|^2 > 1$, the reconstruction problem is solvable: the leaves asymptotically contain some information on the letter sent by the root. In fact in this case the census of the variables in the ℓth generation (the number of leaves which have received each letter) contains some information on the root. Conversely, if $k|\lambda_2(\pi)|^2 < 1$, the census contains asymptotically no information on the root.[21] Therefore, the KS condition $k|\lambda_2(\pi)|^2 = 1$ defines a threshold for the maximum amount of noise allowing census reconstruction. For larger noise ($k|\lambda_2(\pi)|^2 < 1$) one may wonder whether reconstruction is possible exploiting the whole set of symbols received at the ℓth generation, through a clever use of the correlations between the symbols received on the leaves. The answer depends on the channel.

In most of this paper we shall focus onto transition kernels $\pi(\,\cdot\,|\,\cdot\,)$ satisfying the detailed balance condition (reversible) with respect to the uniform distribution $\overline{\eta}(x) = 1/q$. In other words $\pi(y|x) = \pi(x|y)$. With a slight abuse of notation we shall write $\pi(y|x) = \pi(y, x)$. For the problem to be non-trivial, we also assume $\pi(\,\cdot\,|\,\cdot\,)$ to be irreducible and aperiodic. A particularly important example in this family is provided by q-ary symmetric channels (or, borrowing from the statistical mechanics terminology, 'Potts' channels)

$$\pi(y|x) = \begin{cases} 1-\varepsilon & \text{if } y = x. \\ \varepsilon/(q-1) & \text{otherwise} \end{cases}. \tag{1}$$

If $\varepsilon < 1 - 1/q$, $y = x$ is the most likely channel output when the input is x: we shall refer to this case as the 'ferromagnetic' Potts channel. If $\varepsilon > 1 - 1/q$, the opposite happens and we shall speak of 'antiferromagnetic' Potts channel. The particular case $\varepsilon = 1$ is of special interest, since the broadcast process provides a uniformly random proper coloring of the ℓ-generations k-ary tree $\mathbb{T}_k(\ell)$.

It is intuitively clear that the channel (1) 'gets worse' as ε increases from 0 to $1 - 1/q$ (ferromagnetic channel) and 'improves' as ε goes from $1 - 1/q$ to 1 (antiferromagnetic channel). A result by Mossel[17] implies that there exist a ferromagnetic and an antiferromagnetic threshold, respectively $\varepsilon_{\rm r}^{+}(k, q) \in [0, 1 - 1/q]$ and $\varepsilon_{\rm r}^{-}(k, q) \in [1 - 1/q, 1]$, such that the reconstruction problem is solvable when $\varepsilon \in [0, \varepsilon_{\rm r}^{+}[\cup]\varepsilon_{\rm r}^{-}, 1]$ and insolvable if $\varepsilon \in]\varepsilon_{\rm r}^{+}, \varepsilon_{\rm r}^{-}[$. Hereafter we shall drop the $\pm$ superscripts whenever they are clear from the context.

The KS condition $k|\lambda_2(\pi)|^2 > 1$ is satisfied (and the problem is census-solvable) for the channel (1) if and only if $\varepsilon \in [0, \varepsilon_{\rm KS}^{+}(k, q)[\cup]\varepsilon_{\rm KS}^{-}(k, q), 1]$, where:

$$\varepsilon_{\rm KS}^{\pm}(k, q) = \frac{q-1}{q}\left(1 \mp \frac{1}{\sqrt{k}}\right). \tag{2}$$

Notice that the above formula yields $\varepsilon_{\rm KS}^{-}(k, q) > 1$ for some pairs of (k, q). In fact, for the antiferromagnetic channel, the census-reconstruction problem (as well as the general reconstruction problem) is not necessarily solvable for $\varepsilon = 1$.

It is known[5] that, for $q = 2$ (the "binary symmetric" channel, also known as the "symmetric Ising" case), the reconstruction threshold is equal to the KS one: $\varepsilon_{\rm r}(k, 2) = \varepsilon_{\rm KS}(k, 2)$ (for $q = 2$ the ferromagnetic and antiferromagnetic cases are equivalent via the mapping $\varepsilon \mapsto 1 - \varepsilon$). In general, the KS bound implies $\varepsilon_{\rm KS}^{+}(k, q) \geq \varepsilon_{\rm r}^{+}(k, q)$, and $\varepsilon_{\rm r}^{-}(k, q) \geq \varepsilon_{\rm KS}^{-}(k, q)$. Furthermore, in Ref. 17 it was shown that, for all k, when q is large enough, $\varepsilon_{\rm KS}^{+}(k, q) > \varepsilon_{\rm r}^{+}(k, q)$ strictly: reconstruction is possible at noise levels where census reconstruction does not work. However, several fundamental questions remain open even for simple Potts channels: Is there any pair (k, q), with $q > 2$, such that $\varepsilon_{\rm r}(k, q) = \varepsilon_{\rm KS}(k, q)$? How to distinguish systematically between $\varepsilon_{\rm r}(k, q)$ and $\varepsilon_{\rm KS}(k, q)$? How to determine $\varepsilon_{\rm r}(k, q)$ accurately when it does not coincide with $\varepsilon_{\rm KS}(k, q)$? We shall address these issues in the following.

The reconstruction problem is intimately related to statistical physics. Consider a model of Potts spins $y_i \in \{1, \dots, q\}$, on a finite rooted tree with ℓ generations, to be denoted by $\mathbb{T}_k(\ell)$. Suppose that the energy of a configuration $\underline{y}^\ell \equiv \{y_i : i \in \mathbb{T}_k(\ell)\}$ is given by:

$$E(\underline{y}^\ell) = -J \sum_{(i,j)\in\mathbb{T}_k(\ell)} \delta_{y_i,y_j} \, , \tag{3}$$

where (i, j) denotes pairs of spins connected by an edge of the tree. Let $\underline{Y}^\ell$ be the random configuration produced by the broadcast process with channel (1) *up to* generation ℓ, when the transmitted symbol is uniformly random in $\{1, \dots, q\}$. Then

$$\mathbb{P}\{\underline{Y}^\ell = \underline{y}^\ell\} = \frac{1}{Z} \exp\{-\beta E(\underline{y}^\ell)\} \tag{4}$$

provided we make the identification

$$e^{-\beta J} = \frac{\varepsilon}{(q-1)(1-\varepsilon)}. \tag{5}$$

In other words, the broadcast process allows to construct one particular Gibbs measure (state) associated to the energy function (3): in the statistical physics terminology this is the free-boundary measure. In general this is not the unique Gibbs measure for this energy function. For instance, if $\varepsilon < \frac{q-1}{q}\,(1 - \frac{1}{k})$, one can construct q 'ferromagnetic' states as well. Even if more than one Gibbs state exists, the free-boundary state can be extremal (or 'pure'). It turns out that the reconstruction problem is solvable if and only if the Gibbs state with free boundary conditions is not extremal.

Given the strong connection between extremality of Gibbs states and spatial decay of correlations,[11] the last remark is not surprising. What is more surprising (and constitutes the main theme of this paper) is the relation of the reconstructibility with the existence of a dynamical glass phase. In recent years, an ongoing effort has been devoted to the study of glassy models on sparse random graphs. These are graphs which contain cycles but locally 'look like' a tree (e.g. uniformly random graphs with given degree). One of the most widespread features of these models, is the occurrence of glass phases in which the Boltzmann measure gets split into an exponential number of 'lumps' (also referred to as clusters or pure states). This phenomenon is usually studied by solving some 'one-step replica symmetry breaking' (1RSB) distributional equations. In the following we show that these equations, as well as the criterion used to detect glass phases, do indeed coincide with the solvability of an appropriate reconstruction problem.

In spin glass theory, one can encounter two types of transitions to a glass phase. In the first case the transition is continuous in a properly defined order parameter. In spin glass jargon this leads to a phase with 'full replica symmetry breaking' (FRSB). In the second it is discontinuous, leading to 1RSB. Both situations occur in the reconstruction problem, depending on the alphabet and the channel. In the continuous case, the phase transition location is given by a local instability which coincides with the KS threshold, and one has $\varepsilon_{\rm r}(k, q) = \varepsilon_{\rm KS}(k, q)$. This happens, for instance, when $q = 2$. In the opposite case, the 'dynamical' glass transition is discontinuous and its location (which still coincides with the reconstruction threshold) is distinct from the KS one. In the ferromagnetic Potts model one has, for instance, $\varepsilon_r(k, q) > \varepsilon_{KS}(k, q)$ at large enough q.

The coincidence of the reconstruction threshold with the dynamical glass transition, apart from being interesting in itself, allows us to adapt several techniques developed within the theory of spin glasses in order to study the reconstruction problem. On the one hand, importing a numerical procedure currently used in this field, we determine the threshold for several pairs k, q. These results lead us to conjecture that $\varepsilon_{\rm r}(k, q) = \varepsilon_{\rm KS}(k, q)$, for k not too large and $q \leq 4$ (in the ferromagnetic case) or $q \leq 3$ (in the antiferromagnetic case).

Furthermore, we derive a variational principle for the reconstruction problem. In the antiferromagnetic case, this implies a rigorous bound on the reconstruction threshold, which allows to confirm the strict inequality $\varepsilon_{\rm r}^-(k, q) < \varepsilon_{\rm KS}^-(k, q)$ in most of the cases in which this was found to be the case numerically. Although we conjecture such a bound to hold in much greater generality, we weren't able to prove it, and we leave it as a conjecture.

The paper is organized as follows. In Sec. 2 we define the main objects studied in the paper and prove the coincidence between reconstruction and dynamical glass transition. In Sec. 3 we state our variational principle and prove that it provides a rigorous bound for a class of kernels $\pi(\,\cdot\,|\,\cdot\,)$ including the antiferromagnetic model. In Secs. 4 and 5 we apply this principle as well as a numerical procedure to the determination of thresholds for the Potts channel, respectively in the ferromagnetic and antiferromagnetic case. Section 6 discusses the physical meaning of the relation between reconstruction and glass transitions. Section 7 explains how our methods (and the glass—reconstruction correpondance) can be generalized to a broad category of broadcast and reconstruction problems on trees, going much beyond the Potts channel. We conclude in Sec. 8 by summarizing a few conjectures and pointing out some interesting open problems.

2. DISTRIBUTIONAL RECURSION

2.1. Definitions

We denote by V and E the vertex and edge sets of the infinite k-ary tree $\mathbb{T}_k$, by 0 its root and by V_ℓ the set of generation-ℓ vertices ($|V_\ell| = k^\ell$). The broadcast process generates a random color configuration $\underline{X} \equiv \{X_i \ : \ i \in V\}$ with $X_i \in \{1, \dots, q\}$. The root color $X_0 \in \{1, \dots, q\}$, which we also call the transmitted color, is uniformly random. Then, given the values of X up to the ℓ-th generation, the values at the $(\ell + 1)$-th generation are conditionally independent. If a vertex in the ℓ-th generation has color y, the probability that a vertex connected to it in the $(\ell + 1)$-th generation has color z is $\pi(z|y)$.

We shall denote by $\underline{X}_\ell$ the configurations of colors at the ℓth generation, and by $\underline{Y}^\ell$ the configuration *up to* the ℓth generation (i.e. $\underline{Y}^\ell = \{\underline{X}_0, \underline{X}_1, \dots, \underline{X}_\ell\}$). The probability distribution of $\underline{X}_\ell$, conditioned to the choice $X_0 = x$ of the root color will be denoted by $B_x^{(\ell)}(\underline{x}_\ell) \equiv \mathbb{P}\{\underline{X}_\ell = \underline{x}_\ell | X_0 = x\}$.

Suppose now that the configuration of colors at the ℓ-th generation, $\underline{x}_\ell$, is given. We denote by $\eta_\ell(y)$ the probability that the root had sent the color y, given $\underline{X}_\ell$:

$$\eta_\ell(y) = \mathbb{P}[X_0 = y \,| \underline{X}_\ell = \underline{x}_\ell]. \tag{6}$$

$\eta_\ell(\cdot)$ is a probability distribution over $\{1, \dots, q\}$ (i.e. $\eta_\ell(y) \geq 0$ and $\sum_y \eta_\ell(y) = 1$). We shall denote the space of such distributions as $\mathfrak{M}_q$. In order to emphasize the dependency of η_ℓ upon the configuration received in shell ℓ, $\underline{x}_\ell$, we shall sometimes write $\eta_\ell(y) = \eta_{\underline{x}_\ell}(y)$. It is easy to realize that, given the colors received at the ℓth generation, $\eta_\ell(\cdot)$ constitutes a sufficient statistics for the root color x. In other terms, given $\underline{X}_\ell = \underline{x}_\ell$, there is no loss of information in computing $\eta_{\underline{x}_\ell}(\cdot)$ and then guessing X_0 from $\eta_{\underline{x}_\ell}(\cdot)$.

Since $\underline{X}_\ell$ is chosen randomly according to the broadcast process, $\eta_\ell(\cdot)$ is a random probability distribution, i.e. a random point in $\mathfrak{M}_q$. We denote by $Q_x^{(\ell)}(\eta)$ its distribution[3] conditional to the broadcast being started from $X_0 = x$, and call it the 'distribution at the root.' Hereafter a distribution Q over $\mathfrak{M}_q$ will be said *trivial* if it is a singleton on the uniform measure $\overline{\eta}$ (defined by $\forall x : \ \overline{\eta}(x) = 1/q$). Clearly the reconstruction problem is solvable if and only if the large ℓ limit of $Q_x^{(\ell)}$ is non trivial.

There are several ways of characterizing quantitatively the large ℓ behavior of $Q_x^{(\ell)}$. We shall consider below two parameters $I_\ell \equiv I(X_0; \underline{X}_\ell)$ and Ψ_ℓ, which are defined by:

$$I_\ell = \frac{1}{q} \sum_x \int \log_2 \frac{\eta(x)}{\overline{\eta}(x)} \, dQ_x^{(\ell)}(\eta), \qquad \Psi_\ell = \frac{1}{q} \sum_x \int [\eta(x) - \overline{\eta}(x)] \, dQ_x^{(\ell)}(\eta). \tag{7}$$

[3] Notice that we adopt here the standard physicists convention: we carelessly denote probability distributions by their densities even if such densities do not exist. We shall also denote by $\int f(\eta)\, dQ_x^{(\ell)}(\eta)$ the expectation with respect to such a distribution. The fussy reader can easily translate all the formulae below in the standard probability language.

I_ℓ gives the number of information bits that can be transmitted reliably per network use. Ψ_ℓ is the probability that the reconstruction is successful when the receiver guesses color y with probability $\eta_\ell(y)$, minus the the same probability when the receiver guesses uniformly. These are non-negative quantities and can be shown to be non-increasing functions of ℓ. We furthermore let $I_\infty \equiv \lim_{\ell\to\infty} I_\ell$, and $\Psi_\infty \equiv \lim_{\ell\to\infty} \Psi_\ell$.

The tree reconstruction problem can be rephrased by saying that the problem is solvable if and only if $I_\infty > 0$ (or, equivalently, $\Psi_\infty > 0$). For instance, for the ferromagnetic Potts channel, the threshold $\varepsilon_{\rm r}^{+}(k, q)$ is the supremum of the values of ε such that $I_\infty > 0$.

2.2. Merging Rooted Trees

How does one compute the distribution $\eta(y)$ on the root, given a boundary $\underline{X}_\ell = \underline{x}_\ell$? Using the tree-structure, this can be done iteratively by a dynamical programming procedure starting from the leaves. Suppose that at some point in this iteration we have determined the probability distributions $\eta_1(\,\cdot\,), \dots, \eta_k(\,\cdot\,)$ of the k vertices in the tree which lie above a given vertex (see Fig. 2, left). Then the probability $\eta(y)$ that this vertex had color y during the broadcast is given by:

$$\eta(y) = \frac{1}{z(\{\eta_i\})} \prod_{i=1}^{k} \left(\sum_{y_i=1}^{q} \pi(y_i|y)\, \eta_i(y_i) \right), z(\{\eta_i\}) \equiv \sum_{y=1}^{q} \prod_{i=1}^{k} \Big(\sum_{y_i} \pi(y_i|y)\, \eta_i(y_i) \Big). \tag{8}$$

This equation defines a mapping between distributions in $\mathfrak{M}_q$: given k distributions $\eta_1, \dots, \eta_k$, one generates a new one $\eta = \mathsf{F}(\eta_1, \dots, \eta_k)$. Iterating this mapping downwards from the leaves down to the root, one can derive the conditional distribution of the transmitted symbol.

Equation (8) naturally induces a recursion equation for the distribution $Q_x^{(\ell)}$. Consider the reconstruction of the root in a rooted tree with $\ell + 1$ generations (see Fig. 2, right). This graph is formed by k subtrees rooted in the vertices $1, \dots, k$,

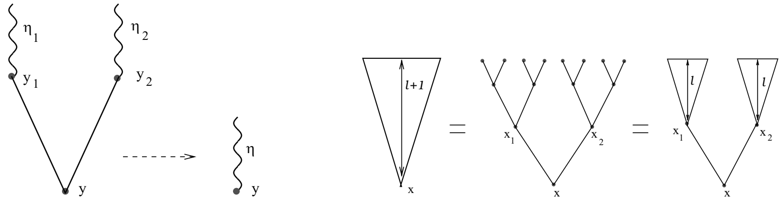

Fig. 2. *Left*: A pictorial representation of the mapping $\eta = \mathsf{F}(\eta_1, \dots, \eta_k)$ defined in (8). Here $k = 2$, a straight line corresponds to the channel π, a wiggly line arriving on a vertex y_j corresponds to a weight $\eta_j(y_j)$. *Right*: A pictorial representation of the recursion (9). A triangle of depth r rooted on variable x_j denotes $Q_{x_j}^{(r)}(\eta_r)$.

which are all joined to the root 0. Each of these subtrees gives an instance of the reconstruction with ℓ generations. Therefore:

$$Q_x^{(\ell+1)}(\eta) = \sum_{x_1 \ldots x_k} \prod_{i=1}^{k} \pi(x_i|x) \int \delta\left[\eta - \mathsf{F}(\eta_1, \ldots, \eta_k)\right] \prod_{i=1}^{k} dQ_{x_i}^{(\ell)}(\eta_i), \qquad (9)$$

where $\delta[\cdots]$ represents a Dirac delta function on $\mathfrak{M}_q$. In words, in order to generate $\eta(\,\cdot\,)$ with distribution $Q_x^{(\ell+1)}$, one can proceed as follows. First draw k independent colors $x_1, \ldots, x_k$ from the distribution $\pi(\,\cdot\,|x)$. Then, draw $\eta_1, \ldots, \eta_k$ independently with distribution, respectively, $Q_{x_1}^{(\ell)}, \ldots, Q_{x_k}^{(\ell)}$. Finally, let $\eta = \mathsf{F}(\eta_1, \ldots, \eta_k)$.

The initial condition is

$$Q_x^{(0)}(\eta) = \delta\left[\eta - \delta_x\right], \qquad (10)$$

where δ_x is the distribution in $\mathfrak{M}_q$ which has weight unity on color x (it is given by $\delta_x(y) = 1$ if $y = x$, and $\delta_x(y) = 0$ otherwise). The Eqs. (9) and (10) fully characterize the distributions $Q_x^{(\ell)}$. The whole reconstruction problem amounts to understanding the large ℓ properties of these recursions.

2.3. Unconditional Distribution and Symmetry Properties

While $Q_x^{(\ell)}$ gives the distribution of $\eta_\ell(\,\cdot\,)$ (defined in Eq. (6)) *conditional* on the transmitted color being equal to x, it is equally interesting to consider the *unconditional* distribution. We will denote it by $\widehat{Q}^{(\ell)}$. Bayes theorem implies the following relation between $Q_x^{(\ell)}$ and $\widehat{Q}^{(\ell)}$

$$Q_x^{(\ell)}(\eta) = q\, \eta(x)\, \widehat{Q}^{(\ell)}(\eta). \qquad (11)$$

This is in fact a rephrasing of the identity

$$\mathbb{P}\{\eta_{\underline{X}_\ell} = \eta | X_0 = x\} = \frac{\mathbb{P}\{X_0 = x | \eta_{\underline{X}_\ell} = \eta\}\mathbb{P}\{\eta_{\underline{X}_\ell} = \eta\}}{\mathbb{P}\{X_0 = x\}}. \qquad (12)$$

An alternative (analytic) proof can be obtained writing $Q_x^{(\ell)}$ in terms of $B_x^{(\ell)}(\underline{x}_\ell)$, the probability that the output of the broadcast process at generation ℓ is $\underline{x}_\ell$, given that the transmitted color is x:

$$Q_x^{(\ell)}(\eta) = \sum_{\underline{x}_\ell} B_x^{(\ell)}(\underline{x}_\ell)\;\; \delta\left[\eta(\cdot) - \frac{B_\cdot^{(\ell)}(\underline{x}_\ell)}{\sum_z B_z^{(\ell)}(\underline{x}_\ell)}\right]. \qquad (13)$$

It is then easy to show that, if $\lambda_0 + \lambda_1 + \ldots + \lambda_{q-1} = 1$, then the expectation value

$$\int \frac{\eta(0)^{\lambda_0}, \ldots, \eta(q-1)^{\lambda_{q-1}}}{\eta(x)}\, dQ_x^{(\ell)}(\eta) \qquad (14)$$

does not depend upon x. This in turns imply that $Q_x^{(\ell)}$ can be written in the form (11) where $\widehat{Q}^{(\ell)}$ is a distribution which does not depend on x (the normalization can be found by summing over x).

If the channel is symmetric with respect to permutations of the colors (as is the case for Potts channels), the distributions $Q_x^{(\ell)}$ and $Q^{(\ell)}$ inherit the same symmetry. More precisely, given a permutation acting on the colors $\sigma \in S_q$, and a distribution $\eta \in \mathfrak{M}_q$, let η^σ be the permuted distribution defined by $\eta^\sigma(x) \equiv \eta(\sigma(x))$. Then, for any permutation σ, $Q_x^{(\ell)}(\eta) = Q_{\sigma(x)}^{(\ell)}(\eta^\sigma)$, and

$$\widehat{Q}^{(\ell)}(\eta^\sigma) = \widehat{Q}^{(\ell)}(\eta). \tag{15}$$

A distribution satisfying condition (15) will be called 'symmetric.'

Let us finally notice that the parameters introduced in Sec. 2 to measure the amount of information on the transmitted color available at the ℓth generation can be expressed in terms of the distribution $\widehat{Q}^{(\ell)}$

$$I_\ell = \int D(\eta||\overline{\eta})\, d\widehat{Q}^{(\ell)}(\eta), \qquad \Psi_\ell = \sum_x \int [\eta(x) - \overline{\eta}(x)]^2\, d\widehat{Q}^{(\ell)}(\eta). \tag{16}$$

Here we use the standard notation for the Kullback-Leibler distance[8] $D(\eta||\overline{\eta}) \equiv \sum_x \eta(x) \log_2[\eta(x)/\overline{\eta}(x)]$. In deriving the second of these expressions, we used the fact that $\int \eta(x)\, d\widehat{Q}^{(\ell)}(\eta) = \overline{\eta}(x) = \frac{1}{q}$ which follows from (11).

2.4. Recursion for the Unconditional Distribution and Spin Glass Correspondence

The recursion relation (9) on $Q_x^{(\ell)}$ implies the following recursion for the unconditional distribution:

$$\widehat{Q}^{(\ell+1)}(\eta) = q^{k-1} \int z(\{\eta_i\})\ \delta\,[\eta - \mathsf{F}(\eta_1, \dots, \eta_k)] \prod_{i=1}^{k} d\widehat{Q}^{(\ell)}(\eta_i), \tag{17}$$

where $z(\{\eta_i\})$ is defined as in Eq. (8). The initial condition (10) converts into $\widehat{Q}^{(0)}(\eta) = \frac{1}{q}\sum_{y=1}^{q} \delta[\eta(\,\cdot\,) - \delta_y(\,\cdot\,)]$. It is also interesting to study the fixed points of this recursion, i.e. the distributions $\widehat{Q}^*$ satisfying:

$$\widehat{Q}^*(\eta) = q^{k-1} \int z(\{\eta_i\})\ \delta\,[\eta - \mathsf{F}(\eta_1, \dots, \eta_k)] \prod_{i=1}^{k} d\widehat{Q}^*(\eta_i), \tag{18}$$

Notice that any solution of this equation has necessarily expectation $\int \eta(x)\, d\widehat{Q}^*(\eta) = \overline{\eta}(x)$ (this is proved by taking expectation on both sides). Any probability distribution over $\mathfrak{M}_q$ satisfying this condition will be hereafter said to be 'consistent.'

The distributional Eq. (18) is well known in spin glass theory and usually referred to as '1RSB equation with Parisi parameter $m = 1$' (in the general 1RSB scheme the factor $z(\{\eta_i\})$ is raised to a power $m \in [0, 1]$). It is used to determine whether an associated statistical mechanics model is in a glass phase. We shall return to the definition of the associated model in Sec. 6. For the time being, we shall adopt the usual physicists criterion as a *definition*: We will say that the statistical mechanics model associated to the reconstruction problem (characterized by a degree/kernel pair k, π) *admits a glass phase* if and only if Eq. (18) has a non-trivial solution.

When considering a continuous family of kernels $\pi(\cdot\,|\,\cdot)$, parametrized by a noise level ε, the value of ε where a non-trivial solution appears is called a dynamical glass transition. The result below implies that this coincides indeed with the reconstruction threshold (i.e. with the extremality threshold for the free boundary Gibbs measure on the infinite tree).

Proposition 1. *The statistical mechanics model associated with the degree/ kernel pair k, π* admits a glass phase*, if and only if the corresponding reconstruction problem is solvable.*

Proof: As noticed for instance in Ref. 6, the sequence of random variables $\eta_\ell(\cdot)$ (not conditioned on the root color), converges almost surely to a limit $\eta_\infty(\cdot)$. As a consequence, the sequence of distributions $\widehat{Q}^{(\ell)}$ converges weakly to the distribution $\widehat{Q}^{(\infty)}$ of $\eta_\infty(\cdot)$. By taking the limit of Eq. (17) (and noticing that $\mathsf{F}(\eta_1, \ldots, \eta_k)$ and $z(\{\eta_i\})$ are continuous and bounded) we find that $\widehat{Q}^{(\infty)}$ must satisfy the fixed point condition (18). If the reconstruction problem is solvable, then $\widehat{Q}^{(\infty)}$ is non-trivial and therefore, according to our definition, the pair k, π *admits a glass phase*.

Conversely[4], let $\widehat{Q}^*$ be a non-trivial solution of (18). Following (11), define the distribution $Q^*_x(\eta) = q\, \eta(x) \widehat{Q}^*(\eta)$. Because of the above calculations, the q distributions Q^*_x, $x \in \{1, \ldots, q\}$ are a fixed point of the recursion (9). We will now show that they can be used to reconstruct the transmitted color from the output at generation ℓ, with probability of success independent of ℓ and strictly larger than $1/q$.

The reconstruction procedure goes as follows. Suppose that the broadcast has generated the values $X_i = x_i$ for $i \in V_\ell$. For each vertex i, generate η_i from the distribution $Q^*_{x_i}$. Consider now a vertex $a \in V_{\ell-1}$, connected to $a_1, \ldots, a_k$ in V_ℓ. Compute $\eta_a = \mathsf{F}(\eta_{a_1}, \ldots, \eta_{a_k})$, where $\mathsf{F}(\cdots)$ is defined as in Eq. (8). Proceeding downwards from the leaves to the root, this allows to construct η_0. At this point

[4] The idea of the converse is due to James Martin who kindly agreed to let us publish it here.

the transmitted symbol can be guessed, for instance, by choosing $X_0 = y$ with probability proportional to $\eta_0(y)$.

We claim that for each vertex $j \in \mathbb{T}_k(\ell)$, and conditional to the broadcast having produced $X_j = x_j$, the $\eta_j(\,\cdot\,)$ provided by the above procedure is distributed according to $Q^*_{x_j}$. This in particular implies that the probability of guessing correctly the root color is

$$\frac{1}{q}\sum_x \int \eta(x)\, dQ^*_x(\eta) = \sum_x \int \eta(x)^2\, d\widehat{Q}^*(\eta) > \frac{1}{q}. \tag{19}$$

The claim is proved by induction starting from the leaves and proceeding downwards to the root. It is true by construction for the vertices of the last generation. Assume it to be true up to generation r and consider a site a in generation $r-1$ connected to $a_1, \ldots, a_k$ in V_r, under the condition $X_a = x_a$. It is clear that the distribution of η_a is obtained through the recursion (9) (with $Q^{(\ell)}_{x_i}$ replaced by $Q^*_{x_{a_i}}$), and since $Q^*_{x_{a_i}}$ is a fixed point of this recursion, this proves the claim. □

3. VARIATIONAL PRINCIPLE

3.1. The General Principle

Here we establish a variational principle from which the fixed point Eq. (18) for the distribution at the root can be deduced. We shall not try to explain here its physical origin, which is related to spin glass theory,[19] but just discuss the relation with the reconstruction problem. Throughout this section we use the notation $\pi(x|y) = \pi(y|x) \equiv \pi(x,y)$. Given a distribution $\widehat{Q}$ over $\mathfrak{M}_q$, we define its *complexity* as

$$\Sigma(\widehat{Q}) = -\frac{k+1}{2}\int \widehat{W}_{\rm e}(\eta_1,\eta_2)\;\; d\widehat{Q}(\eta_1)\, d\widehat{Q}(\eta_2) + \int \widehat{W}_{\rm v}(\eta_1,\ldots,\eta_{k+1})\prod_{i=1}^{k+1} d\widehat{Q}(\eta_i)\,, \tag{20}$$

where

$$\widehat{W}_{\rm e} \equiv -\left[\frac{\sum_{x_1,x_2}\eta_1(x_1)\eta_2(x_2)\pi(x_1,x_2)}{\sum_{x_1,x_2}\overline{\eta}(x_1)\overline{\eta}(x_2)\pi(x_1,x_2)}\right]\log\left[\frac{\sum_{x_1,x_2}\eta(x_1)\eta(x_2)\pi(x_1,x_2)}{\sum_{x_1,x_2}\overline{\eta}(x_1)\overline{\eta}(x_2)\pi(x_1,x_2)}\right], \tag{21}$$

$$\widehat{W}_{\rm v} \equiv -\left[\frac{\sum_x\prod_i\sum_{x_i}\eta_i(x_i)\pi(x,x_i)}{\sum_x\prod_i\sum_{x_i}\overline{\eta}(x_i)\pi(x,x_i)}\right]\log\left[\frac{\sum_x\prod_i\sum_{x_i}\eta_i(x_i)\pi(x,x_i)}{\sum_x\prod_i\sum_{x_i}\overline{\eta}(x_i)\pi(x,x_i)}\right]. \tag{22}$$

The complexity is interesting for the reconstruction problem because of the following remark:

Proposition 2. *Let $\widehat{Q}^*$ be a distribution over $\mathfrak{M}_q$ which satisfies the fixed point Eq. (18). Then $\widehat{Q}^*$ is a stationary point of the complexity $\Sigma(\,\cdot\,)$. More precisely, given any consistent distribution $\widehat{Q}$ over $\mathfrak{M}_q$, define $\Sigma^*(t) \equiv \Sigma((1-t)\widehat{Q}^* + t\widehat{Q})$. Then*

$$\left.\frac{d\Sigma^*}{dt}\right|_{t=0} = 0. \tag{23}$$

Proof: This proposition is a direct consequence of Lemma 2 in Appendix A (the proof consists in explicitly computing the derivative of $\Sigma^*(t)$ and checking that it vanishes under the fixed point conditions). □

The complexity Σ can also be written in terms of the conditional distributions $Q_x(\eta) = q\eta(x)\widehat{Q}(\eta)$. Define $p(x_1, x_2)$ to be the marginal distribution of two neighboring variables on the tree: $p(x_1, x_2) = \pi(x_1, x_2)/q$. Similarly, let $p(x_1, \dots, x_{k+1}) = [\sum_x \pi(x, x_1) \dots \pi(x, x_{k+1})]/q$, the distribution of $k+1$ variables with one common neighbor. The complexity is then given, in terms of $Q_x(\eta)$, by:

$$\Sigma(Q) = -\frac{k+1}{2} \sum_{x_1, x_2} p(x_1, x_2) \int W_{\rm e}(\eta_1, \eta_2)\, dQ_{x_1}(\eta_1)\, dQ_{x_2}(\eta_2)$$

$$+ \sum_{\{x_i\}} p(x_1, \dots, x_{k+1}) \int W_{\rm v}(\eta_1, \dots, \eta_{k+1}) \prod_{i=1}^{k+1} dQ_{x_i}(\eta_i), \tag{24}$$

where

$$W_{\rm e} \equiv -\log\left[\frac{\sum_{x_1,x_2} \eta_1(x_1)\eta_2(x_2)\pi(x_1, x_2)}{\sum_{x_1,x_2} \overline{\eta}(x_1)\overline{\eta}(x_2)\pi(x_1, x_2)}\right], \tag{25}$$

$$W_{\rm v} \equiv -\log\left[\frac{\sum_x \prod_i \sum_{x_i} \eta_i(x_i)\pi(x, x_i)}{\sum_x \prod_i \sum_{x_i} \overline{\eta}(x_i)\pi(x, x_i)}\right]. \tag{26}$$

3.2. Implications on Reconstructibility

Experience from spin glass theory, and the physical interpretation of the complexity, suggests the following conjecture:

Conjecture 1. *Consider the reconstruction problem for the k-ary tree and a reversible channel $\pi(y|x) = \pi(x|y)$. If there exists a consistent distribution $\widehat{Q}^{\rm tr}$ over $\mathfrak{M}_q$, such that $\Sigma(\widehat{Q}^{\rm tr}) < 0$, then the reconstruction problem is solvable.*

Let us give here a few comments in favor of the plausibility of this conjecture. Notice first that, if $\widehat{Q}$ is trivial, then $\Sigma(\widehat{Q}) = 0$. Let $\mathsf{P}(\mathfrak{M}_q)$ denote the space of consistent probability distributions over $\mathfrak{M}_q$. Suppose that there exists $\widehat{Q}^{\text{tr}}$ with $\Sigma(\widehat{Q}^{\text{tr}}) < 0$. Consider now the distribution $\widehat{Q}^* \in \mathsf{P}(\mathfrak{M}_q)$ such that the complexity is minimal. Of course $\Sigma(\widehat{Q}^*) \le \Sigma(\widehat{Q}^{\text{tr}}) < 0$ and therefore $\widehat{Q}^*$ is non-trivial. If (i) $\widehat{Q}^*$ is a stationary point of the complexity, and (ii) the stationary points of $\Sigma(\widehat{Q})$ in $\mathsf{P}(\mathfrak{M}_q)$ coincide with the solutions of the fixed point Eq. (18), then the existence of $\widehat{Q}^{\text{tr}}$ implies that the reconstruction problem is solvable. Point (i) amounts to banishing the possibility that $\widehat{Q}^*$ is on the 'border' of $\mathsf{P}(\mathfrak{M}_q)$. Point (ii) is a stronger version of Proposition 2.

Notice that a priori one could formulate a similar conjecture with a $\widehat{Q}^{\text{tr}}$ having $\Sigma(\widehat{Q}^{\text{tr}}) > 0$, replacing 'minimum' with 'maximum' and 'negative' with 'positive' in the above. It is easy to find counterexamples showing that this 'reverse' conjecture is false. The reason is probably that the distribution $\widehat{Q}^*$ maximizing $\Sigma(\widehat{Q})$ is on the border of $\mathsf{P}(\mathfrak{M}_q)$, and therefore (i) does not hold.

Assuming Conjecture 1 to hold, it implies a simple variational technique for proving that reconstruction is possible. Just consider an explicit finite-dimensional family of distributions $\widehat{Q}_\mu$ depending on some parameters $\mu \in \mathbb{R}^d$, and minimize $\Sigma(\widehat{Q}_\mu)$ over μ. If the minimum is negative, then reconstruction is possible. We will apply the *variational principle* in this form in the next Sections. In the rest of this Sec. (and in Appendix A) we shall prove the principle for a special family of kernels $\pi(\,\cdot\,|\,\cdot\,)$ including the antiferromagnetic Potts channel.

We define a kernel $\pi(y|x) = \pi(x,y)$, $x,y \in \{1,\dots,q\}$ to be 'frustrated' if it can be decomposed as $\pi(x,y) = \pi_* - \widehat{\pi}(x,y)$ where $\pi_* \in \mathbb{R}$ is a constant and $\widehat{\pi}(x,y)$, $x,y \in \{1,\dots,q\}$ is a positive-definite matrix. The antiferromagnetic Potts kernel is a particular instance of this family, with $\pi_* = \varepsilon/(q-1)$, and $\widehat{\pi}(x,y) = |\lambda_2|\,\delta_{x,y}$ where $\lambda_2 = 1 - q\varepsilon/(q-1)$.

Our basic result is the following.

Lemma 1. *Let $\pi(\,\cdot,\,\cdot\,)$ be a frustrated kernel and $\widehat{Q}^*$ a consistent distribution over $\mathfrak{M}_q$ which is not a solution of the associated fixed point Eq. (18). Then there exists a consistent distribution $\widehat{Q}$ over $\mathfrak{M}_q$ such that:*

$$\frac{d}{dt}\Sigma((1-t)\widehat{Q}^* + t\widehat{Q})\bigg|_0 < 0. \tag{27}$$

The proof of this statement is postponed to Appendix A. Here we limit ourselves to proving that it implies the desired principle.

Proposition 3. *Conjecture 1 holds true in the case of frustrated kernels: Let $\pi(\,\cdot,\,\cdot\,)$ be a frustrated kernel, and $\Sigma(\,\cdot\,)$ the associated complexity function. If* *there exists a consistent distribution $\widehat{Q}$ over $\mathfrak{M}_q$ such that $\Sigma(\widehat{Q}^{\rm tr}) < 0$, then the reconstruction problem is solvable.*

Proof: Let $\Sigma_{\min} \equiv \inf \Sigma(\widehat{Q})$, the inf being taken in $\mathsf{P}(\mathfrak{M}_q)$. Since this space is subsequentially compact with respect to the weak topology,[28] and Σ is continuous with respect to this topology, there exists a consistent distribution $\widehat{Q}^*$, such that $\Sigma(\widehat{Q}^*) = \Sigma_{\min}$. Because of Lemma 1, $\widehat{Q}_*$ is a solution of Eq. (18). Furthermore $\Sigma(\widehat{Q}_*) \leq \Sigma(\widehat{Q}^{\rm tr}) < 0$ and therefore $\widehat{Q}_*$ is non-trivial. The result is a consequence of Proposition 1. □

3.3. An Application to Potts Channels

Here we describe a simple family of distributions which can be used variationally when studying the Potts channels. We will show in the next sections that, in spite of its simplicity, it leads to rather accurate results.

The family is indexed by a single real parameter $\mu \in [0, 1]$. We shall denote by $\widehat{Q}_\mu$ the corresponding distribution and will write, with some abuse of notation, $\Sigma(\mu) \equiv \Sigma(\widehat{Q}_\mu)$. The distribution $\widehat{Q}_\mu$ attributes equal weight $1/q$ to the q points in $\mathfrak{M}_q$ denoted by $\gamma^{(x)}$, $x \in \{1, \dots, q\}$, defined as follows

$$\gamma^{(x)}(y) = \begin{cases} 1-\mu & \text{if } y = x, \\ \mu/(q-1) & \text{otherwise.} \end{cases} \tag{28}$$

Some calculus shows that $\Sigma(\mu) = -\frac{k+1}{2}\, w_{\rm e}(\mu) + w_{\rm v}(\mu)$, where

$$w_{\rm e}(\mu) = -\frac{1}{q}\, A \log A - \frac{q-1}{q}\, B \log B, \tag{29}$$

$$A = q\left\{\frac{\varepsilon}{q-1} + \left(1 - \frac{q\varepsilon}{q-1}\right)\left[(1-\mu)^2 + \frac{\mu^2}{q-1}\right]\right\}, \tag{30}$$

$$B = q\left\{\frac{\varepsilon}{q-1} + \left(1 - \frac{q\varepsilon}{q-1}\right)\left[\frac{2\mu(1-\mu)}{q-1} + \frac{(q-2)\mu^2}{(q-1)^2}\right]\right\}, \tag{31}$$

and

$$w_{\rm v}(\mu) = -\frac{1}{q^{k+1}} \sum_{n_1,\dots,n_q} \binom{k+1}{n_1,\dots,n_q} z[n] \log z[n], \tag{32}$$

$$z[n] = q^k \left[\frac{\varepsilon+\mu}{q-1} - \frac{q\varepsilon\mu}{(q-1)^2}\right]^{k+1} \sum_{x=1}^{q} \left[\frac{\varepsilon + (q-1-q\varepsilon)(1-\mu)}{\varepsilon + (q-1-q\varepsilon)\mu/(q-1)}\right]^{n_x}, \tag{33}$$

the first sum being restricted to $n_1, \dots, n_q \geq 0$ and $n_1 + \cdot + n_q = k+1$.

Let us briefly discuss how these formulae are used in the following. To be definite, we refer here to the ferromagnetic case, the antiferromagnetic one being

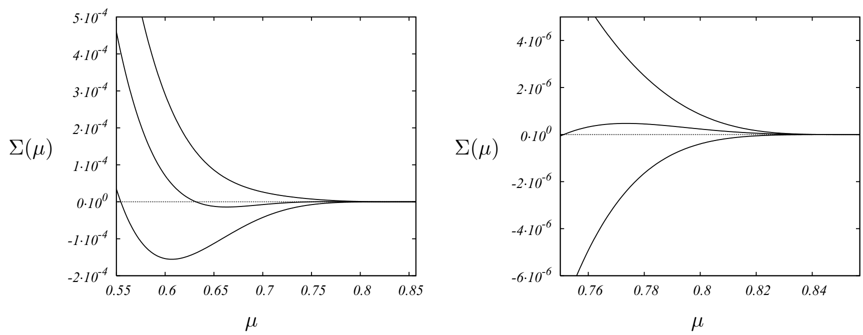

Fig. 3. The complexity for the ferromagnetic Potts channel with $k = 2$ and $q = 7$ within the variational ansatz $\widehat{Q}_\mu$ described in Sec. 3.3. A negative complexity implies that the reconstruction problem is solvable. The three curves correspond (from bottom to top) to $\varepsilon = 0.250, 0.253, 0.256$. The right plot is a zoom near $\mu = 6/7$. The KS threshold for $k = 2, q = 7$ is $\varepsilon_{\mathrm{KS}} \approx 0.2510$. For $\varepsilon = 0.250 < \varepsilon_{\mathrm{KS}}$, $\Sigma(\mu)$ is negative in the neighborhood of $\mu = 6/7$. For $\varepsilon = 0.253 > \varepsilon_{\mathrm{KS}}$, as μ decreases from its maximum value $6/7$, $\Sigma(\mu)$ is first positive, but then becomes negative with a minimum for $\mu \approx 0.6$, implying $\varepsilon_{\mathrm{r}} > 0.253$. This behavior is typical of a first order phase transition. For $\varepsilon = 0.256$, $\Sigma(\mu)$ is always positive, and one cannot draw any conclusion.

completely analogous. Given k, q and ϵ, we compute $\Sigma(\mu)$, and minimize it numerically for $\mu \in [0, 1]$. The largest value (more precisely, the supremum) of ε such that the minimum value is negative, is denoted by $\varepsilon_{\mathrm{var}}(k, q)$. According to conjecture 1, we expect $\varepsilon_{\mathrm{r}}(k, q) \geq \varepsilon_{\mathrm{var}}(k, q)$. Although we have proved it only for frustrated kernels (which do not include the ferromagnetic Potts channel), we shall loosely use the term 'variational bound' also in the other cases.

One can show that the variational bound is always at least as good as the KS one: $\varepsilon_{\mathrm{var}}(k, q) \geq \varepsilon_{\mathrm{KS}}(k, q)$ by looking at the behavior of $\Sigma(\mu)$ near to $\mu = (1 - 1/q)$. By Taylor expanding $\Sigma(\mu)$ for $\mu = (1 - 1/q) + \delta\mu$, we obtain $\Sigma(\mu) = c_{k,q}(\varepsilon)\delta\mu^4 + O(\delta\mu^5)$. Furthermore $c_{k,q}(\varepsilon) < 0$ for $\varepsilon < \varepsilon_{\mathrm{KS}}(k, q)$ and $c_{k,q}(\varepsilon) > 0$ for $\varepsilon > \varepsilon_{\mathrm{KS}}(k, q)$. In Fig. 3 we plot $\Sigma(\mu)$ for the ferromagnetic Potts channel with $k = 2, q = 7$, showing that the variational bound $\varepsilon_{\mathrm{var}}(k, q)$ is strictly larger than the KS one. We shall discuss in the next section for which values of k, q this happens. If the variational principle were proved for the ferromagnetic channel, this would prove $\varepsilon_{\mathrm{r}}(k, q) > \varepsilon_{\mathrm{KS}}(k, q)$ in these cases.

Let us notice that we do not expect the variational lower bound to be tight. More precisely, even minimizing it over the space of distributions over $\mathfrak{M}_q$, $\min \Sigma(\widehat{Q})$ becomes negative only below a threshold $\varepsilon_{\mathrm{c}}(k, q)$ with $\varepsilon_{\mathrm{KS}}(k, q) < \varepsilon_{\mathrm{c}}(k, q) < \varepsilon_{\mathrm{r}}(k, q)$ (in the case where $\varepsilon_{\mathrm{KS}}(k, q) < \varepsilon_{\mathrm{r}}(k, q)$). Our numerical simulations confirm this expectation which is motivated by the physical interpretation of the complexity.

4. THRESHOLDS FOR THE FERROMAGNETIC POTTS CHANNEL

In order to determine reconstruction thresholds numerically, we simulate the recursion (9), by representing the distributions $Q_x^{(\ell)}$ through a large enough sample. We will estimate reconstruction to be possible if the sample does not concentrate, for ℓ large around the point $\overline{\eta}$.

This procedure is very similar to the 'population dynamics' method used to solve similar equations in spin glass theory.[27,19] We work with q samples ('populations') $P_1^{(\ell)}, \dots P_q^{(\ell)}$, each containing M points $\eta_i \in \mathfrak{M}_q$, $i \in \{1, \dots, M\}$ (i.e. M vectors $\eta_i(x)$, $x \in \{1, \dots, q\}$ with $\eta_i(x) \geq 0$ and $\sum_x \eta_i(x) = 1$). The population $P_x^{(\ell)}$ represents an i.i.d. sample from the distribution $Q_x^{(\ell)}$. The population $P_x^{(\ell+1)}$, $x \in \{1, \dots, q\}$ is computed, for each $\ell \geq 0$ as follows.

- Choose k iid colors $x_1, \dots, x_k$ with distribution $\pi(\,\cdot\,|x)$.
- Choose k vectors $\eta_1, \dots \eta_k$, with η_i uniformly random in P_{x_i}.
- Compute $\eta = \mathsf{F}(\eta_1, \dots, \eta_k)$ according to (8).
- Store this new η in the population $P_x^{(\ell+1)}$, and repeat until the population contains M elements.

This whole cycle is repeated until the populations $P_x^{(\ell)}$ become stationary (by this we mean that their moments no longer depend on ℓ) within some prescribed accuracy.

Reconstructibility can be monitored by computing the parameters I_ℓ and Ψ_ℓ on in the populations $P_x^{(\ell)}$. If $I_\ell, \Psi_\ell \to 0$ as $\ell \to \infty$, then we estimate that reconstruction is not possible. If they instead converge to a finite value, we take this value as an estimate of I_∞, Ψ_∞. Figure 4 shows an example of such a calculation.

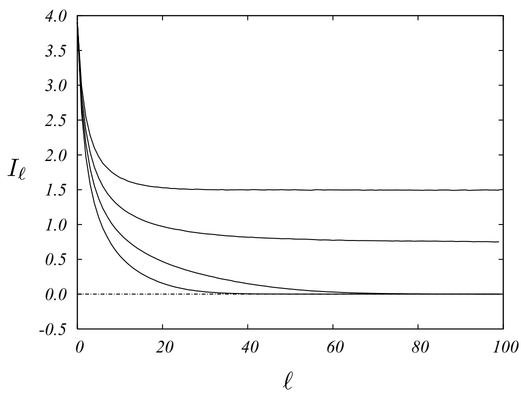

Fig. 4. The information (in bits) that can be transmitted reliably through a k-ary tree network of q-ary symmetric channels (ferromagnetic Potts channels), as determined with the population dynamics algorithm. Here $k = 2, q = 15$ and the noise parameter is (from top to bottom) $\varepsilon = 0.333082, 0.308057,$ $0.282444, 0.256422$. We used populations of size $M = 10^5$, and averaged over 10 runs.

Table I. Thresholds (numerical results and bounds) for the ferromagnetic Potts channel

q	k	ε_{r}	$\varepsilon_{\mathrm{KS}}$	$\varepsilon_{\mathrm{var}}$	$\varepsilon_{\mathrm{alg}}$	$\varepsilon_{\mathrm{MP}}$	I_*	Ψ_*
5	2	0.2348 (1)	0.2343146	0.23491	—	0.30264	0.052 (5)	0.0152 (16)
5	3	0.33881 (5)	0.3381198	0.33887	0.19047	0.41712	0.06 (2)	0.016 (4)
5	4	0.4008 (1)	0.4	0.40081	0.29046	0.48	0.06 (1)	0.020 (4)
5	7	0.4986 (1)	0.4976284	0.49847	0.41114	0.57143	0.07 (1)	0.020 (4)
5	15	0.5955 (1)	0.5934409	0.59422	0.53965	0.65238	0.14 (1)	0.040 (8)
7	2	0.25432 (5)	0.2510513	0.25369	—	0.34577	0.14 (1)	0.028 (4)
7	4	0.43325 (5)	0.4285714	0.43250	0.30769	0.53909	0.195 (5)	0.045 (2)
10	2	0.2716 (2)	0.2636039	0.26977	—	0.38325	0.23 (2)	0.040 (5)
15	2	0.2881 (1)	0.2733670	0.28472	—	0.41652	0.37 (3)	0.053 (4)

Note. The reconstruction threshold ε_{r}, whose numerical estimate is shown in the first column, satisfies the rigorous bounds $\varepsilon_{\mathrm{r}} \geq \varepsilon_{\mathrm{KS}}$, $\varepsilon_{\mathrm{r}} \geq \varepsilon_{\mathrm{alg}}$, and $\varepsilon_{\mathrm{r}} \leq \varepsilon_{\mathrm{MP}}^{-}$. The 'algorithmic bound' ϵ_{alg} is computed by analyzing reconstruction through recursive majority along the lines of Ref. 16. The variational principle that is not proven for this ferromagnetic channel would imply $\varepsilon_{\mathrm{r}} \geq \varepsilon_{\mathrm{var}}$. The symbol – means that the corresponding bound does not provide any information.

Reconstructibility thresholds are determined by repeating the same experiment for several values of the channel noise ε.

Numerical simulations clearly show that the reconstructibility and Kesten Stigum threshold coincide for $q = 3$ and $q = 4$. We checked this to be the case for $q = 3$ and $k = 2$–$7, 10, 15, 20, 30, 50$, and $q = 4$ and $k = 2, 3, 5, 10, 15, 30$ and expect it to be the case generically, at least for k not too large. When this is the case, the order parameters I_∞, Ψ_∞ decrease continuously and vanish at $\varepsilon_{\mathrm{r}}(k, q) = \varepsilon_{\mathrm{KS}}(k, q)$.

For $q \geq 5$ we always find $\varepsilon_{\mathrm{r}}(k, q) > \varepsilon_{\mathrm{KS}}(k, q)$. In these cases $I_\infty(\varepsilon) \downarrow I_* > 0$, $\Psi_\infty(\varepsilon) \downarrow \Psi_* > 0$ as $\varepsilon \uparrow \varepsilon_{\mathrm{r}}(k, q)$. In spin glass language, the transition is discontinuous: we refer to next Sec. for some illustrations. We report our numerical results in Table I. This table also contains the variational lower bound $\varepsilon_{\mathrm{var}}(k, q) \leq \varepsilon_{\mathrm{r}}(k, q)$, as well as the upper bound derived in Ref. 21: $\varepsilon_{\mathrm{r}}(k, q) \leq \varepsilon_{\mathrm{MP}}^{+}(k, q)$, where

$$\varepsilon_{\mathrm{MP}}^{\pm}(k, q) = (q-1)\frac{(2-q+2kq) \mp \sqrt{(2-q+2kq)^2 - 4k(k-1)q^2}}{2kq^2}. \quad (34)$$

In Fig. 5 we plot the thresholds as a function of q for $k = 2$.

5. THRESHOLDS FOR THE ANTIFERROMAGNETIC POTTS CHANNEL

In Table II we present our numerical results for the reconstruction thresholds of the antiferromagnetic Potts channel in the cases in which it differs from $\varepsilon_{\mathrm{KS}}$, together with the bounds.

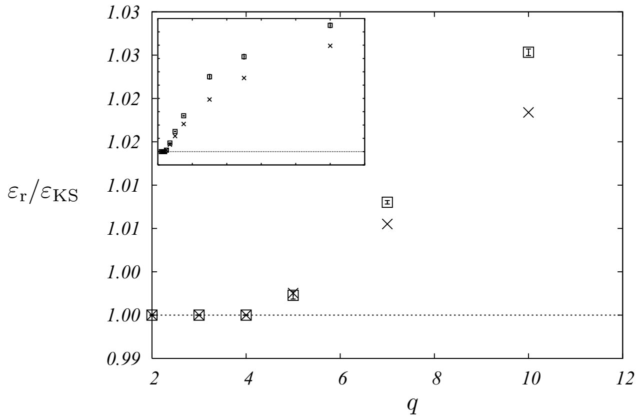

Fig. 5. Ratio between the reconstructibility and the KS thresholds $\varepsilon_{\rm r}(k, q)/\varepsilon_{\rm KS}(k, q)$, for the ferromagnetic Potts channel and $k = 2$. *Squares* correspond to the numerical determination of $\varepsilon_{\rm r}(k, q)$ and *crosses* to the variational lower bound $\varepsilon_{\rm var}(k, q)$. The inset refer to larger number of colors (up to 100).

One distinctive feature of this channel is that, even in the limit $\varepsilon \to 1$ reconstruction may be impossible. For any given $q \geq 3$ reconstruction becomes possible only for $k \geq k_*(q)$. Numerically we found $k_*(3) = 5$, $k_*(4) = 8$, $k_*(5) = 13$, $k_*(6) = 17$. In fact the case $\varepsilon = 1$ has a special interest. In this case the configuration produced by the broadcast process is distributed according to the free boundary Gibbs measure for proper colorings of the (infinite) tree $\mathbb{T}_k$. Our numerical results imply that this measure is extremal only for $k < k_*(q)$, with $k_*(q)$ as above. Using the variational principle (which in this case is proved, cf. Proposition 3), we can show that $k_*(3) \leq 5$, $k_*(4) \leq 9$, $k_*(5) \leq 13$, $k_*(6) \leq 17\dots$

For $q = 3$ we found the reconstructibility threshold to coincide always with the KS threshold. This was checked for $k = 4$–7, 10, 20. The parameters I_∞ and Ψ_∞ are continuous functions of ε vanishing at $\varepsilon_{\rm KS}$. An example is provided in Fig. 6, left frame. For $q \geq 4$ and $k \geq k_*(q)$ the transition at the reconstructibility threshold is discontinuous, cf. Fig. 6, right frame. Table II gives the values of I_*, Ψ_* and $\Sigma_* \equiv \lim_{\varepsilon\to\varepsilon_r} \Sigma(\widehat{Q}^{(\infty)})$ (in the ferromagnetic channel, this number is so small that it cannot be measured reliably in the numerics). Most of the remarks made for the ferromagnetic channel apply to this case.

6. RELATION TO SPIN GLASS THEORY

In this section we explore the link between reconstruction and spin glass theory. For simplicity we keep to the Potts channel, but the discussion can be easily generalized. Consider a configuration $\underline{Y}^L$ generated by the broadcast along a finite rooted tree $\mathbb{T}_k(L)$ with L generations, starting from a uniformly random symbol

Table II. Thresholds (numerical results and bounds) for the antiferromagnetic Potts channel

q	k	ε_{r}	$\varepsilon_{\mathrm{KS}}$	$\varepsilon_{\mathrm{var}}$	$\varepsilon_{\mathrm{alg}}$	$\varepsilon^{-}_{\mathrm{MP}}$	I_*	Ψ_*	Σ_*
4	8	0.99953 (4)	—	—	—	0.91552	1.56(4)	0.56 (1)	0.026 (3)
4	9	0.9908 (4)	1	0.99298	—	0.90717	1.31 (2)	0.47 (2)	0.009 (1)
4	10	0.9820(8)	0.9871708	0.98304	—	0.9	1.2 (2)	0.42 (4)	0.005 (4)
4	11	0.9725(3)	0.9761335	0.97363	0.99736	0.89376	1.07 (5)	0.39 (1)	$\lesssim 0.005$
4	12	0.9643(3)	0.9665063	0.96498	0.98946	0.88826	0.26 (3)	4.2 (5)	$\lesssim 0.005$
4	15	0.9431(3)	0.9436492	0.94338	0.96903	0.875	0.5 (1)	0.16 (3)	$\lesssim 0.001$
4	18	0.9267(2)	0.9267766	0.92686	0.95264	0.86502	0.3 (1)	0.11 (4)	$\lesssim 0.001$
5	13	0.99741(5)	—	0.99982	—	0.92308	1.76 (4)	0.59 (1)	0.042 (5)
5	14	0.9932(1)	—	0.99555	—	0.91916	1.7 (1)	0.54 (2)	0.03 (1)
5	15	0.9888(1)	—	0.99092	—	0.91561	1.48 (5)	0.48 (2)	0.03 (1)
5	20	0.9685(3)	0.9788854	0.96991	0.98581	0.90177	1.1 (5)	0.36 (2)	0.01 (1)
6	17	0.999924(5)	—	—	—	0.93482	2.20 (4)	0.667 (15)	0.095 (5)
6	20	0.9932(3)	—	0.99546	—	0.92792	1.87 (6)	0.569 (15)	0.04 (2)

Note. The reconstruction threshold ε_{r}, whose numerical estimate is shown in the first column, satisfies the rigorous bounds $\varepsilon_{\mathrm{r}} \leq \varepsilon_{\mathrm{KS}}$ (from Ref. 12), $\varepsilon_{\mathrm{r}} \leq \varepsilon_{\mathrm{alg}}$ (cf. Ref. 16), $\varepsilon_{\mathrm{r}} \leq \varepsilon_{\mathrm{var}}$ (from Proposition 3), and $\varepsilon_{\mathrm{r}} \geq \varepsilon^{-}_{\mathrm{MP}}$ (from Ref. 21). The symbol – means that the corresponding bound does not provide any information.

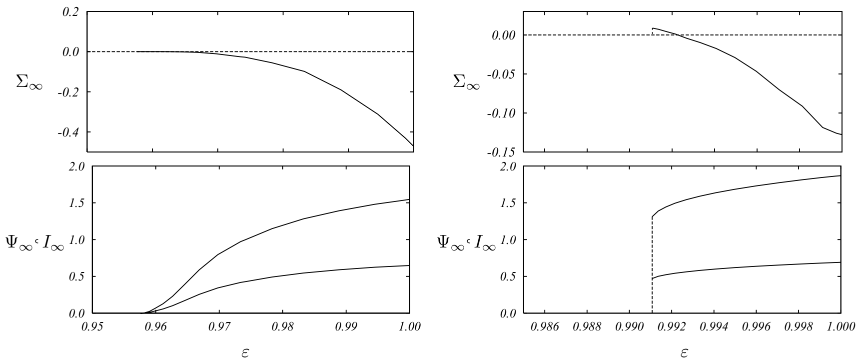

Fig. 6. Asymptotic complexity Σ_∞, information capacity I_∞ and conditional variance Ψ_∞ as a function of the noise parameter for the antiferromagnetic Potts channel. On the left: typical continuous reconstructibility transition, $q = 3, k = 6$. On the right: typical discontinuous transition, $q = 4, k = 9$.

in $\{1, \ldots, q\}$ at the root. As we saw in the introduction, $\underline{Y}^L$ is an equilibrium configuration of the Potts model with free boundary conditions on $\mathbb{T}_k(L)$, i.e. is distributed according to the Boltzmann law for the energy function (3). The coupling J of this model is given by $e^{-\beta J} = \frac{\varepsilon}{(q-1)(1-\varepsilon)}$, and is ferromagnetic (resp. antiferromagnetic) if $J > 0$ (resp $J < 0$).

Once the broadcast process has fixed the variables on the boundary at distance L from the root, the reconstruction problem can be phrased in terms of the conditional distribution $\mathbb{P}\{X_0 = x | \underline{X}_L = \underline{x}_L\}$. The distribution of the first $L - 1$ generations given the received symbols, $\mathbb{P}\{\underline{Y}^{L-1} | \underline{X}_L = \underline{x}_L\}$, is also given by Boltzmann law for the energy function (3). However the boundary condition is now given by the received symbols. One fundamental reason why reconstruction is related to a spin glass problem is that this boundary condition tends to frustrate the system, in the sense of creating conflicting constraints.

It is well known that on trees, frustration comes only through the choice of boundary conditions. Here we discuss the spin glass phase induced in the Potts model on the tree by various possible choices. We shall first show how a 'naive' choice of boundary conditions leads to a simple replica symmetric recursion relation. Then we show how well-chosen self-consistent boundary conditions lead to the correct 1RSB fixed point equation of reconstruction. Finally we discuss the explicit realization of the corresponding spin glass model as a model of Potts spins on a random graph (not a tree).

6.1. Boundary Conditions with Independent Spins

As before, we call $\underline{X}_L$ the set of all spins in the Lth generation. A boundary condition (BC) is a probability distribution on these spins. One first possibility is when spins on the boundary are independent random variables: a given spin X_i, $i \in V_L$ takes value x_i with probability $\eta_i(x_i)$. The overall distribution is therefore $\prod_{i \in V_L} \eta_i(x_i)$. Once a set of $\eta_i(\cdot)$'s is given, the partition function of the tree is obtained as:

$$Z_L(\{\eta_i\}_{i \in V_L}) = \sum_{\underline{y}^L} \prod_{(i,j) \in \mathbb{T}_k(L)} \pi(x_i, x_j) \prod_{i \in V_L} \eta_i(x_i), \tag{35}$$

and the Boltzmann distribution is

$$\mathbb{P}^L_{\{\eta_i\}}(\underline{y}^L) = \frac{1}{Z_L(\{\eta_i\}} \prod_{(ij) \in \mathbb{T}_k(L)} \pi(x_i, x_j) \prod_{i \in V_L} \eta_i(x_i). \tag{36}$$

We have still the freedom of chosing the $\eta_i(.)$. One simple possibility would be to take them identical: $\eta_i(\cdot) = \overline{\eta}(\cdot)$,[10,26,25] but in order to have a disordered and frustrated problem one can choose to sample the η_i's independently from a *symmetric* distribution $P^{(0)}(\eta)$. This definition of a spin glass problem on a tree was adopted for instance in Ref. 7 in the Ising case ($q = 2$), where each of the boundary spins was fixed to ± 1 independently. In our formulation, this corresponds to the choice $P^{(0)}(\eta) = \frac{1}{2}\left(\delta\left[\eta(\cdot), \delta_{+1}\right] + \delta\left[\eta(\cdot), \delta_{-1}\right]\right)$.

The recursive procedure for merging rooted trees applies in the same way as in Sec. 2.2. Consider the marginal distribution on the first $L - \ell \leq L$ generations, $\mathbb{P}(\underline{y}^{L-\ell})$. It is clear that this has the same form as in Eq. (36), with some new $\eta'_i(\cdot)$, $i \in V_{L-\ell}$. When one generate BCs randomly as described above, the η'_i are iid random variables with common distribution $P^{(\ell)}(\eta)$. A little thought shows that $P^{(\ell)}(\eta)$ is related to the one in the shell just above by:

$$P^{(\ell+1)}(\eta) = \int \delta\left[\eta - \mathsf{F}(\eta_1, \ldots, \eta_k)\right] \prod_{i=1}^{k} dP^{(\ell)}(\eta_i). \tag{37}$$

Notice that in each shell, $P^{(\ell)}$ is symmetric.

The marginal distribution at the root of the tree $\mathbb{T}_k(L)$, under the Boltzmann law (36), is a random variable with distribution $P^{(L)}(\eta)$. In this model, the existence of a spin glass phase is characterized by a non trivial limit of $P^{(L)}(\eta) \to P^{(\infty)}(\eta)$ as $L \to \infty$. Such a limit solves a fixed point equation corresponding to (37).

The reader will notice that Eq. (37) is similar to the reconstruction Eq. (17), with one crucial difference: the 'reweighting' factor $z(\{\eta_i\})$ in (17) is absent here. In the spin-glass jargon, Eq. (37) is the 'replica symmetric' (RS) equation, while Eq. (17) is the 1RSB equation with Parisi parameter $m = 1$.

We want to argue that the model defined by Eq. (36), with iid η_i's is not a 'good' model of spin glass on a tree. Technically this is seen from the fact that its glass phase is a RS one, while the spin glass models on graphs with loops typically show RSB. Fundamentally, the drawback of this model is precisely that it neglects correlations between spins on the boundary. As we will discuss in Sec. 6.3, such correlations are necessary in order to study the existence of many pure states, a distinguished mark of spin glasses on graphs with loops.

A minimalistic way of introducing such correlations is to keep uniquely those correlations induced by the tree itself; this is precisely what is done in reconstruction, as we now discuss.

6.2. Self-Consistent Boundary Conditions

It is clear that the broadcast/reconstruction process generates correlated BCs. By this we mean that the conditional distribution $\mathbb{P}(\underline{Y}^{L-1} = \underline{y}^{L-1} | \underline{X}_L)$ has still the form (36) but the η_i are no longer independent. More explicitely, the η_i's, $i \in V_L$ have distribution:

$$\mathbb{P}(\{\eta_i\}_{i\in V_L}) = \frac{1}{\Xi_L}\, Z_L(\{\eta_i\}_{i\in V_L}) \prod_{i\in V_L} \widetilde{Q}^{(0)}(\eta_i)\,, \tag{38}$$

where $Z(\{\eta_i\})$ is the partition function (35) of the tree $\mathbb{T}_k(L)$ with BC $\{\eta_i\}$, and $\widetilde{Q}^{(0)}(\eta)$ is the uniform distribution on the q 'corners' of the simplex $\eta(x) = \delta_{x,r}$, $r \in \{1, \dots, q\}$.

We can analyze the system with BC (38) as we did in the previous section for uncorrelated BCs. Consider, as before, the marginal distribution of the first $L - \ell \leq L$ generations of the tree. It also has the form (36) and the new η_i's at distance ℓ retain the same correlation structure. Their joint distribution is

$$\mathbb{P}(\{\eta_i\}_{i\in V_\ell}) = \frac{1}{\Xi_\ell}\, Z_\ell(\{\eta_i\}_{i\in V_\ell}) \prod_{i\in V_\ell} \widetilde{Q}^{(\ell)}(\eta_i). \tag{39}$$

Finally, the distribution $\widetilde{Q}^{(\ell+1)}$ is related to $\widetilde{Q}^{(\ell)}$ through the recursion (17) that we found when discussing reconstruction, with the correct reweighting factor. Since we took $\widetilde{Q}^{(0)}(\eta) = \widehat{Q}^{(0)}(\eta)$, this implies $\widetilde{Q}^{(\ell)}(\eta) = \widehat{Q}^{(\ell)}(\eta)$ for any $\ell \geq 0$.

It is interesting to define the spin glass problem on the tree associated with the non-trivial fixed point of Eq. (18). This just amounts to generating the BCs from (38) using $\widetilde{Q}^{(0)} = \widehat{Q}^*$. This problem has the virtue of being statistically translation invariant[5] (although, for any given realization, the resulting Gibbs measure is not translation invariant). In particular the properties of a spin don't depend on its distance to the root.

[5] Provided one replaces the rooted tree (with a root of degree k) with a regular Cayley tree (with all the vertices of degree $k + 1$).

6.3. Spin Glass on the Bethe Lattice

While the previous definition of a spin glass on a tree is perfectly correct, it is clear that a lot of the physics has been put into the choice of the BC distribution. This is necessary because of the crucial role of BCs on trees. An alternative definition of the Bethe lattice spin glass, proposed in Ref. 19, is to use, instead of trees, random graphs which have a tree structure on finite length scales. Let us consider for instance the problem of N Potts spins on the vertices of a random regular graph $\mathcal{G}_N$ with degree $k+1$, with pairwise interactions given by the kernel $\pi(x, y)$. The partition function of such a model is

$$Z = \sum_{x_1,\dots,x_N} \prod_{(i,j)\in\mathcal{G}_N} \pi(x_i, x_j) \tag{40}$$

In any finite neighborhood of a randomly chosen node i, the local structure[6] of $\mathcal{G}_N$ is (with high probability) the one of a regular tree with degree $k+1$. In fact, the shortest loop through i is typically of size $\log N$ which diverges when $N \to \infty$. This setting is interesting for two reasons: (i) Loops, although large, can create some frustration; (ii) The system is approximately homogeneous (unlike on a regular tree, where vertices on the boundary have a neighborhood very different from the others).

Spin glasses on random lattices with a local tree-like structure have been the object of many studies in recent times. The cavity method of Refs. 19 and 20 is an iterative procedure which exploits the tree-like structure. Here we shall just mention some of its main results without justification, the aim being to clarify the correspondence between the spin glass model on a random graph and the reconstruction problem.

In the cavity method one first considers the graph, rooted in a site i, obtained by cutting the edge (i, j) to one of its neighbors, cf. Fig. 7. The marginal distribution of the root, $\eta_{i\to j}(x_i)$, with respect to the model on the 'amputated' graph, is then written in terms of the distributions $\eta_{l\to i}(x_l)$ where l are the neighbors of i different from j. It is possible to write such a recursion only if the variables x_l, in the absence of the edges (l, j), become uncorrelated in the large N limit. Such a property is made possible by the local tree-like structure, but also requires a fast decay of correlations in the graph. This is expected to happen either when the system admits a single Gibbs state, or when the Boltzmann measure is restricted to a pure (extremal) Gibbs state.[7] In the first case, the problem is described by a unique distribution $\eta(x)$, and $\eta_{i\to j} = \eta$ for all directed edges $i \to j$. This distribution is a fixed point of (8), satisfying thus $\eta = \mathsf{F}(\eta, \dots, \eta)$. It is called the 'paramagnetic'

[6] By this we mean the subgraph within any fixed distance from i. The property described here can also be phrased in terms of *local weak convergence.*[1]

[7] The definition of extremal Gibbs state on a *finite* graph goes beyond the scope of this paper.

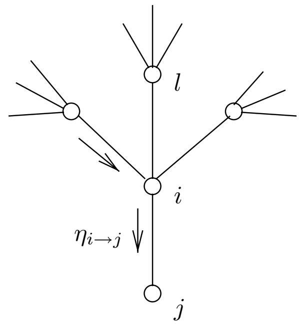

Fig. 7. The basic cavity recursion.

or 'liquid' phase. In the case where there exist several pure states, the recursion holds when the measure is restricted to one pure state α: on a given (large) graph one could thus generate a set of 'messages' $\eta^{\alpha}_{i\to j}(x_i)$ for each state α. Notice that, for a given α, the messages now depend explicitely on the edge: the measure is no longer uniform, but it is modulated. The 1RSB cavity method assumes that there exist exponentially many such pure states, the number $\mathcal{N}(f)$ of states with free energy density $F^{\alpha}/N = f$ is written in terms of the complexity function $\Sigma(f)$ as $\mathcal{N}(f) = \exp(N\Sigma(f))$. In such a case one can perform a statistics in the space of pure states, by introducing, for each edge (i, j), the probability $R_{i\to j}(\eta)$ that the message $\eta^{\alpha}_{i\to j} = \eta$, when α is chosen randomly with a weight proportional to the total Boltzmann weight of state α. After performing this average over states the various edges become again equivalent, and one finds that the distribution $R_{i\to j}(\eta) = \widehat{Q}^*$ satisfies exactly the 1RSB fixed point Eq. (18). So there exists a 1RSB glass phase if and only if this equation has a non trivial symmetric solution. Notice that this equation can also have other non-symmetric solutions. For instance in the case of the ferromagnetic Potts channel, at low enough temperature there is a solution where Q^* is peaked on a η with a ferromagnetic bias, but it does not satisfy the symmetry property that we impose for the study of the glass state. In such a system the glass solution exists, but it is not realized on a random graph: the system will transit to a ferromagnetic phase. On the contrary in some other cases the glass phase will be realized. For instance we expect this to be the case for the antiferromagnetic Potts model on the random graph.[4,22]

The tree reconstruction problem on the one hand, and the spin glass on a random graph on the other, thus naturally lead to the same equations. Some aspects of this correspondence call for a better understanding. Consider the model on the random graph defined in Eq. (40). Let us suppose that it has several pure states, and that the 1RSB cavity solution of the problem is correct. Now isolate around an arbitrary point the set of all its neighbors up to distance ℓ. Generically it is a tree $\mathbb{T}_k(\ell)$. The vertices outside this tree create some boundary condition

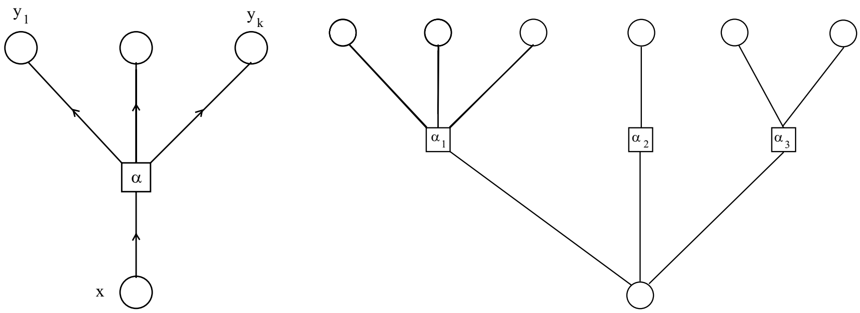

Fig. 8. *Left*: A function node representing the 'one-to-k' channel $\pi^{(\alpha)}(y_1, \ldots, y_k|x)$. *Right*: An example of a small random tree network, with $l_0 = 3$.

on the leaves of this tree, depending on the pure state α that we are considering. We have found that the statistics of these BC on the pure states corresponds to the statistics of the boundaries in the broadcast/reconstruction, and both are described by the distribution $\widehat{Q}^*$. In spin glass theory (within 1RSB) one can count the pure states through the computation of the complexity function $\Sigma(f)$. It would be very interesting to have an interpretation of this function in terms of the reconstruction problem.

7. GENERALIZATIONS

So far we have focused on k-ary trees whose links corresponds to identical copies of the same q-ary channel satisfying the symmetry condition $\pi(y|x) = \pi(x|y)$. However, none of these hypotheses is crucial to our approach. In this section we define a considerably more general context, and sketch how to adapt the above formalism to this case. Some cases of broadcast through non regular trees, or with asymmetric channels have been considered for instance in Refs. 9, 14, 17, 21. The present formalism encompasses all these cases and generalizes them to broadcast through hypergraphs.

We consider a finite set of kernels $\{\pi^{(\alpha)}(\,\cdot\,|\,\cdot\,)\,;\ \alpha = 1, \ldots, n\}$, each kernel describing a one-to-many communication channel. For each $x \in \{1, \ldots, q\}$, and each k_α-uple $y_1, \ldots, y_{k_\alpha}$, $\pi^{(\alpha)}(y_1, \ldots, y_{k_\alpha}|x)$ gives the probability that users $1, \ldots, k_\alpha$ receive outputs $y_1, \ldots, y_{k_\alpha}$ if the channel input was x. These kernels must satisfy the conditions

$$\pi^{(\alpha)}(y_1, \ldots, y_{k_\alpha}|x) \geq 0, \qquad \sum_{y_1,\ldots,y_{k_\alpha}} \pi^{(\alpha)}(y_1, \ldots, y_{k_\alpha}|x) = 1, \tag{41}$$

and x is called the parent of $y_1, \ldots, y_{k_\alpha}$. Such 'one-to-$k$' communication channels can be represented graphically using factor nodes of degree $k + 1$, cf. Fig. 8.

Next, we define a *random* tree network ensemble depending on two probability distributions q_α, $\alpha \in \{1, \dots, n\}$ ($q_\alpha \geq 0$, $\sum q_\alpha = 1$) and p_l, $l \geq 0$ ($p_l \geq 0$, $\sum p_l = 1$). One (infinite) random network $\mathbb{T}$ from this ensemble is generated as follows starting from the root 0.

- Draw an integer l_0 with distribution p_l. This is the degree of the vertex 0
- For each $a \in \{1, \dots, l_0\}$, draw α_a independently with distribution q_α, and attach a channel of type α_a to the vertex 0. The root will transmit through such channels.
- For each $a \in \{1, \dots, l_0\}$, and each $i_a \in \{1, \dots, k_{\alpha_a}\}$, associate a vertex to the i_a-th output of channel a. Repeat the above construction for each of these vertices.

In Fig. 8 we show a small example of such a network. We denote by $\mathbb{T}(i)$ the random (sub)network rooted at i, by $\mathbb{T}(i, \ell)$ its first ℓ generations (starting from i), and by $\underline{X}_{i,\ell}$ the received colors, ℓ generations above i ($\underline{X}_{0,\ell} \equiv \underline{X}_\ell$).

The network is used to communicate. A color $x_0 \in \{1, \dots, q\}$ is chosen at the root with probability $\varphi_0(x_0)$ and broadcast through the l_0 channels connected to the root itself. Each of the first generation vertices receives a corrupted version $x_i \in \{1, \dots, q\}$ of this color, with joint distribution

$$\prod_{a=1}^{l_0} \pi^{(\alpha)}(x_{a,1}, \dots, x_{a,i_a} | x_0), \tag{42}$$

where (a, r) denotes the r-th output vertex of the a-th channel. In other words distinct channels act independently. The same transmission process is repeated at the first generation and so forth, through the entire network. The problem is to reconstruct the transmitted color from the output at generation ℓ, denoted as $\underline{x}_\ell$.

Analogously to the case investigated in the previous sections, we say that the reconstruction problem is solvable if the conditional mutual information $I(X_0; \underline{X}_\ell | \mathbb{T})$ does not vanish as $\ell \to \infty$. Equivalently, the problem is solvable if there is a reconstruction procedure which succeeds with probability strictly larger than $\max_{x_0} \varphi_0(x_0)$ in the $\ell \to \infty$ limit. In these definitions we assume the network structure to be known at the receiver (this is why we consider a mutual information which is *conditional* to this structure).

Unlike for q-ary reversible channels, the distribution of the color X_i received at vertex i, is not uniform and depends upon the vertex. We shall denote it by $\varphi_i(x_i) \equiv \mathbb{P}[X_i = x_i | \mathbb{T}]$. In fact $\varphi_i(\,\cdot\,)$ can be determined recursively: if j is the parent of i, and i_1, i_{k-1} the other vertices that share this parent, then

$$\varphi_i(x_i) = \sum_{x_j} \sum_{x_{i_1}, \dots, x_{i_{k-1}}} \pi^{(\alpha)}(x_i, x_{i_1}, \dots, x_{i_{k-1}} | x_j)\, \varphi_j(x_j)\,. \tag{43}$$

Let us consider the reconstruction problem. The conditional distribution of the transmitted color given the observation at generation ℓ, $\mathbb{P}[X_0 = x|\underline{X}_\ell = \underline{x}_\ell]$ can be computed recursively proceeding from the leaves downwards to the root as in Sec. 2.2. In order to simplify the analysis, it is convenient to 'factor out' the *a priori* information $\varphi_i(x_i)$, and define

$$\eta_{i,\ell}(x) = \mathbb{P}_i[X_i = x|\underline{X}_{i,\ell}]. \tag{44}$$

where $\mathbb{P}_i$ denotes probability with respect to a modified process in which the boundary $\underline{X}_{i,\ell}$ is obtained from a broadcast starting from X_i chosen uniformly at random in $\{1, \dots, q\}$. Of course we have

$$\mathbb{P}[X_i = x|\underline{X}_{i,\ell}] = \frac{1}{z_i}\,\varphi_i(x)\,\eta_{i,\ell}(x), \tag{45}$$

where $z_i \equiv \sum_x \varphi_i(x)\eta_{i,\ell}(x)$ ensures the correct normalization. Notice that $\eta_{i,\ell}(x)$ depends uniquely on the portion of the tree above i, more precisely on $\mathbb{T}(i, \ell)$ and $\underline{X}_{i,\ell}$.

With a slight abuse of notation, let us denote by $\pi^{(a)} = \pi^{\alpha(a)}$, $a \in \{1, \dots, l_i\}$ the channels whose input is x_i, and by $j_1, \dots, j_{k(a)}$ the corresponding output vertices. It is easy to derive the following recursion which generalizes Eq. (8)

$$\eta_{i,\ell+1}(x) = \frac{1}{z(\{\eta_{j,\ell}\}, \{\pi^{(a)}\})} \prod_{a=1}^{l_i} \sum_{x_1,\dots x_{k(a)}} \pi^{(a)}(x_1, \dots x_{k(a)}|x)$$
$$\times\, \eta_{j_1,\ell}(x_1) \cdots \eta_{j_{k(a)},\ell}(x_{k(a)}), \tag{46}$$

where ℓ is the distance from the leaves. The constant $z(\{\eta_{j,\ell}\}, \{\pi^{(a)}\})$ is defined by the constraint $\sum_x \eta_{i,\ell+1}(x) = 1$. We shall denote the above mapping synthetically by writing $\eta_{i,\ell+1} = \mathsf{F}(\{\eta_{j,\ell}\}, \{\pi^{(a)}\})$.

One can define two types of probability distributions associated to $\eta_{i,\ell}$. We first assume that the tree $\mathbb{T}$ is given, and consider the distribution of $\eta_{i,\ell}$ conditional to $X_i = x$ and $\mathbb{T}$. This will be denoted by $Q_x^{(i,\ell)}(\eta)$. Arguing as in the case of regular trees, one derives the recursion

$$Q_x^{(i,\ell+1)}(\eta) = \sum_{\{x_j\}} \prod_{a=1}^{l_i} \pi^{(a)}(x_{j_1}, \dots x_{j_{k(a)}}|x) \int \delta[\eta - \mathsf{F}(\{\eta_j\}, \{\pi^{(a)}\})]$$
$$\times \prod_j dQ_{x_j}^{(j,\ell)}(\eta_j). \tag{47}$$

Next, we consider the distribution of $\eta_{i,\ell}$ *unconditional* of $\mathbb{T}$, which we denote by $Q_x^{(\ell)}(\eta)$. This can also be regarded as the expectation of the previous distribution: $Q_x^{(\ell)}(\eta) = \mathbb{E}_{\mathbb{T}} Q_x^{(i,\ell)}(\eta)$. Notice that $Q_x^{(i,\ell)}(\eta)$ depends on $\mathbb{T}$ only through the first ℓ generations of the subtree rooted at i $\mathbb{T}(i, \ell)$. Since the structures of distinct subtrees are independent, if we average Eq. (47) with respect to $\mathbb{T}$, the averages factorize, yielding a recursion equation for $Q_x^{(\ell)}(\eta)$:

$$Q_x^{(\ell+1)}(\eta) = \mathbb{E} \sum_{\{x_j\}} \prod_{a=1}^{l_i} \pi^{(a)}(x_{j_1}, \dots x_{j_{k(a)}}|x) \times \int \delta[\eta - \mathsf{F}(\{\eta_j\}, \{\pi^{(a)}\})] \prod_j dQ_{x_j}^{(\ell)}(\eta_j). \tag{48}$$

Here $\mathbb{E}$ denotes expectation with respect to the degree l_i and the channel types. The last expression is particularly convenient for numerical simulations and it is not more complex than the iteration (9) studied in the previous sections.

We can also consider the distributions unconditional to the transmitted color: $\widehat{Q}^{(i,\ell)}(\eta)$ and $\widehat{Q}^{(\ell)}(\eta)$. The same relation as for regular trees hold in this case $Q_x^{(i,\ell)}(\eta) = q\eta(x)\widehat{Q}^{(i,\ell)}(\eta)$ and $Q_x^{(\ell)}(\eta) = q\eta(x)\widehat{Q}^{(\ell)}(\eta)$. It is easy to derive the corresponding recursions. We just write down the equation for the last (non-random) distribution

$$\widehat{Q}^{(\ell+1)}(\eta) = \mathbb{E} \int \frac{z(\{\eta_j\}\{\pi^{(a)}\})}{z(\{\overline{\eta}\}\{\pi^{(a)}\})} \delta[\eta - \mathsf{F}(\{\eta_j\}, \{\pi^{(a)}\})] \times \prod_j d\widehat{Q}^{(\ell)}(\eta_j). \tag{49}$$

In order to discuss the correspondence with the dynamical glass transition in this more general setting, it is necessary to distinguish two cases. In the simplest one, the RS cavity equations for the associated statistical mechanics model admit the solution $\eta_{i\to j}(x) = \overline{\eta}(x)$ for any directed link $i \to j$ in the graph. Under this hypothesis it is not hard to show that Eq. (49) is equivalent (in the same sense discussed in Sec. 2) to the $m = 1$ 1RSB equation.[8]

In the general case (i.e. if $\eta_{i\to j}(x) = \overline{\eta}(x)$ does not solve the RS equations), the dynamical glass transition still corresponds to the extremality of the free boundary measure on an infinite tree. The last problem, however, cannot be formulated in the same framework as described here. One can still write an equation of the form (47), conditioned to the graph structure, as is usually done in statistical physics. But the average over the graph structure can only be performed conditioning upon the value of $\eta_{i\to j}(\,\cdot\,)$ in the RS solution,[15] and so the relation between the spin glass problem and the reconstruction problem unconditioned to the structure of the tree, is not as simple as before. We shall not enter these details here.

[8] The expert will perhaps be surprised by this remark since it is usually said that the order parameter for such systems is a 'measure over the space of distributions.' However it turns out that, for $m = 1$, the expectation of this measure satisfies an equation which is Eq. (49).

8. CONCLUSION AND OPEN PROBLEMS

The coincidence between the reconstruction threshold and the dynamical glass transition explored in this paper is interesting from several points of view.

First of all, it provides some perspective for putting on a firmer basis the theory of glassy systems on locally tree-like graphs developed in the last few years. As a concrete example, notice that the population dynamics algorithm defined in Sec. 4 presents two important advantages with respect to the procedure adopted by statistical physicists. First, in spin glass theory the usual prescription is to look for a solution of the fixed point Eq. (18). This poses the problem of the initial condition: even if a given initial condition yields a trivial fixed point, this may not be the case for all the initial conditions. In the present formulation the iteration is initialized with a very specific initial condition, and one is guaranteed that, if it converges to a trivial fixed point, no non-trivial fixed point exists. Second, simulating Eq. (18) requires a 'reweighting' of the sample which is usually the trickier part of the calculation. No reweighting is needed in the new approach: Eq. (9) can be handled easily.

Also, it provides some indication on the correctness of simple analytic and algorithmic approaches to these systems, such as the replica symmetric cavity method, or the belief propagation algorithm. While it has long been known that their correctness should be related to a fast correlation decay in a model on a tree, a precise criterion has never been formulated.[9] Ous work suggests that the extremality of the associated Gibbs measure on a tree provides such a criterion.

More broadly, it illustrate some subtleties of the physics of the glass transition. It is well known that, in a 1RSB glass transition, point to point correlations (static scattering factors) do not present any diverging correlation length. This paper shows that such a length can be derived quite generally from point-to-set correlations. Indeed the definition considered here is essentially equivalent to the one of Ref. 3. It was shown in Ref. 23 that this length scale divergence implies a lower bound on the time scale divergence.

Finally, statistical physics ideas can inspire new results on the original reconstruction problem. The most interesting such idea is, in our view, the complexity functional introduced in Sec. 3. Apart from being conceptually innovative with respect to classical techniques, it seems to provide by far the best rigorous quantitative estimates of the reconstruction threshold, cf. Proposition 3. This is well illustrated by the results in Table II. It will certainly be interesting to find a concrete

[9] A frequently used sufficient criterion is the uniqueness of the Gibbs state on the infinite tree (see Ref. 2 and references therein). As it emerges from our discussion, this criterion is often much stronger than needed. It is also interesting to recall that Tatikonda and Jordan[29] first connected the convergence properties of belief propagation to the extremality of the free boundary Gibbs on a properly defined tree.

interpretation of this object in terms of the original reconstruction problem. Also, the values of the channel parameter at which the asymptotic complexity $\Sigma(\widehat{Q}^{(\infty)})$ vanish have a particularly important role in statistical mechanics, but did not find any role here.

There are several results that we have not been able to prove rigorously. We already formulated one such results as Conjecture 1, and proved it for a particular class of channel models as Proposition 3. Another interesting fact which has emerged from our numerical simulations is the coincidence of the KS and reconstruction thresholds when the number of colors is small.

Conjecture 2. *Consider the reconstruction problem for the k-ary tree and the ferromagnetic Potts channel (q-ary symmetric channel) with $q \leq 4$, or the antiferromagnetic Potts channel with $q \leq 3$. Then, there exists a $k_{\max} \geq 30$ such that, if $k < k_{\max}$ the reconstruction threshold coincides with the Kesten-Stigum threshold.*

A stronger version of this conjecture would be to require the thesis to be valid for *all* values of k (i.e. to state that $k_{\max} = \infty$). Although we didn't find any k contradicting this stronger version, this might of course be due to the limitation on the values of k that we can treat numerically.

Finally, let us single out the case of completely antiferromagnetic ($\varepsilon = 1$) Potts channels:

Conjecture 3. *Let $k_*(q)$ be the maximum value of k such that the free boundary Gibbs measure for uniformly random proper colorings on the infinite k-ary tree is extremal. Then $k_*(3) = 5$, $k_*(4) = 8$, $k_*(5) = 13$, $k_*(6) = 17$.*

APPENDIX A: VARIATIONAL PRINCIPLE FOR FRUSTRATED KERNELS

This appendix is devoted to the proof of Lemma 1. It is convenient to introduce some notations. If $A(x, y)$, $x, y \in \{1, \ldots, q\}$ is a symmetric matrix and $\eta_1(x)$, $\eta_2(x)$, $x \in \{1, \ldots, q\}$ are two vectors, we shall write $A\eta_1(x) \equiv \sum_y A(x, y)\eta_1(y)$ and $\eta_1 A\eta_2 \equiv \sum_{x,y} A(x, y)\eta_1(x)\eta_2(y)$. Furthermore, if $\widehat{Q}$ is a distribution over $\mathfrak{M}_q$ we will denote by $\mathsf{T}\widehat{Q}$ the distribution obtained by using Eq. (18): $\mathsf{T}\widehat{Q}$ is the left hand side of Eq. (18) when in the right hand side $\widehat{Q}^*$ has been substituted by $\widehat{Q}$. Finally, given $\eta_1, \eta_2 \in \mathfrak{M}_q$, we let

$$\Delta(\eta_1, \eta_2) \equiv \left(\frac{\eta_1 \pi \eta_2}{\overline{\eta} \pi \overline{\eta}}\right) \log \left(\frac{\eta_1 \pi \eta_2}{\overline{\eta} \pi \overline{\eta}}\right). \tag{A1}$$

We also write $\eta \stackrel{\mathrm{d}}{=} \widehat{Q}$ when η has distribution $\widehat{Q}$. We first derive two simple lemmas.

Lemma 2. *Let $\widehat{Q}^*$ and $\widehat{Q}$ be two consistent distributions over $\mathfrak{M}_q$ and $\Sigma^*(t) \equiv \Sigma((1-t)\widehat{Q}^* + t\widehat{Q})$. Then*

$$-\frac{1}{(k+1)}\left.\frac{d\Sigma^*}{dt}\right|_0 = \mathbb{E}\left\{\Delta(\nu,\eta_2') - \Delta(\nu,\eta_2) - \Delta(\eta_1,\eta_2') + \Delta(\eta_1,\eta_2)\right\}, \quad \text{(A2)}$$

where the expectation is taken with respect to the independent random variables $\eta_1, \eta_2 \stackrel{\mathrm{d}}{=} \widehat{Q}^*$, $\eta_2' \stackrel{\mathrm{d}}{=} \mathsf{T}\widehat{Q}^*$ and $\nu \stackrel{\mathrm{d}}{=} \widehat{Q}$.

Proof: Elementary calculus yields

$$-\frac{1}{(k+1)}\left.\frac{d\Sigma^*}{dt}\right|_0 = \psi(1) - \psi(0), \quad \text{(A3)}$$

where

$$\psi(t) \equiv -\mathbb{E}\left\{\left(\frac{\nu^t\pi\eta}{\overline{\eta}\pi\overline{\eta}}\right)\log\left(\frac{\nu^t\pi\eta}{\overline{\eta}\pi\overline{\eta}}\right)\right\} \quad \text{(A4)}$$
$$+\mathbb{E}\left\{\left(\frac{\sum_x \pi\nu^t(x)\prod_{i=1}^k \pi\eta_i(x)}{\sum_x \prod_{i=0}^k \pi\overline{\eta}(x)}\right)\log\left(\frac{\sum_x \pi\nu^t(x)\prod_{i=1}^k \pi\eta_i(x)}{\sum_x \prod_{i=0}^k \pi\overline{\eta}(x)}\right)\right\}.$$

Here $\eta, \eta_1, \ldots, \eta_k \stackrel{\mathrm{d}}{=} \widehat{Q}^*$ and $\nu^t \stackrel{\mathrm{d}}{=} (1-t)\widehat{Q}^* + t\widehat{Q}$ are independent random variables. The first term, when integrated on t, gives the contribution $\mathbb{E}\{\Delta(\eta_1,\eta_2) - \Delta(\nu,\eta_2)\}$ to (A2).

As for the second term, observing that $\prod_{i=1}^k \pi\eta_i(x) = z(\{\eta_i\})\,\mathsf{F}(\eta_1,\ldots,\eta_k)$, it can be rewritten as

$$q^{k-1}\mathbb{E}\left\{z(\{\eta_i\})\left[\frac{\nu^t\pi\mathsf{F}(\eta_1\ldots\eta_k)}{\overline{\eta}\pi\overline{\eta}}\right]\log\left[\frac{\nu^t\pi\mathsf{F}(\eta_1\ldots\eta_k)}{\overline{\eta}\pi\overline{\eta}}\right]\right\}$$
$$+q^{k-1}\mathbb{E}\left\{z(\{\eta_i\})\left[\frac{\nu^t\pi\mathsf{F}(\eta_1\ldots\eta_k)}{\overline{\eta}\pi\overline{\eta}}\right]\log\left[q^{k-1}z(\{\eta_i\})\right]\right\}. \quad \text{(A5)}$$

Since $\mathbb{E}\nu^t(x) = \overline{\eta}(x)$, the second term is t-independent and does not contribute to (A3). The first term is equal to $\mathbb{E}\Delta(\nu^t,\eta_2')$ where $\eta_2' \stackrel{\mathrm{d}}{=} \mathsf{T}\widehat{Q}^*$. □

Lemma 3. *Let $\eta_1, \eta_2 \in \mathfrak{M}_q$ and define $\delta\eta_i(x) = \eta_i(x) - \overline{\eta}(x)$. If $\pi(x,y) = \pi_* - \widehat{\pi}(x,y)$ is a frustrated kernel, then*

$$|\delta\eta_1\widehat{\pi}\delta\eta_2| \le \overline{\eta}\pi\overline{\eta}. \quad \text{(A6)}$$

Proof: Since $\widehat{\pi}$ is positive definite, $\phi\widehat{\pi}\psi \equiv \sum_{x,y}\widehat{\pi}(x,y)\phi(x)\psi(y)$ is a well defined scalar product. Cauchy-Schwarz inequality implies

$$|\delta\eta_1\widehat{\pi}\delta\eta_2| \le \sqrt{(\delta\eta_1\widehat{\pi}\delta\eta_1)(\delta\eta_2\widehat{\pi}\delta\eta_2)} \le \max\{(\delta\eta_1\widehat{\pi}\delta\eta_1), (\delta\eta_2\widehat{\pi}\delta\eta_2)\}. \quad \text{(A7)}$$

Therefore it is sufficient to prove Eq. (A6) for $\delta\eta_1 = \delta\eta_2 = \delta\eta$. Let $\eta(x) = \overline{\eta}(x) + \delta\eta(x)$. Since $\pi(x, y), \eta(x) \geq 0$, and $\sum_x \delta\eta(x) = 0$, we have

$$0 \leq \eta\pi\eta = \overline{\eta}\pi\overline{\eta} + \delta\eta\pi\delta\eta = \overline{\eta}\pi\overline{\eta} - \delta\eta\widehat{\pi}\delta\eta. \tag{A8}$$

□

We can now turn to the proof of Lemma 1. In the following, given $\eta \in \mathfrak{M}_q$, we define $\delta\eta(x) \equiv \eta(x) - \overline{\eta}(x)$. Obviously we have

$$\Delta(\eta_1, \eta_2) = \left(1 - \frac{\delta\eta_1\widehat{\pi}\delta\eta_2}{\overline{\eta}\pi\overline{\eta}}\right) \log\left(1 - \frac{\delta\eta_1\widehat{\pi}\delta\eta_2}{\overline{\eta}\pi\overline{\eta}}\right). \tag{A9}$$

Because of Lemma 3 we can expand this expression in an absolutely convergent series

$$\Delta(\eta_1, \eta_2) = -\frac{\delta\eta_1\widehat{\pi}\delta\eta_2}{\overline{\eta}\pi\overline{\eta}} + \sum_{n=2}^{\infty} C_n \left(\frac{\delta\eta_1\widehat{\pi}\delta\eta_2}{\overline{\eta}\pi\overline{\eta}}\right)^n, \tag{A10}$$

where $C_n \equiv 1/n(n-1) > 0$. If $\eta_1 \stackrel{\mathrm{d}}{=} \widehat{Q}_1$ and $\eta_2 \stackrel{\mathrm{d}}{=} \widehat{Q}_2$ are independent random variables with consistent distributions, we get

$$\mathbb{E}\Delta(\eta_1, \eta_2) = \sum_{n=2}^{\infty} C_n\, q^n \phi_1^{(n)} \widehat{\pi}^{\otimes n} \phi_2^{(n)}, \tag{A11}$$

where $\widehat{\pi}^{\otimes n}(x_1 \dots x_n; y_1 \dots y_n) \equiv \widehat{\pi}(x_1, y_1) \cdots \widehat{\pi}(x_n, y_n)$ is the n-fold tensor product of $\widehat{\pi}$, $\phi_a^{(n)}(x_1 \dots x_n) \equiv \mathbb{E}\{\delta\eta_a(x_1) \cdots \delta\eta_a(x_n)\}$ are the moments of the distribution $\widehat{Q}_a$, and

$$\phi_1^{(n)} \widehat{\pi}^{\otimes n} \phi_2^{(n)} \equiv \sum_{x_1 \dots x_n} \sum_{x_1 \dots x_n} \widehat{\pi}^{\otimes n}(x_1 \dots x_n; y_1 \dots y_n)$$
$$\times\, \phi_1^{(n)}(x_1 \dots x_n) \phi_2^{(n)}(x_1 \dots x_n). \tag{A12}$$

Now consider Lemma 2 and take $\widehat{Q} = \mathsf{T}\widehat{Q}^*$. We get

$$-\frac{1}{(k+1)} \left.\frac{d\Sigma^*}{dt}\right|_0 = \mathbb{E}\big\{\Delta(\eta_1', \eta_2')$$
$$- \Delta(\eta_1', \eta_2) - \Delta(\eta_1, \eta_2') + \Delta(\eta_1, \eta_2)\big\}, \tag{A13}$$

where $\eta_1, \eta_2 \stackrel{\mathrm{d}}{=} \widehat{Q}^*$ and $\eta_1', \eta_2' \stackrel{\mathrm{d}}{=} \mathsf{T}\widehat{Q}^*$. Applying Eq. (A11) we get

$$-\frac{1}{(k+1)} \left.\frac{d\Sigma^*}{dt}\right|_0 = \sum_{n=2}^{\infty} C_n\, q^n (\phi_{\mathsf{T}}^{(n)} - \phi^{(n)}) \widehat{\pi}^{\otimes n} (\phi_{\mathsf{T}}^{(n)} - \phi^{(n)}), \tag{A14}$$

where $\phi^{(n)}$ and $\phi_{\mathsf{T}}^{(n)}$ denote the moments (respectively) of $\widehat{Q}^*$ and $\mathsf{T}\widehat{Q}^*$. Since $\widehat{\pi}$ is positive definite, $\widehat{\pi}^{\otimes n}$ is positive definite as well and therefore the right hand side is a sum of non-negative terms. In order for this right hand side to vanish, each of the terms must vanish, which implies $\phi_{\mathsf{T}}^{(n)} = \phi^{(n)}$ for each n. But, since $\widehat{Q}^*$ and $\mathsf{T}\widehat{Q}^*$ have bounded support, this implies $\widehat{Q}^* = \mathsf{T}\widehat{Q}^*$, which is false by hypothesis. Therefore the right hand side of Eq. (A14) is strictly positive and $\frac{d\Sigma^*}{dt}\Big|_0 < 0$ as desired.

ACKNOWLEDGMENTS

This work has been carried out mostly during our stay at the MSRI on the occasion of the program on "Probability, Algorithms, and Statistical Physics." We have benefited from numerous discussions, suggestions and comments from James Martin, Elchanan Mossel and Yuval Peres; it is a great pleasure to thank them for these stimulating exchanges. Useful discussions with Olivier Rivoire and Guilhem Semerjian are also gratefully acknowledged.

REFERENCES

1. D. Aldous and J. M. Steele. The objective method: Probabilistic combinatorial optimization and local weak convergence, in H. Kesten, ed., *Probability on Discrete Structures*, New York: Springer (2004).
2. A. Bandyopadhyay and D. Gamarnik, Counting without sampling. New algorithms for enumeration problems using statistical physics. arXiv:math.PR/0510471.
3. J.-P. Bouchaud and G. Biroli, On the Adam-Gibbs-Kirkpatrick-Thirumalai-Wolynes scenario for the viscosity increase in glasses. *J. Chem. Phys.* **121**:7347–7354 (2004).
4. A. Braunstein, R. Mulet, A. Pagnani, M. Weigt and R. Zecchina, Polynomial iterative algorithms for coloring and analyzing random graphs. *Phys. Rev. E* **68**:036702 (2003).
5. P. M. Bleher, J. Ruiz and V. A. Zagrebnov, On the purity of limiting Gibbs state for the Ising model on the Bethe lattice. *J. Stat. Phys.* **79**:473–482 (1995).
6. G. R. Brightwell and P. Winkler, Gibbs extremality for the hard-core model on a Bethe lattice. (submitted) 1 (2003).
7. J. T. Chayes, L. Chayes, J. P. Sethna and D.. J. Thouless, A mean field spin glass with short-range interactions. *Commun. Math. Phys.* **106**:41–89 (1986).
8. T. M. Cover and J. A. Thomas, *Elements of Information Theory*, New York: Wiley Interscience (1991).
9. W. Evans, C. Kenyon, Y. Peres and L. J. Schulman, Broadcasting on trees and the Ising model. *Ann. Appl. Probab.* **10**:410–433 (2000).
10. G. Monroy F. Peruggi and F. di Liberto, The potts model on bethe lattices. i. general results. *J. Math. Phys.* **16**:811–817 (1984).
11. H.-O. Georgii. *Gibbs Measures and Phase Transition*, Berlin: de Gruyter (1988).
12. H. Kesten and B. P. Stigum, Additional limit theorems for decomposable multi-dimesional Galton–Watson processes. *Ann. Math. Statist.* **37**:1463–1481 (1966).
13. H. Kesten and B. P. Stigum, Limit theorems for decomposable multi-dimesional Galton–Watson processes. *J. Math. Anal. Appl.* **17**:309–338 (1966).

14. J. B. Martin, Reconstruction thresholds on regualr trees, *Discrete Math. Theor. Comput. Sci.* **AC**:191–204 (2003).
15. M. Mézard and A. Montanari, Unpublished (2005).
16. E. Mossel, Recursive reconstruction on periodic trees. *Rand. Struct. Alg.* **13**:81–97 (1998).
17. E. Mossel, Reconstruction on trees: beating the second eigenvalue. *Ann. Appl. Probab.* **11**:285–300 (2001).
18. E. Mossel, Survey: Information flows on trees, in J. Nestril and P. Winkler, eds., *Graphs, Morphisms and Statistical Physics. DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, American Mathematical Society (2004).
19. M. Mézard and G. Parisi, The Bethe lattice Spin glass revisited. *Eur. Phys. J. B* **20**:217–233 (2001).
20. M. Mézard and G. Parisi, The cavity method at zero temperature. *J. Stat. Phys.* **111**:1–34 (2003).
21. E. Mossel and Y. Peres, Information flow on trees. *Ann. Appl. Probab.* **13**:817–844 (2003).
22. R. Mulet, A. Pagnani, M. Weigt and R. Zecchina, Coloring random graphs. *Phys. Rev. Lett.* **89**:268701 (2002).
23. A. Montanari and G. Semerjian, On the dynamics of the glass transition on Bethe lattices, cond-mat/0509366 (2005).
24. E. Mossel and M. Steel, How much can evolved characters tell us about the tree that generated them? in O. Gascuel, ed., *Mathematics of Evolution and Phylogeny*, Oxford: Oxford University Press (2005).
25. F. Peruggi, Probability measures and Hamiltonian models on Bethe lattices: II. The solution of thermal and configurational problems. *J. Phys. A: Math. Gen.* **25**:3316–3323 (1983).
26. F. Peruggi, Probability measures and Hamiltonian models on Bethe lattices: I. Properties and construction of MRT probability measures. *J. Math. Phys.* **25**:3303–3315 (1984).
27. D. J. Thouless R. Abou-Chacra and P. W. Anderson, A selfconsistent theory of localization, *J. Phys. C* **6**:1734–1752 (1973).
28. A. N. Shiryaev, *Probability*, New York: Springer (1996).
29. S. Tatikonda and M. I. Jordan, *Loopy Belief Propagation and Gibbs Measures*. 18th Conference on Uncertainty in Artificial Intelligence, August 2002 (Edmonton, Canada).