

Typical behavior of relays in communication channels

R. C. Alamino and D. Saad

Neural Computing Research Group, Aston University, Birmingham, United Kingdom

(Received 2 August 2007; revised manuscript received 26 October 2007; published 21 December 2007)

The typical behavior of the relay-without-delay channel under low-density parity-check coding and its multiple-unit generalization, termed the *relay array*, is studied using methods of statistical mechanics. A demodulate-and-forward strategy is analytically solved using the replica symmetric ansatz which is exact in the system studied at Nishimori's temperature. In particular, the typical level of improvement in communication performance by relaying messages is shown in the case of a small and a large number of relay units.

DOI: [10.1103/PhysRevE.76.061124](https://doi.org/10.1103/PhysRevE.76.061124)

PACS number(s): 02.50.-r, 02.70.-c, 89.20.-a

I. INTRODUCTION

Methods of statistical mechanics have recently become increasingly more important in the study of communication channels. The development of the replica and cavity methods for analyzing disordered systems [1,2] and the related recent introduction of systematic rigorous bounds [3,4] made new theoretical tools available for their analysis.

More specifically, the replica method has been applied to a wide range of problems in information theory, from error correcting codes [5,6] to multiuser communication [7]. It facilitates the derivation of practical and theoretical limits in various communication channels and provides typical results in cases that are difficult to tackle via traditional methods of information theory.

The growing use of information networks, both physically connected and wireless, and the increasing number of services taking place in the Internet, have made the study of multiuser communication highly attractive and relevant from a practical point of view, in addition to being a challenging and exciting field for theoretical research.

Up to date, there is no generalized theory for multiuser channels within the framework of information theory and analytical results are only known for special cases. The main difficulty being that multiuser networks do not admit the source-channel separation principle, which allows one to separate the information transmission process into the two successive steps of source coding (compression) and channel coding (error correction); this principle plays an essential role in the information theoretic analysis of communication channels. In spite of their incomplete theoretical foundations, multiuser communication networks play an important role in a variety of communication devices ranging from mobile phones to computers. We strongly believe that a statistical physics-based analysis may offer answers where the current information theory methodology fails, especially in the limit of a large number of users.

With the technological demand and the possibility of providing a principled analysis by the methods of statistical mechanics, early results for multiuser communication are being revisited and analyzed from different and complementary points of view, resulting in new insights and developments [7,8]. One of the more interesting and relevant communication channels is the *relay channel* [9]. The generic relay channel is characterized by an auxiliary user between trans-

mitter and receiver, which assists in the transmission of the message. Due to the increase in the number of multiuser networks, such as mobile phones and computer networks, the transfer of information with the help of relays has become an attractive option. As these networks are becoming more distributed, the assisted transmission supported by arrays of relays has become feasible and merits further analytical exploration.

This paper is organized as follows. In Sec. II we define the general relay array and introduce as particular cases the classical relay channel and the recently investigated relay-without-delay channel. In Sec. III we outline the statistical physics methods used to analyze the problem which will be based on a replica approach detailed in Sec. IV. Section V contains our conclusions and final comments.

II. MODEL

A. Low-density parity-check codes

Low-density parity-check (LDPC) codes [10] are state-of-the-art error-correcting codes with performance that is second to none, especially within the high code rate regime. In the notation we will be using here, N -dimensional messages \mathbf{s} are encoded into M -dimensional codewords \mathbf{t} . LDPC codes are defined by a binary *parity-check matrix* $A = [C_1|C_2]$, concatenating side-by-side two very sparse matrices known to both sender and receiver: C_2 that is invertible and of dimensionality $(M-N) \times (M-N)$ and C_1 of dimensionality $(M-N) \times N$. The matrix A can be either random or regular, characterized by the number of nonzero elements per row (K) and column (C). Irregular codes show superior performance to regular structures [11,12] if constructed carefully. In order to simplify our treatment, we focus here on regular constructions; the generalization to irregular codes is straightforward [13,14].

Encoding refers to the linear mapping of an N -dimensional original message $\mathbf{s} \in \{0,1\}^N$ to an M -dimensional codeword $\mathbf{t} \in \{0,1\}^M$ ($M > N$)

$$\mathbf{t} = G\mathbf{s} \pmod{2}, \quad (1)$$

where all operations are performed in the field $\{0,1\}$, indicated by $\pmod{2}$, and the $M \times N$ generator matrix is

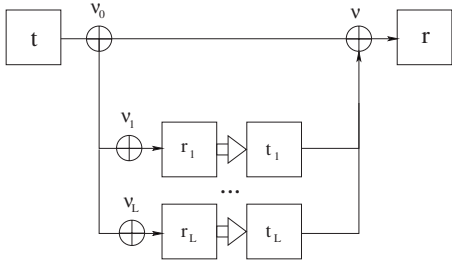


FIG. 1. The L -component relay array. The transmitter sends a codeword \mathbf{t} to the final receiver and to each of the L relays. Each relay receives a message \mathbf{r}_i which is a corrupted version of the original codeword subject to the AWGNs $\boldsymbol{\nu}_0$ and $\boldsymbol{\nu}_i$. It then sends to the final receiver the encoded vector \mathbf{t}_i . The final receiver receives the original transmitted codeword summed with all the relayed messages \mathbf{t}_i and corrupted by the AWGN vectors $\boldsymbol{\nu}_0$ and $\boldsymbol{\nu}$.

$$G = \begin{pmatrix} I \\ C_2^{-1} C_1 \end{pmatrix} (\text{mod } 2), \quad (2)$$

where I is the $N \times N$ identity matrix. By construction $AG = 0 (\text{mod } 2)$ and the first N bits of \mathbf{t} correspond to the original message \mathbf{s} .

Decoding is carried out by estimating the most probable transmitted vector from the received corrupted codeword [6,13]. For mathematical convenience, in the present work we map the Boolean variable $\mathbf{t} \in \{0, 1\}^M$ into a spin variable $\hat{\mathbf{t}} \in \{1, -1\}^M$ by the transformation $x \rightarrow (-1)^x$.

B. Relay array

The *relay array* is a multiple-unit generalization of the (single unit) relay channel of [9]. Since the single relay is a special case of this general framework, we will first explain the principles of relay-based communication using the more general scenario.

The LDPC codeword \mathbf{t} is transmitted to each one of L relay units through noisy channels, corrupted by a global additive white Gaussian noise (AWGN) $\boldsymbol{\nu}_0$ and by local independent AWGNs $\boldsymbol{\nu}_i$, both of zero mean and variances σ_0^2 and σ_i^2 , respectively. Each relay processes the received corrupted message \mathbf{r}_i and encodes the acquired information into a vector \mathbf{t}_i which is then transmitted to a final receiver. The final receiver receives an algebraic summation of the relay outputs plus a direct transmission from the original sender, corrupted also by $\boldsymbol{\nu}_0$, subject to a final AWGN $\boldsymbol{\nu}$ of zero mean and variance σ^2 . The exact form of the channel is depicted in Fig. 1 and the corresponding equations are

$$\mathbf{r} = a\mathbf{t} + \sum_{i=1}^L b_i \mathbf{t}_i + \boldsymbol{\nu} + \boldsymbol{\nu}_0, \quad (3)$$

$$\mathbf{r}_i = c_i \mathbf{t} + \boldsymbol{\nu}_i + \boldsymbol{\nu}_0. \quad (4)$$

The variables a , b_i , and c_i ($i=1, \dots, L$) are the relative gains of each transmission and can be random or set to constant values. The power from the original source to each relay is c_i^2 , to the final receiver is a^2 , and the power from each relay

to the final receiver is b_i^2 . For simplicity, as well as for comparison with results reported in the literature, we will mostly consider the case of unit relative gain parameters.

When $L=1$, we refer to the channel simply as the *relay channel*. In the *classical relay channel* (CRC), studied by Cover and El-Gamal [15], the messages sent by the relays to the final receiver are only allowed to depend on the set of symbols received by each of the relays *before* the current time step, $t_i^\mu = f(t^1, \dots, t^{\mu-1})$, which corresponds to the fact that it takes the relay some time to process the information before relaying it. However, if the time delay in the direct transmission to the final receiver is much longer than in the transmission to the relay units, one can allow the message sent by the relay to depend on the present received symbol as well such that $t_i^\mu = f(t^1, \dots, t^\mu)$. This last case, termed *relay-without-delay* (RWD), created significant interest recently and was studied by El-Gamal and Hassanpour [16]. For the case of a relay array where all communication is carried out through the relays and there is no direct transmission to the final receiver, the restriction of the CRC, to consider all but the last received symbol, is unnecessary.

The most studied strategies used by the relay units are the *amplify-and-forward* (AF) and the *decode-and-forward* (DF) strategies. In AF, which we will only mention briefly in this paper, the relay just retransmits its received vector after possible amplification, e.g., $\mathbf{t}_i = \mathbf{r}_i$. In DF, the relays decode the message and transmit their estimates to the final receiver. The mathematical treatment of both strategies will be discussed in the corresponding sections.

III. STATISTICAL PHYSICS OF DECODING

We transform the decoding problem of the final receiver into a statistical physics system by defining a dynamical variable $\boldsymbol{\tau} \equiv (\tau^1, \dots, \tau^M)$, which represents the candidate variable vectors at the receiver. Each τ^μ plays a role equivalent to a spin located in the μ th site of an M -site lattice.

The final receiver generates an estimate $\hat{\mathbf{t}}$ of the original codeword using the marginal posterior maximizer (MPM) estimator

$$\hat{t}^\mu = \text{sgn} \langle \tau^\mu \rangle_{\mathcal{P}(\boldsymbol{\tau}|\mathbf{r})}, \quad (5)$$

which minimizes the probability of bit error [13,17]. Other estimators can be used depending on the error measure considered. For example, minimization of *block error* is obtained using the maximum *a posteriori* estimator $\hat{\mathbf{t}} = \max_{\boldsymbol{\tau}} \mathcal{P}(\boldsymbol{\tau}|\mathbf{r})$.

The posterior probability density is calculated by Bayes' rule as

$$\mathcal{P}(\boldsymbol{\tau}|\mathbf{r}) = \frac{\mathcal{P}(\mathbf{r}|\boldsymbol{\tau})\mathcal{P}(\boldsymbol{\tau})}{\mathcal{P}(\mathbf{r})}, \quad (6)$$

with

$$\begin{aligned} \mathcal{P}(\mathbf{r}) &= \sum_{\mathbf{t}} \mathcal{P}(\mathbf{r}|\mathbf{t})\mathcal{P}(\mathbf{t}) \\ &= \sum_{\mathbf{t}, \{\mathbf{t}_i\}, \{\mathbf{r}_i\}} \mathcal{P}(\mathbf{t})\mathcal{P}(\mathbf{r}|\mathbf{t}, \{\mathbf{t}_i\}) \prod_i \mathcal{P}(\mathbf{t}_i|\mathbf{r}_i)\mathcal{P}(\mathbf{r}_i|\mathbf{t}). \end{aligned} \quad (7)$$

One of the basic quantities of interest is the overlap between the codeword and the decoded message. Our analysis focuses on the typical behavior of the decoding process and, accordingly, we take averages over all possible codewords, all received messages, and all allowed encodings, which we consider as quenched disorder in the corresponding physical system. The overlap between decoded and original messages takes the form

$$d = \frac{1}{M} \sum_{\mu=1}^M \langle t^\mu \operatorname{sgn}\langle \tau^\mu \rangle_{\mathcal{P}(\tau|\mathbf{r})} \rangle_{A,\mathbf{r},\mathbf{t}}. \quad (8)$$

This quantity can be derived from the free energy

$$f = - \lim_{M \rightarrow \infty} \frac{1}{\beta M} \langle \ln Z \rangle_{A,\mathbf{r},\mathbf{t}}, \quad (9)$$

with the partition function

$$Z = \sum_{\mathbf{t}} e^{-\beta \mathcal{H}(\mathbf{t};\mathbf{r})}, \quad (10)$$

and the corresponding Hamiltonian

$$\mathcal{H}(\mathbf{t};\mathbf{r}) = - \ln \mathcal{P}(\mathbf{r}|\mathbf{t})\mathcal{P}(\mathbf{t}). \quad (11)$$

Usually we disregard the normalization of the distributions within the Hamiltonian as they merely add constants that shift the zero energy. In the case of LDPC codes, $\mathcal{P}(\mathbf{t})$ turns out to be a constraint on the summation variables.

In the above Hamiltonian, the parity-check matrix A defines an interaction between the τ variables while \mathbf{t} and \mathbf{r} induce local fields at the corresponding sites. The inverse temperature β is the ratio between the true and the decoder's assumed noise level. In our numerical calculations, we adopt $\beta=1$, also known as *Nishimori's temperature*, which means that the decoder assumes the correct noise level for the channel. It can be shown that at Nishimori's temperature the system never enters the glassy phase [2,18] and the thermodynamically dominant solution is always replica symmetric (RS); we therefore restrict our analysis to the RS treatment.

One of the important properties in the statistical physics formulation of the problem is that looking at the problem as a dynamical spin system, one can interpret the results in terms of phase transitions, which are directly related to the overlap between the original and estimated messages and the entropy function of the obtained solutions. Combining this extra information we can have a better understanding of the way the system changes from a phase of perfect decoding (termed the *ferromagnetic phase*) to a phase where the message is recovered only up to a certain amount of error (the *paramagnetic phase*).

As the replica treatment of AF turns out to be the same as for the simple Gaussian channel with a modified power and noise level, the solution is obtained straightforwardly by applying the results of [13] and will not be studied here.

Full use of LDPC decoding in the relays in the DF strategy is made when each relay decodes the received vector \mathbf{r}_i by the MPM estimator using the fact that the codeword was encoded by a LDPC code. The message transmitted to the final receiver by each relay would then be

$$t_i^\mu = \operatorname{sgn}\langle \tau_i^\mu \rangle_{\mathcal{P}(\tau_i|\mathbf{r}_i)}. \quad (12)$$

In Eq. (6) this is equivalent to setting

$$\mathcal{P}(\mathbf{t}_i|\mathbf{r}_i) = \prod_{\mu=1}^M \delta(t_i^\mu - \operatorname{sgn}\langle \tau_i^\mu \rangle_{\mathcal{P}(\tau_i|\mathbf{r}_i)}). \quad (13)$$

As $t_i^\mu \in \{\pm 1\}$, we can rewrite this probability density as

$$\mathcal{P}(\mathbf{t}_i|\mathbf{r}_i) = \prod_{\mu=1}^M \theta(\langle t_i^\mu \tau_i^\mu \rangle_{\mathcal{P}(\tau_i|\mathbf{r}_i)}), \quad (14)$$

where θ is the Heaviside step function.

The replica treatment of the LDPC DF turns out to be extremely involved due to the introduction of a θ function with an average over the variables τ_i inside it, which includes a term dependent on the parity-check matrix. Analytical studies of this rather difficult case are under way.

In the present work we focus on a simplification of this strategy, also known in the literature as demodulate and forward. In it, the relays do not have the complete information about the encoding mechanism and therefore assume a uniform prior for the transmitted codeword. In this case, the posterior distribution of the bits in the message for the relay is

$$\mathcal{P}(\mathbf{t}_i|\mathbf{r}_i) = \prod_{\mu=1}^M \frac{1}{1 + \exp[-2t_i^\mu r_i^\mu / (\sigma_i^2 + \sigma_0^2)]}, \quad (15)$$

and it is straightforward to show that the MPM estimator is given simply by

$$t_i^\mu = \operatorname{sgn}(r_i^\mu). \quad (16)$$

The fact that the disorder with respect to the selected code does not appear in the estimate of the relays makes the replica calculations feasible in this case, as follows.

IV. REPLICA SYMMETRIC ANALYSIS

As the RS analysis of LDPC coding systems has been introduced and carried out in a number of presentations (e.g., [13]) we will omit the detailed derivation and concentrate on the final expressions. The derivation follows exactly the same steps as in [19] where quenched averages over all possible parity-check matrices are first carried out, followed by the RS assumption which enables the representation of the order parameters in the form of field distributions (see also [20]) with each order parameter containing m replicated dynamical variables τ_i being written as

$$q(\tau_{ia_1}, \dots, \tau_{ia_m}) = \int_{-\infty}^{\infty} dx \pi(x) x^{m-\Delta} \equiv \langle x^{m-\Delta} \rangle_x, \quad (17)$$

$$\hat{q}(\tau_{ia_1}, \dots, \tau_{ia_m}) = \int_{-\infty}^{\infty} d\hat{x} \hat{\pi}(\hat{x}) \hat{x}^{m-\Delta} \equiv \langle \hat{x}^{m-\Delta} \rangle_{\hat{x}}, \quad (18)$$

where Δ is 1 if the zeroth replica is included and 0 otherwise. The field distributions π and $\hat{\pi}$ act as generators of the order parameters in the replica symmetric analysis of diluted systems.

A set of self-consistent equations is obtained by the saddle point method in the thermodynamic limit where the extremizations are made with respect to the field distributions π and $\hat{\pi}$ resulting in

$$\hat{\pi}(\hat{x}) = \left\langle \delta \left(\hat{x} - \prod_{m=1}^{K-1} x^m \right) \right\rangle_x,$$

$$\pi(x) = \left\langle \delta \left(x - \frac{\sum_{\tau} [\Psi(\tau, r)]^{\beta} \prod_{l=1}^{C-1} (1 + \tau \hat{x}^l)}{\sum_{\tau} [\Psi(\tau, r)]^{\beta} \prod_{l=1}^{C-1} (1 + \tau \hat{x}^l)} \right) \right\rangle_{r, \hat{x}}, \quad (19)$$

where

$$\Psi(\tau, r) \equiv \int \left[\prod_{i=1}^L dr_i \exp \left(- \frac{(r_i - c_i \tau)^2}{2(\sigma_i^2 + \sigma_0^2)} \right) \right] \times \exp \left[- \frac{1}{2(\sigma^2 + \sigma_0^2)} \left(r - a\tau - \sum_i b_i \operatorname{sgn} r_i \right)^2 \right]. \quad (20)$$

The expression $\langle t_i \rangle$ is the mean of the variable t_i and $\mathcal{P}(r) \propto \Psi(1, r)$. The overlap is

$$d = \langle \operatorname{sgn} u \rangle_u, \quad (21)$$

with

$$\mathcal{P}(u) = \left\langle \delta \left(u - \frac{\sum_{\tau} [\Psi(\tau, r)]^{\beta} \prod_{l=1}^C (1 + \tau \hat{x}^l)}{\sum_{\tau} [\Psi(\tau, r)]^{\beta} \prod_{l=1}^C (1 + \tau \hat{x}^l)} \right) \right\rangle_{r, \hat{x}}, \quad (22)$$

the free energy is given by

$$\beta f = \frac{C}{K} \ln 2 + C \langle \ln(1 + x\hat{x}) \rangle_{x, \hat{x}} - \frac{C}{K} \left\langle \ln \left(1 + \prod_{m=1}^K x^m \right) \right\rangle_x - \left\langle \ln \left(\sum_{\tau} [\Psi(\tau, r)]^{\beta} \prod_{l=1}^C (1 + \tau \hat{x}^l) \right) \right\rangle_{\hat{x}, r}, \quad (23)$$

and the internal energy, the derivative with respect to β of the above equation becomes

$$u = - \left\langle \frac{\sum_{\tau} \Psi^{\beta}(\ln \Psi) \prod_{l=1}^C (1 + \tau \hat{x}^l)}{\sum_{\tau} \Psi^{\beta} \prod_{l=1}^C (1 + \tau \hat{x}^l)} \right\rangle_{\hat{x}, r}. \quad (24)$$

For any number L of relays, the results can be obtained by a numerical solution of Eqs. (19). Note the summation over the internal variables, i.e., the messages received and sent by the relays. This comes from the Bayesian formulation of the problem where the final receiver has access just to \mathbf{r} and, therefore, one must integrate over all unknown variables.

We also note that the above equations are fairly general. Using the appropriate function Ψ one can recover all previous results for single user channels and apply them to more general channels when the intermediate processing of the message does not involve any knowledge of the parity-check matrices.

The ferromagnetic state, which corresponds to perfect decoding, is given by the following solution to the saddle point Eqs. (19):

$$\hat{\pi}(\hat{x}) = \delta(\hat{x} - 1), \quad \pi(x) = \delta(x - 1). \quad (25)$$

Substitution of these distributions in Eq. (21) gives $d=1$. By substituting the ferromagnetic solution into the formulas for the free and internal energies, we obtain (at Nishimori's temperature)

$$u = f = - \langle \ln \Psi(1, r) \rangle_r, \quad (26)$$

meaning that the entropy of this phase is zero.

The Hamiltonian of the relay array is *gauge invariant* with respect to the gauge transformation

$$r^{\mu} \rightarrow \gamma^{\mu} r^{\mu},$$

$$t^{\mu} \rightarrow \gamma^{\mu} t^{\mu}, \quad (27)$$

where the vector γ obeys the parity-check constraints. We can verify that the transition probabilities $\mathcal{P}(r^{\mu} | t^{\mu})$ are also invariant under this gauge transformation. Note that if a channel is symmetric, i.e., exhibits a similar probability for cross-symbol error flips (for a detailed definition see [21]), it is automatically gauge invariant under the above transformation. For gauge invariant channels the internal energy is

$$U = \langle \mathcal{H}(\boldsymbol{\tau}; \mathbf{r}) \rangle_{\boldsymbol{\tau}, \mathbf{r}, \mathbf{t}} = \sum_{\boldsymbol{\tau}, \mathbf{t}} \int d\mathbf{r} \mathcal{P}(\boldsymbol{\tau} | \mathbf{r}, \beta) \mathcal{P}(\mathbf{r} | \mathbf{t}) \mathcal{P}(\mathbf{t}) \mathcal{H}(\boldsymbol{\tau}; \mathbf{r}), \quad (28)$$

where

$$\mathcal{P}(\boldsymbol{\tau} | \mathbf{r}, \beta) \propto e^{-\beta \mathcal{H}}, \quad (29)$$

is the thermal Gibbs probability at inverse temperature β which obeys $\mathcal{P}(\boldsymbol{\tau} | \mathbf{r}, \beta=1) = \mathcal{P}(\boldsymbol{\tau} | \mathbf{r})$. Since under such a gauge transformation the Hamiltonian remains invariant, we have $\mathcal{H}(\mathbf{t}; \mathbf{r}) = \mathcal{H}(\mathbf{1}; \mathbf{r})$, where $\mathbf{t} \equiv (t^1 r^1, \dots, t^M r^M)$ and $\mathbf{1}$ is an M -dimensional vector with all entries equal to 1. Therefore, one can write the following expression for the internal energy:

$$U = \sum_{\boldsymbol{\tau}, \mathbf{t}} \int d\mathbf{r} \frac{\mathcal{P}(\mathbf{r} | \boldsymbol{\tau}, \beta) \mathcal{P}(\boldsymbol{\tau} | \beta)}{\mathcal{P}(\mathbf{r} | \beta)} \mathcal{P}(\mathbf{r} | \mathbf{t}) \mathcal{P}(\mathbf{t}) \mathcal{H}(\mathbf{1}; \boldsymbol{\tau}). \quad (30)$$

Gauging the variables $\boldsymbol{\tau} \rightarrow \mathbf{r}$, reorganizing the terms, and taking $\beta=1$, we finally obtain

$$U = \int d\mathbf{r} \mathcal{P}(\mathbf{r} | 1) \mathcal{H}(\mathbf{1}; \mathbf{r}). \quad (31)$$

The meaning of this is that, for a gauge invariant channel of the type described above (which includes general symmetric channels), the internal energy is independent of the con-

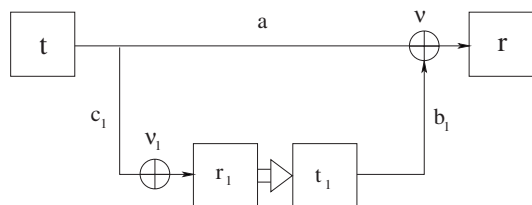


FIG. 2. Schematic drawing of the relay-without-delay setup to be analyzed.

figuration. In special cases, as can be found in [2], the gauge symmetry allows for an analytical expression to be found. The same method can be used to prove that the probability distribution for the magnetization is equal to the probability distribution for the two-point correlations in Nishimori's temperature, which indicates the absence of a spin-glass phase and no replica symmetry breaking.

A. Relay channel

In order to compare our results with those of [16], we analyze the RWD for the setup sketched in Fig. 2 with $\sigma_1^2 = \eta\sigma^2$, $a=b_1=1$, and $c_1=(1+\sigma^2)^{-1/2}$. The corresponding function Ψ is then given by

$$\Psi(\tau, r) = e^{-(r-\tau-1)^2/2\sigma^2} \operatorname{erfc}\left(-\frac{\tau}{\sqrt{2\eta\sigma^2}}\right) + e^{-(r-\tau+1)^2/2\sigma^2} \operatorname{erfc}\left(+\frac{\tau}{\sqrt{2\eta\sigma^2}}\right), \quad (32)$$

where η is an arbitrary positive constant and $\operatorname{erfc}(x)$ is the complementary error function

$$\operatorname{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty e^{-y^2} dy. \quad (33)$$

For these values of noise and gains, the capacity of this channel as derived in [16] is

$$C = \frac{1}{2} \log_2 \left(1 + \frac{1+c_1^2}{\sigma^2} \right). \quad (34)$$

The numerical results for the overlap between the retrieved and the original codewords, obtained by solving recursively Eqs. (19), are given in Fig. 3 for $K=4$, $C=3$, $\beta=1$, and $\eta=0.1$. Shannon's limit, marking the noise level below which error-free communication is theoretically possible, is indicated by the vertical dashed line and corresponds to a noise level $\sigma^2 \approx 8.79$. The dashed curve shows the overlap for a simple Gaussian channel with noise level σ^2 and the continuous one shows the overlap for the RWD. The improvement in the practical limit for error-free communication indicated by the highest noise level for which $d=1$ is clear. However, the distance between the dynamical transition threshold $\sigma_d^2 \approx 2.22$, marking the point where subdominant metastable states emerge, and Shannon's limit for the channel is greater than in the case of the simple Gaussian channel (for numerical results for the Gaussian channel see [21]). Numerical

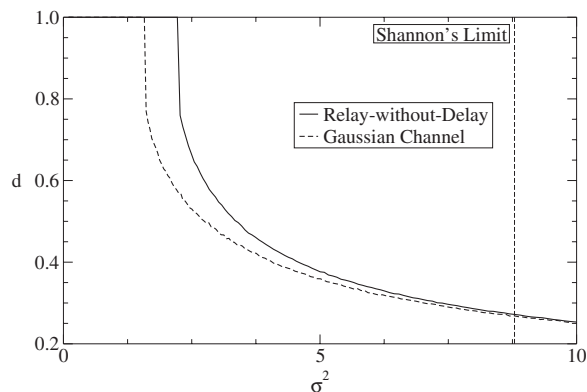


FIG. 3. The overlap of retrieved and original codewords for the RWD channel with $\sigma_1^2=0.1\sigma^2$ is given by the continuous curve. The dashed curve shows the same for the simple Gaussian channel. The vertical line indicates Shannon's limit for the RWD as calculated by El-Gamal and Hassanpour [16].

calculations point to the expected result that decreasing the noise level from the source to the relay brings the dynamical transition threshold σ_d^2 closer to Shannon's limit. However, one must remember that the relay strategy examined does not use the full potential of the relay and the additional information embedded in the LDPC codes. We expect that a LDPC decoding *in the relay* will improve the communication performance and currently focus on the analysis of this scenario.

We can also see in Fig. 3 that, as the noise level increases, the channel becomes closer to the Gaussian channel. This is just a consequence of the fact that, for high noise level, the additional information provided by the relay becomes negligible as both relay and receiver decode the message poorly.

Figure 4 shows the entropy and the free and internal energies for the same values as in Fig. 3. At the dynamical transition point, where practical perfect decoding becomes unfeasible, the entropy becomes negative, indicating the emergence of subdominant metastable states that can be further explored using the replica symmetry breaking ansatz. Between this point and the thermodynamical transition point (where the ferromagnetic free energy becomes subdominant and the entropy becomes positive again) the dominant state is still ferromagnetic but the population dynamics algorithm used to solve the saddle point equations becomes trapped in a local minimum with free energy higher than the ferromagnetic one. This is also reflected in the failure of local search algorithms in finding the original message as they are likely to be trapped in the exponentially many suboptimal solutions. Due to the equality between the internal energy and the ferromagnetic free energy, the point where the entropy becomes positive again is also the point where both energies cross as shown in the bottom graph.

In Fig. 5 we plot the dynamical and thermodynamical transition noise levels against η , the ratio between the relay and final receiver noise levels. We see that the dynamical and thermodynamical transition points decrease with η but become closer to each other, stabilizing at asymptotic values that match those of the simple AWGN channel values as the relay contribution becomes meaningless.

Although the capacity for the RWD is known only in special cases, its upper bound can be higher than in the case

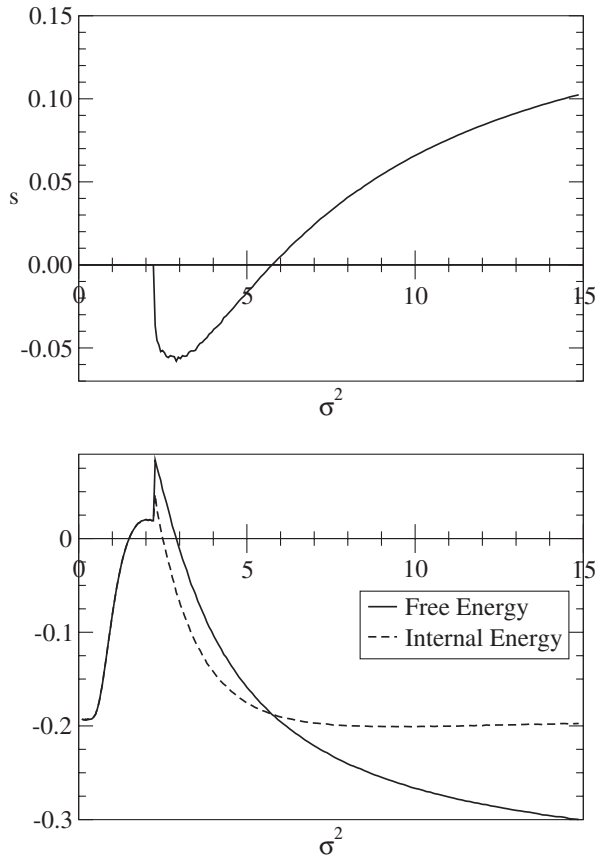


FIG. 4. Entropy and energies for the RWD channel at Nishimori’s temperature. The upper graph shows the entropy which is given by the difference between the internal energy and the free energy depicted in the bottom graph.

of the CRC. In order to verify it for the LDPC-based framework that we analyze in this paper, we now use a setup equivalent to the one studied in [15] and shown in Fig. 6 where $a=c_1=b_1=1$ and $\sigma_0^2 \equiv \lambda\sigma^2$, with λ an arbitrary positive constant. The capacity of the CRC in this case is

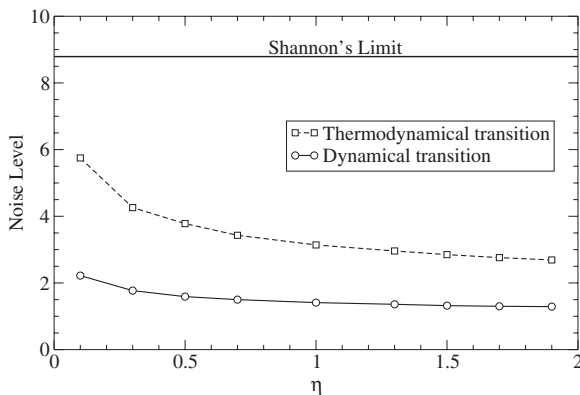


FIG. 5. The continuous-dashed line shows the dynamical-thermodynamical transition noise levels of the RWD against η , the ratio between relay and final receiver noise levels. The upper horizontal line corresponds to Shannon’s limit $\sigma^2 \approx 8.79$.

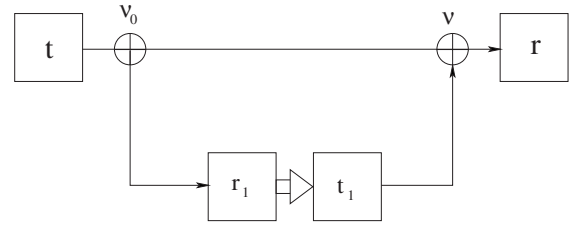


FIG. 6. Schematic drawing of the classical relay channel setup. The relative gains are all equal to 1 and not shown in the figure.

$$C = \begin{cases} \frac{1}{2} \log_2 \left(1 + \frac{1}{\lambda \sigma^2} \right), & \lambda \geq 1, \\ \frac{1}{2} \log_2 \left(1 + \frac{4}{(1 + \lambda)^2 \sigma^2} \right), & \lambda < 1. \end{cases} \quad (35)$$

In Fig. 7 we compare the dynamical and thermodynamical threshold noise levels of a RWD with Shannon’s limit for the CRC, both with the setup described above, for different values of λ , the ratio between the noise levels applied at the transmission and reception points.

We can see that, although the practical decoding line (dynamical transition) falls below Shannon’s limit for all calculated values, the thermodynamical transition goes above it for the CRC case at higher values of λ . Figure 7 shows that the capacity of the RWD is indeed higher than the CRC for the case studied and quantifies the gain in allowing the message sent by the relay to depend on the current transmitted symbol (which is excluded in the CRC); the RWD result being calculated with the practical LDPC coding scheme. Although allowing this instantaneous dependence would at first sight seem just a small modification, insignificant in the infinite block length limit, it indeed gives relevant extra information which facilitates more efficient retrieval at the final receiver. The insight gained is that for the RWD and large λ , the relay transmission t_1^μ is correlated with the original codeword t^μ , which is not the case in the CRC; this allows

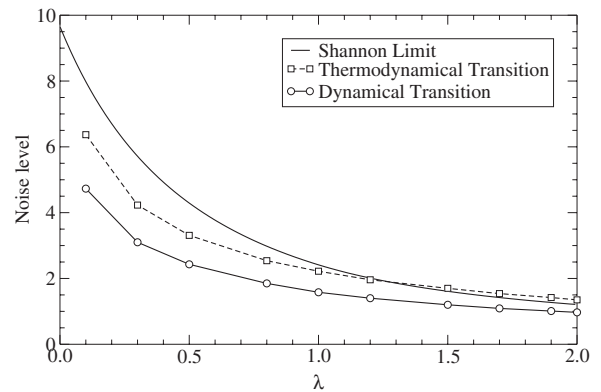


FIG. 7. The continuous-dashed line shows the dynamical-thermodynamical transition noise levels of the RWD against λ in the setup of Fig. 6. The continuous line without marked symbols is Shannon’s limit for a CRC with the same noise levels and transmission powers.

for an improvement in the information extraction at the receiver.

B. Large relay array

Now, we will use the central limit theorem to obtain the result for large L in the relay array setup given in Fig. 1. As the relay messages are correlated and to guarantee that the quantities have the same order, we introduce a $1/L$ scaling in the summation over relay messages. The function Ψ for this model becomes

$$\begin{aligned} \Psi(\tau, r) \equiv & \int \left[\prod_{i=1}^L dr_i \exp\left(-\frac{(r_i - \tau)^2}{2(\sigma_i^2 + \sigma_0^2)}\right) \right] \\ & \times \exp\left[-\frac{1}{2(\sigma^2 + \sigma_0^2)}\left(r - \tau - \frac{1}{L} \sum_i \text{sgn } r_i\right)^2\right], \end{aligned} \quad (36)$$

where we assumed, for simplicity, $a=b_i=c_i=1$.

For $L \gg 1$, the central limit theorem gives rise to a modified distribution of the variable r given by

$$\mathcal{P}(r) = \int \left(\prod_{i=1}^L dr_i \mathcal{P}(r_i) \right) F\left(\frac{1}{L} \sum_{i=1}^L \text{sgn } r_i\right) = \langle F(\omega) \rangle_\omega, \quad (37)$$

where

$$F(\omega) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(r-1-\omega)^2}{2(\sigma^2 + \sigma_0^2)}\right), \quad (38)$$

$$\mathcal{P}(r_i) = \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left(-\frac{(r_i-1)^2}{2(\sigma_i^2 + \sigma_0^2)}\right), \quad (39)$$

with

$$\mathcal{P}(\omega) = \mathcal{N}\left(\frac{1}{L} \sum_{i=1}^L \langle \text{sgn } r_i \rangle_{r_i}, \frac{1}{L^2} \sum_{i=1}^L (1 - \langle \text{sgn } r_i \rangle_{r_i})^2\right). \quad (40)$$

For simplicity, we consider the case where the noise level is the same for all relays $\sigma_i^2 = \sigma_1^2$ and define

$$\sigma_f^2 \equiv \sigma_1^2 + \sigma_0^2, \quad \sigma^2 \equiv \sigma^2 + \sigma_0^2. \quad (41)$$

The corresponding distribution for ω then becomes

$$\mathcal{P}(\omega) = \mathcal{N}\left(\text{erf}(1/\sqrt{2}\sigma_r), \frac{1}{L} \text{erfc}^2(1/\sqrt{2}\sigma_r)\right). \quad (42)$$

Consequently, the contribution for the final noise level coming from the relay transmission decreases as L^{-1} . In the limit $L \rightarrow \infty$, this distribution becomes a δ function centered at the error function value and therefore

$$\mathcal{P}(r) = \mathcal{N}[1 + \text{erf}(1/\sqrt{2}\sigma_r), \sigma_f^2]. \quad (43)$$

Accordingly, the function Ψ becomes

$$\Psi(\tau, r) = \exp\left(-\frac{1}{2\sigma_f^2} [r - \tau - \text{erf}(\tau/\sqrt{2}\sigma_r)]^2\right). \quad (44)$$

Figure 8 compares the dynamical and thermodynamical tran-

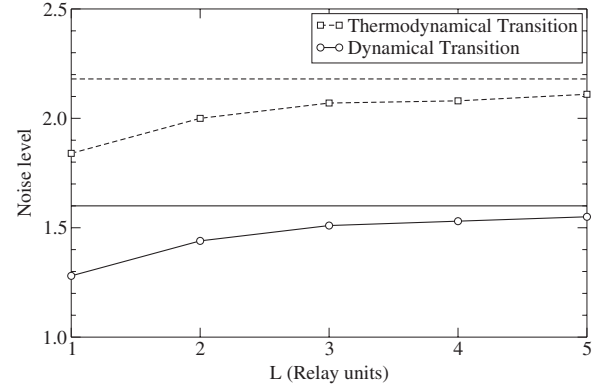


FIG. 8. Dynamical and thermodynamical transition points for many relays. The exact formula is used to calculate the points $L = 1, 2, 3, 4, 5$. The horizontal lines represent the large L limit.

sition points for $L=1, 2, 3, 4, 5$ calculated by the exact formula and the result obtained by the approximation for large L . Again, we consider the case of $K=4$, $C=3$, and $\beta=1$.

It is clear from Fig. 8 that already at $L=5$, both dynamical and thermodynamical transition points approach the large L limit solution, thus making this approximation attractive already for low L values.

V. CONCLUSIONS

In this work we analyzed the behavior of relay arrays using methods of statistical mechanics. These communication networks are of growing significance due to the increase of multiuser, mobile, and distributed communication systems.

We found an analytical solution for the relay-without-delay channel given by the RS ansatz which, due to the gauge symmetry of the channel, is exact at Nishimori's temperature that corresponds to a choice of the correct prior within the Bayesian framework. We showed the level of improvement with respect to a simple Gaussian channel without relaying which, even for the naive relay strategy of demodulate and forward analyzed here, is significant.

We compared the RWD dynamical and thermodynamical transition points, for different noise ratios, between the relay and the direct channels, and found that although these points are far from Shannon's theoretical limit, the difference between the dynamical and the thermodynamical transition decreases. The relevance of the relay is clearly decreasing as its noise level increases as the level of additional information it conveys diminishes.

We also were able to compare the RWD case to the classical relay channel for different noise ratios between the relay and the direct channel. We found that the capacity of the RWD is higher than the CRC for a high relay noise, showing the significance of the extra information conveyed by the relay on the current transmitted symbol, which is absent in the CRC framework.

The performance of a large array of relays was analyzed and compared against results obtained for a small number of units. The results obtained are consistent and indicate that

this useful approximation provides accurate results already for a small number of units. For a large array, we also found that the increase in noise tolerance levels off.

We have demonstrated the usefulness of methods adopted from statistical physics for analyzing multiuser communication systems. While we have concentrated on limited scenarios of relay channels, we believe that these methods hold a promising alternative to the information theory methodology which, in general, has not been successful in dealing

with multiuser communication systems. The study of different relay channels and other multiuser communication networks is underway.

ACKNOWLEDGMENT

Support from EVERGROW Contract No. 1935 in FP6 of the EU EPSRC Grant No. EP/E04951611 are gratefully acknowledged.

-
- [1] M. Mézard, G. Parisi, and M. Virasoro, *Spin Glass Theory and Beyond* (World Scientific, Singapore, 1987).
 - [2] H. Nishimori, *Statistical Physics of Spin Glasses and Information Processing* (Oxford University Press, Oxford, UK, 2001).
 - [3] F. Guerra, *Commun. Math. Phys.* **233**, 1 (2003).
 - [4] S. Franz, M. Leone, F. Ricci-Tersenghi, and R. Zecchina, *Phys. Rev. Lett.* **87**, 127209 (2001).
 - [5] N. Sourlas, *Nature (London)* **339**, 693 (1989).
 - [6] Y. Kabashima and D. Saad, *J. Phys. A* **37**, R1 (2004).
 - [7] T. Tanaka, *IEEE Trans. Inf. Theory* **48**, 2888 (2002).
 - [8] K. Nakamura, Y. Kabashima, R. Morelos-Zaragoza, and D. Saad, *Phys. Rev. E* **67**, 036703 (2003).
 - [9] T. M. Cover and J. Thomas, *Elements of Information Theory* (Wiley, New York, 1991).
 - [10] R. Gallager, *IRE Trans. Inf. Theory* **8**, 21 (1962).
 - [11] T. Richardson, A. Shokrollahi, and R. Urbanke, *IEEE Trans. Inf. Theory* **47**, 619 (2001).
 - [12] I. Kanter and D. Saad, *Phys. Rev. Lett.* **83**, 2660 (1999).
 - [13] R. Vicente, D. Saad, and Y. Kabashima, in *Advances in Imaging and Electron Physics*, edited by P. Hawkes (Academic, New York, 2002), Vol. 125, pp. 232–353.
 - [14] R. Vicente, D. Saad, and Y. Kabashima, *J. Phys. A* **33**, 6527 (2000).
 - [15] T. M. Cover and A. A. El-Gamal, *IEEE Trans. Inf. Theory* **25**, 572 (1979).
 - [16] A. El-Gamal and N. Hassanpour, *Proceedings of the 2005 IEEE International Symposium on Information Theory*, p. 1078.
 - [17] Y. Iba, *J. Phys. A* **32**, 3875 (1999).
 - [18] A. Montanari, *Eur. Phys. J. B* **23**, 121 (2001).
 - [19] R. C. Alamino and D. Saad, *J. Phys. A* **40**, 12259 (2007).
 - [20] K. Y. M. Wong and D. Sherrington, *J. Phys. A* **20**, L793 (1987).
 - [21] T. Tanaka and D. Saad, *J. Phys. A* **36**, 11143 (2003).