Analysis of Belief Propagation for Non-Linear Problems: The Example of CDMA (or: How to Prove Tanaka's Formula)

Andrea Montanari
Laboratoire de Physique Théorique
de l'Ecole Normale Supérieure,
CNRS-UMR 8549
Email: montanar@lpt.ens.fr

David Tse
Department of Electrical Engineering
and Computer Science,
University of California, Berkeley
Email: dtse@eecs.berkeley.edu

***Abstract*— We consider the CDMA (code-division multiple-access) multi-user detection problem for binary signals and additive white gaussian noise. We propose a spreading sequences scheme based on random sparse signatures, and a detection algorithm based on belief propagation (BP) with linear time complexity. In the new scheme, each user conveys its power onto a *finite* number of chips $\bar{l}$, in the large system limit.**

We analyze the performances of BP detection and prove that they coincide with the ones of optimal (symbol MAP) detection in the $\bar{l} \to \infty$ limit. In the same limit, we prove that the information capacity of the system converges to Tanaka's formula for random 'dense' signatures, thus providing the *first* rigorous justification of this formula. Apart from being computationally convenient, the new scheme allows for optimization in close analogy with irregular low density parity check code ensembles.

I. Introduction

A. Motivation

The crucial new characteristics of modern (iterative) coding systems [1] are: (i) Probabilistic construction based on sparse random graphs; (ii) Iterative (belief propagation, BP) decoding; (iii) Focus onto the large system limit. Despite their generality, the impact of these principles outside the area of linear error correcting codes has been limited. It is therefore extremely interesting to extend their scope to other communications and information theory problems[1].

The tools developed for the analysis of iterative coding systems must be considerably strenghtened in order to cope with such generalizations. Consider for instance the question of whether BP decoding is asymptotically optimal (in the large system limit), i.e. if it implements symbol MAP decoding. For LDPC codes, density evolution (DE) allows to show that this is the case if the noise level is smaller than a threshold, below which the asymptotic BP bit error rate $\mathrm{P}_{\mathrm{b}}^{\mathrm{BP}}$ vanishes. When $\mathrm{P}_{\mathrm{b}}^{\mathrm{BP}} > 0$ (as we expect in a general setting), one cannot say much about MAP performances, and their relation to BP (apart from the obvious sub-optimality of BP).

Recently, some definite progress was made on these problems in the context of LDPC codes [3], [4], [5]. The basic new ingredient is a 'general area theorem' that yields the rate of change of the mutual information across the system, under a change in the channel parameter. Earlier examples of such a relation were found by Ashikhmin, Kramer , and ten Brink [6] (for the erasure channel), and Guo, Shamai and Verdù [7], [8] (for the gaussian and Poisson channels). The approach based on the area theorem seems rather general. In order to illustrate it, and further explore its capabilities, we consider here a new application: multi-user detection [9].

B. Multi-user detection with binary inputs

In a simple multi-user detection scenario, each of K users transmits a symbol $x_i \in \mathbb{R}$ to a common receiver, after encoding it using a signature $\underline{s}_i \in \mathbb{R}^N$. The received signal is

$$\underline{y} = \sum_i x_i \, \underline{s}_i + \underline{w} \, . \tag{1}$$

where the noise $\underline{w}$ is a vector of N i.i.d. gaussian variables of mean 0 and variance σ^2. The input symbols x_i are also modeled as i.i.d.'s. Writing $\mathbb{S}$ for the $N \times K$ matrix with columns $\underline{s}_1, \dots, \underline{s}_K$, and $x = (x_1, \dots, x_K)^{\mathrm{T}}$ for the input, the above equation can also be written $\underline{y} = \mathbb{S}\,\underline{x} + \underline{w}$. Of great interest is the large system limit $N, K \to \infty$ with $K/N = \alpha$ fixed.

How reliably can the input $\underline{x}$ be reconstructed given $\underline{y}$ and the signature matrix $\mathbb{S}$? In order to answer this question, the signatures $\underline{s}_i$ are usually taken to be i.i.d. random vectors. The standard choice is to set $\underline{s}_i = \frac{1}{\sqrt{N}}(s_{i1}, \dots, s_{iN})^{\mathrm{T}}$ where the s_{ia} are i.i.d. with zero mean and unit variance (we will call these 'dense signatures'). Tse and Hanly [10], and Verdù and Shamai [11] considered the case in which the input symbols x_i are gaussian random variables. Using random matrix theory, they were able to compute the minimum mean square error, and the information capacity of the system. In [12], we considered a multi-user detection algorithm based on

[1] An earlier example that support this view is the use of low density codes with non-linear checks for lossy data compression in [2].

BP, and proved it to be optimal (i.e. to implement minimum mean square error detection) with high probability in the large system limit.

The case of binary input symbols $x_i \in \{+1,-1\}$ uniformly at random, is of obvious interest for practical applications, and out of reach of classical methods (such as random matrix theory). Tanaka [13] used the replica method from statistical physics in order to determine the asymptotic information capacity. More precisely, let us define per-user conditional entropy $h \equiv \lim_{K\to\infty} K^{-1}\mathbb{E}H(X|Y)$, where the expectation is taken with respect to the random signatures and throughout the paper we measure entropies in nats (obviously $I(X;Y) = K\log 2 - H(X|Y)$). He obtained $h = h_{\rm RS}(\sigma^2,\alpha) \equiv \sup_q h_{\rm RS}(q;\sigma^2,\alpha)$, where

$$\begin{aligned} h_{\rm RS}(q;\sigma^2,\alpha) &= \mathsf{E}_z \log 2\cosh(\lambda(q)+\sqrt{\lambda(q)}\,z) - \\ &\quad -\frac{1}{2}\lambda(1+q) - \frac{1}{2\alpha}\log\left(1+\frac{\alpha}{\sigma^2}(1-q)\right), \end{aligned} \tag{2}$$

$\lambda(q) = [\sigma^2+\alpha(1-q)]^{-1}$, and E_z denotes throughut the paper expectation with respect to the standard normal variable z. It is easy to show that the value of q maximizing $h_{\rm RS}(q;\sigma^2,\alpha)$ must satisfy the stationarity condition

$$q = \mathsf{E}_z \tanh^2(\lambda(q)+\sqrt{\lambda(q)}\,z)\,. \tag{3}$$

Unhappily, the replica method is non-rigorous. In this paper we will prove Tanaka's formula for $\alpha \le \alpha_{\rm s} \approx 1.49$ (a precise definition of $\alpha_{\rm s}$ is provided in the next Section). For earlier applications of BP to multi-user detection with binary signals, we refer, for instance to [14], [15], [16]. We will prove that, in the same regime $\alpha < \alpha_{\rm s}$, optimal (symbol MAP) detection can be implemented using BP.

In order to prove these results, we will introduce a new 'sparse signature' scheme, see Section II, and view standard dense signatures as a limiting case. The identity between the two limiting procedures will be the object of a separate publication. The new scheme (which is reminiscent of LT codes [17]) is on the other hand interesting *per se*. It allows to implement BP in a very natural way with complexity linear in N. Furthermore, it opens the way to optimization of the degree sequence thus improving the performances over dense signatures. We refer to Section IV for numerical indications in this direction.

II. The sparse signature scheme, and main results

A. Sparse signatures and belief propagation

As already mentioned, in order to prove Tanaka's formula we shall introduce a new signature scheme. This is caracterized by a distribution $\{\Omega_l : l \ge 0\}$ over the non negative integers (to avoid pathological behaviors, we assume it to have bounded support). We also let $\overline{l} > 0$ be its mean and define $\omega_l \equiv l\Omega_l/\overline{l}$ for $l \ge 0$. The user i constructs her signature s_i independently from the other users as follows. She chooses an integer l from the distribution Ω_l, and a subset ∂i of $\{1,\dots,N\}$ of size $|\partial i| = l$ uniformly at random among the $\binom{N}{l}$ such subsets. Her signature is $s_i = \frac{1}{\sqrt{l}}(s_{i1},\dots,s_{iN})^{\rm T}$ where $s_{ia} \in \{+1,-1\}$ uniformly at random if $a \in \partial i$, and $s_{ia} = 0$ otherwise.

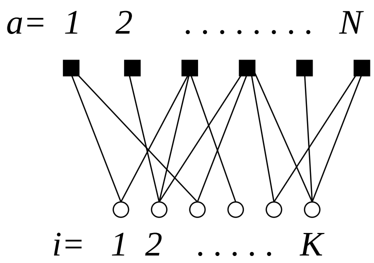

Fig. 1. Factor graph representation of the sparse signature scheme: circles represent users (variable nodes) and squares chips (function nodes).

Notice that the normalization ensures that the average power employed by each user is equal to 1 as for the dense signature scheme. However this power is conveyed onto a finite number of chips. Viceversa, each chip $a \in \{1,\dots,N\}$ receives power from a finite number of users, to be denoted as ∂a (this is the set of $i \in \{1,\dots,K\}$ such that $a \in \partial i$). The conditional distribution of the input symbols, given the received signal y take the form

$$\mu^{y,\mathbb{S}}(x) = \frac{1}{Z}\prod_{a=1}^{N}\psi_{y_a}(x_{\partial a})\,, \tag{4}$$

$$\psi_{y_a}(x_{\partial a}) = \exp\left\{-\frac{1}{2\sigma^2}\left(y_a - \sum_{i\in\partial a}\frac{s_{ia}}{\sqrt{l}}x_i\right)^2\right\}\,. \tag{5}$$

Such distribution is conveniently represented through the associated factor graph, cf. Fig. 1. This includes K variable nodes (one for each user i), N function nodes (one for each chip a) and an edge joining user i and chip a whenever $i \in \partial a$.

If signatures are chosen according to the proposed scheme, the resulting factor graph is a sparse random graph. The degree distribution is Ω_l on the variable node (user) side, and converges to a Poisson distribution with mean $\overline{l}\alpha$ on the function node (chip) side.

BP is introduced in the standard way: we limit ourselves to writing down the update equations in terms of log-likelihoods[2]. Two types of messages are updated: variable to function node, $v_{i\to a}$, and function to variable node, $u_{a\to i}$. The update equations read

$$v^t_{i\to a} = \sum_{b\in\partial i\setminus a} u^{t-1}_{b\to i}\,, \tag{6}$$

$$u^t_{a\to i} = f(v_{j\to a}, s_{ja}, j\in\partial a\setminus i;\, s_{ia};\, y_a)\,, \tag{7}$$

where the index t denotes the iteration number and

$$f(v_1,s_1,\dots,v_k,s_k;\, s_0;\, y) \equiv \frac{1}{2}\log\frac{W_+}{W_-}\,, \tag{8}$$

$$W_{\xi_0} \equiv \sum_{\xi_1\dots\xi_k=\pm1} e^{-\frac{1}{2\sigma^2}\left(y-\sum_{i=0}^{k}\frac{s_i}{\sqrt{l}}\xi_i\right)^2}\prod_{i=1}^{k} e^{v_i\xi_i}\,. \tag{9}$$

We furthermore adopt the initial condition $u^0_{a\to i} = v^0_{i\to a} = 0$. After a fixed number of iterations, all the messages incoming at variable node i are combined to compute the decision $x_i^{{\rm BP},t} \equiv \mathrm{sign}\{\sum_{a\in\partial i} u^t_{a\to i}\}$.

[2] More precisely, we use here *one half* of log-likelihoods.

B. Main results

In order to state and prove our main results more easily, it is convenient to focus onto ‘Poisson’ signature schemes. By this we mean that $\{\Omega_l, l \geq 0\}$ is a Poisson distribution of mean $\bar{l}$. We shall come back to the general case in Sections III-A and IV. Within this setting, we consider the expected conditional entropy per user $\mathbb{E}\, H(X|Y)/K$ (the expectation being taken with respect to the random signatures). Since we do not know *a priori* whether the large system limit exists, we define $\overline{h}(\sigma^2,\alpha,\bar{l}) \equiv \limsup_{N\to\infty} \mathbb{E}\, H(X|Y)/K$, and $\underline{h}(\sigma^2,\alpha,\bar{l}) \equiv \liminf_{N\to\infty} \mathbb{E}\, H(X|Y)/K$. In both cases, the limit is taken keeping the ratio $K/N = \alpha$ fixed.

If we let $\bar{l} \to N$ and then $N \to \infty$, we would recover the standard dense signature scheme (strictly speaking this corresponds to Ω_l concentrated on $l = N$). Here we shall invert the order of the two limits and let $N \to \infty$ and then $\bar{l} \to \infty$ *afterwards*. Our first result shows that, if the limit is taken in this way, Tanaka formula is correct. For our proof technique to work α must be smaller than the ‘spinodal value’ $\alpha_{\rm s}$. This is the largest number such that, for any $\alpha < \alpha_{\rm s}$ the solution to Eq. (3) is unique for all $\sigma^2 \in [0,\infty)$, and is a differentiable function of σ^2. By solving Eq. (3) numerically, we get $\alpha_{\rm s} \approx 1.49$.

Theorem 1: If $\alpha < \alpha_{\rm s}$, then the per-user conditional entropy converges to Tanaka’s formula in the dense signature limit

$$\lim_{l\to\infty} \overline{h}(\sigma^2,\alpha,\bar{l}) = \lim_{l\to\infty} \underline{h}(\sigma^2,\alpha,\bar{l}) = h_{\rm RS}(\sigma^2,\alpha)\,. \tag{10}$$

The hypothesis of Poisson signatures is presently used only in the proof of Lemma 1. It shouldn’t however be difficult to extend this result to more general sequences of degree distributions Ω_l.

A key step in the proof of the above result consists in analizing the BP-based detection algorithm defined by Eqs. (6), (7). Our second result shows that, in the small α regime this algorithm is indeed optimal (the proof of this result is deferred to a longer paper).

Theorem 2: Let $\overline{\mathrm{P}_{\rm b}}(\bar{l},N)$ be the expected bit error rate under symbol MAP detection, and $\overline{\mathrm{P}^{\rm BP}_{\rm b}}(\bar{l},N;t)$ the same quantity for t iterations BP detection. Define the asymptotic BP error overhead as

$$\Delta(\bar{l};t) = \limsup_{N\to\infty} [\overline{\mathrm{P}^{\rm BP}_{\rm b}}(\bar{l},N;t) - \overline{\mathrm{P}_{\rm b}}(\bar{l},N)]\,. \tag{11}$$

If $\alpha < \alpha_{\rm s}$, then BP is optimal in the dense signature limit, namely $\lim_{t\to\infty}\lim_{\bar{l}\to\infty}\Delta(\bar{l};t) = 0$.

III. A SKETCH OF THE PROOF

A. A few simple remarks

We start by collecting a few remarks whose proof is routine, and therefore omitted apart from a few hints.

All $+1$ input. For the sake of analysis (and for proving Theorem 1) we can assume that the input signal is $x = x_+ \equiv (+1,\dots,+1)^{\rm T}$. In particular, if we write $\mathbb{E}^+_{y,\mathbb{S}}$ for the joint expectation with respect to y and $\mathbb{S}$, conditional to $x = x_+$, then $\mathbb{E}_{\mathbb{S}} H(X|Y) = -\mathbb{E}_{x,y,\mathbb{S}} \log \mathbb{P}(X|Y,\mathbb{S}) = -\mathbb{E}^+_{y,\mathbb{S}} \log \mathbb{P}(X = x_+|Y,\mathbb{S})$.

Density evolution.(DE) Any finite neighborhood of a randomly chosen node in the factor graph associated to the sparse signature scheme, converges in distribution to a tree with the degree distribution mentioned above. As a consequence, the messages distribution can be analyzed through a standard DE approach.

Define the sequence of random variables $\{v^t, u^t;\, t \geq 0\}$ as follows: $v^0 = u^0 = 0$, and

$$v^{t+1} \overset{\rm d}{=} \sum_{b=1}^{l} u^t_b\,, \quad u^t \overset{\rm d}{=} f(v^t_1, s_1, \dots, v^t_k, s_k;\, s_0; y)\,, \tag{12}$$

for $t \geq 0$. Here $\overset{\rm d}{=}$ denotes identity in distribution; $u^t_1, u^t_2, \dots$ (respectively, $v^t_1, v^t_2, \dots$) are i.i.d. copies of u^t (respectively, of v^t); l is an integer random variable with distribution ω_l, and k is a Poisson random variable with mean $\bar{l}\alpha$; finally $s_0, \dots, s_k$ are i.i.d.’s with $s_i \in \{+1,-1\}$ uniformly at random, $y = \frac{1}{\sqrt{\bar{l}}}\sum_{i=0}^{k} s_i + w$ with w a normal random variable with mean 0 and variance σ^2.

Let (ia) be a uniformly random edge in the factor graph and $v^t_{i\to a}$, $u^t_{a\to i}$ the corresponding BP messages, under the assumption that x_+ has been transmitted. Then $v^t_{i\to a}$ (respectively $u^t_{a\to i}$) converges in distribution to v^t (respectively, to u^t) as $N \to \infty$.

Symmetry condition. A random variable X is ‘symmetric’ if $\mathbb{E}[f(-X)] = \mathbb{E}[\,e^{-2X} f(X)]$ for any function f such that both expectation exist. It is easy to show that the random variables u^t, v^t defined above are symmetric (this is analogous to what happens in LDPC codes).

Area theorem. Following [7], the derivative, with respect to the noise parameter, of the conditional entropy is proportional to the expectation of the conditional variance

$$\frac{{\rm d}H(X|Y)}{{\rm d}\sigma^2} = \frac{1}{2\sigma^4}\mathbb{E}_y\left\{{\rm Var}(\mathbb{S}X|Y)\right\}\,. \tag{13}$$

Let us take the expectation with respect to the signatures $\mathbb{S}$, and normalize by the number of users. Using the all $+1$ assumption, we get (derivative and expectation can be interchanged because $H(X|Y)$ has positive bounded derivative, see below)

$$\begin{aligned}\frac{1}{K}\frac{{\rm d}\mathbb{E}H(X|Y)}{{\rm d}\sigma^2} &= \frac{1}{2\sigma^4}\cdot \\ &\cdot \frac{1}{N\bar{l}\alpha}\sum_{a=1}^{N}\mathbb{E}^+_{y,\mathbb{S}}\left\{|\partial a| - \Big(\sum_{i\in\partial a} s_{ia}\widehat{x}_i\Big)^2\right\}\,,\end{aligned} \tag{14}$$

where $\widehat{x}_i = \widehat{x}_i(Y,\mathbb{S}) \equiv \mathbb{E}[X_i|Y,\mathbb{S}]$. We shall sometimes refer to the right hand side as to the GEXIT function and denote it by $g_N(\alpha,\sigma^2)$. From the above expressions it is easy to realize that $0 \leq g_N(\alpha,\sigma^2) \leq 1/2\sigma^4$. The same inequalities also hold at fixed $\mathbb{S}$, which justifies the exchange of derivative and expectation above.

As in Refs. [3], [4], [5], we introduce furthermore the BP GEXIT function $g^t_{\rm BP}(\alpha,\sigma^2)$, with t a non-negative integer. This is defined by replacing the expectation $\sum_{i\in\partial a} s_{ia}\widehat{x}_i$ on the right hand side of Eq. (14) by its estimate after t iterations of BP (in the $N \to \infty$ limit). In terms of the DE variables

$$g^t_{\rm BP}(\alpha, \sigma^2) = \frac{1}{2\sigma^4 \bar{l}\alpha}\mathbb{E}\left\{ k - \left\langle \sum_{i=1}^{k} s_i \xi_i \right\rangle^2 \right\}, \qquad (15)$$

where $\langle \cdot \rangle$ denotes an average over $\xi_i \in \{+1, -1\}$ with distribution

$$\nu(\{\xi_i\}) = \frac{1}{\Xi}\, e^{-\frac{1}{2\sigma^2}\left(w + \frac{1}{\sqrt{l}}\sum_{i=1}^{k} s_i(1-\xi_i)\right)^2} \prod_{i=1}^{k} e^{v_i^t \xi_i}\,, \qquad (16)$$

and the expectation $\mathbb{E}$ is taken with respect to $\{v_i^t\}$ (i.i.d. and distributed as v^t from DE), $\{s_i\}$ (i.i.d. uniform in $\{+1, -1\}$), w (gaussian with mean zero and variance σ^2), and k (Poisson with mean $\bar{l}\alpha$).

B. The proof

The proof of Theorem 1 makes use of three lemmas, which we state without demonstration for lack of space. As in Section II-B, $\overline{h}(\alpha, \sigma^2, \bar{l})$ and $\underline{h}(\alpha, \sigma^2, \bar{l})$ denote, respectively, the $\limsup$ and $\liminf$ of the expected conditional entropy per bit, in the system with Poisson signatures.

The first lemma states that, in the low noise limit, the input can be reconstructed faithfully from the transmitted message and therefore the conditional entropy per bit vanishes (recall that we are dealing with discrete inputs).

Lemma 1: For any $\alpha > 0$, $\lim_{\sigma^2 \to 0} \lim_{\bar{l} \to \infty} \overline{h}(\alpha, \sigma^2, \bar{l}) = 0$.
The proof is based on a union bound, and a combinatorial calculation.

The second lemma provides upper and lower bounds on the conditional entropy per user, in terms of BP GEXIT functions. For the sake of definiteness, we state the lemma for Poisson signatures (and denote the corresponding BP GEXIT functions as $g^t_{\rm BP}(\alpha, {\sigma'}^2, \bar{l})$) although it obviously holds in greater generality [5].

Lemma 2: For any $\bar{l} > 0$, $\sigma_0^2 > 0$, and non-negative integer t

$$1 - \int_{\sigma^2}^{\infty} g^t_{\rm BP}(\alpha, {\sigma'}^2, \bar{l})\, {\rm d}{\sigma'}^2 \le \underline{h}(\alpha, \sigma^2, \bar{l}) \le \qquad (17)$$

$$\le \overline{h}(\alpha, \sigma^2, \bar{l}) \le \overline{h}(\alpha, \sigma_0^2, \bar{l}) + \int_{\sigma_0^2}^{\sigma^2} g^t_{\rm BP}(\alpha, {\sigma'}^2, \bar{l})\, {\rm d}{\sigma'}^2\,.$$

This is in fact an easy consequence of the general result that GEXIT functions preserve physical degradation [5].

Finally, a Lemma on the large $\bar{l}$ limit of DE.

Lemma 3: Define the sequence $\{\lambda_t;\ t \ge 0\}$ by setting $\lambda_0 = 0$ and

$$\lambda_{t+1} = \left\{ \sigma^2 + \alpha \left[1 - \mathsf{E}_z \tanh^2(\lambda_t + \sqrt{\lambda_t}\, z) \right] \right\}^{-1}, \qquad (18)$$

for any $t \ge 0$. Let $\{v^t;\ t \ge 0\}$ be the solution of DE for the system with Poisson signatures (with mean $\bar{l}$) and the same values of σ^2 and α. Then, for any $t \ge 0$, v^t converges in distribution to a gaussian random variable with mean λ_t and variance λ_t as $\bar{l} \to \infty$.

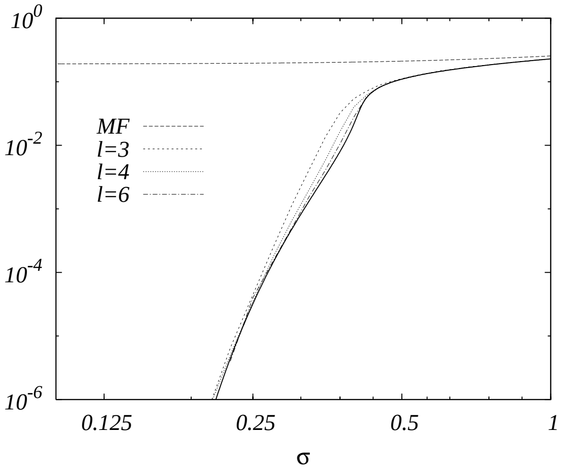

Fig. 2. The bit error rate as a function of the noise parameter σ at $\alpha = 1.3$. The bold continuous line is Tanaka's result for dense signatures under symbol MAP detection. MF refer to the same signature scheme under matched filter detection. The other (dashed) lines correspond to sparse signatures and BP detection.

The proof is based on a repeated application of the central limit theorem (the argument can be written as an induction over t). The reader is invited to try, for instance, with $t = 1, 2, \dots$.

Let us now turn to the proof of Theorem 1. We start by using Lemma 3 to compute the large $\bar{l}$ limit of the BP GEXIT functions. After a simple application of central limit theorem, we get

$$\lim_{\bar{l} \to \infty} g^t_{\rm BP}(\alpha, \sigma^2, \bar{l}) = \frac{1}{2\sigma^2}\, \frac{(1 - q_t)}{\sigma^2 + \alpha(1 - q_t)}\,, \qquad (19)$$

where $q_t \equiv \mathsf{E}_z \tanh^2(\lambda_t + \sqrt{\lambda_t}\, z)$. We shall denote the expression on the right hand side of Eq. (19) as $g^t_{\rm BP}(\alpha, \sigma^2)$.

Next, we use Lemma 2. Noticing that $0 \le g^t_{\rm BP}(\alpha, \sigma^2, \bar{l}) \le 1/4\sigma^2$ we can apply the dominated convergence theorem to take the $\bar{l} \to \infty$ limit in Eq. (17). If we take $\sigma_0^2 \to 0$ afterwards and apply Lemma 1, we get

$$1 - \int_{\sigma^2}^{\infty} g^t_{\rm BP}(\alpha, {\sigma'}^2)\, {\rm d}{\sigma'}^2 \le \underline{h}(\alpha, \sigma^2, \infty) \le \qquad (20)$$

$$\le \overline{h}(\alpha, \sigma^2, \infty) \le \int_{0}^{\sigma^2} g^t_{\rm BP}(\alpha, {\sigma'}^2)\, {\rm d}{\sigma'}^2\,,$$

where $\underline{h}(\alpha, \sigma^2, \infty) \equiv \liminf_{\bar{l} \to \infty} \underline{h}(\alpha, \sigma^2, \bar{l})$, and $\overline{h}(\alpha, \sigma^2, \infty) \equiv \limsup_{\bar{l} \to \infty} \overline{h}(\alpha, \sigma^2, \bar{l})$.

Simple calculus shows that λ_t is strictly positive and increasing in t for $t \ge 1$, and $\lambda_t \simeq \sigma^{-2}$ as $\sigma \to 0$. Furthermore $\lim_{t \to \infty} \lambda_t = \lambda_{\rm BP}$ is the smallest positive fixed point of the recursion (18), i.e. the smallest positive solution of Tanaka's stationarity equation (3).

From these remarks, it follows that $g^t_{\rm BP}(\alpha, \sigma^2)$ is integrable over $\sigma \in [0, \infty)$ and strictly decreasing in $t \ge 1$. We can

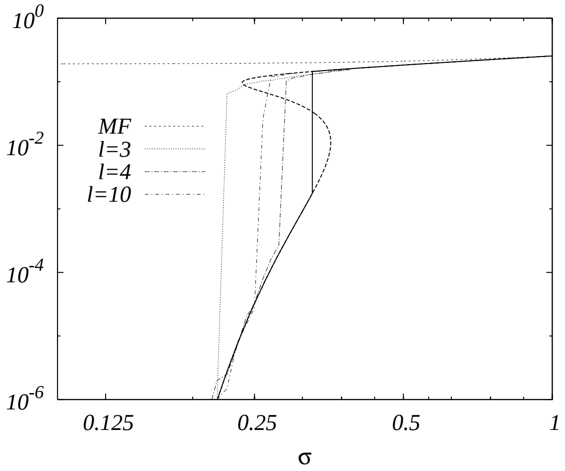

Fig. 3. Same as in Fig. 2 but for $\alpha = 1.9$. The S shaped dashed curve is the analytical continuation of the bit error rate for dense signatures.

therefore take the $t \to \infty$ limit of Eq. (20) to get

$$1 - \int_{\sigma^2}^{\infty} g_{\rm BP}(\alpha, \sigma'^2)\, {\rm d}\sigma'^2 \le \underline{h}(\alpha, \sigma^2, \infty) \le \qquad (21)$$

$$\le \overline{h}(\alpha, \sigma^2, \infty) \le \int_0^{\sigma^2} g_{\rm BP}(\alpha, \sigma'^2)\, {\rm d}\sigma'^2 \,,$$

where we defined

$$g_{\rm BP}(\alpha, \sigma^2) \equiv \lim_{t\to\infty} g^t_{\rm BP}(\alpha, \sigma^2) = \frac{1}{2\sigma^2}\, \frac{1 - q_{\rm BP}}{\sigma^2 + \alpha(1 - q_{\rm BP})}\,,$$

and $q_{\rm BP} = {\sf E}_z \tanh(\lambda_{\rm BP} + \sqrt{\lambda_{\rm BP}} z)$.

We are left with the task of showing that the first and the last expressions in Eq. (21) do indeed coincide and are both equal to Tanaka's formula $h_{\rm RS}(\alpha, \sigma^2)$. Recall that, for $\alpha < \alpha_{\rm s}$, the stationarity equation (20) admits a unique solution depending smoothly on σ^2. Furthermore, we saw above that this coincides with the BP fixed point. Using these remarks, we can differentiate Eq. (2) with respect to σ^2, to get

$$\frac{\partial h_{\rm RS}}{\partial \sigma^2}(\alpha, \sigma^2) = g_{\rm BP}(\alpha, \sigma^2)\,. \qquad (22)$$

The proof is completed by applying the fundamental theorem of calculus to Eq. (21) and noticing that $h_{\rm RS}(\alpha, 0) = 0$ and $h_{\rm RS}(\alpha, \infty) = 1$. □

IV. Numerical simulations

One may wonder how quickly is the $\bar{l} \to \infty$ limit in Theorems 1 and 2 attained. In Fig. 2 we show the results of numerical simulations using DE, and regular signatures (Ω_l concentrated on a single value), for $\alpha = 1.3 < \alpha_{\rm s}$. Already at $l = 4$ the bit error rate is extremely close to the dense limit!

Even more surprising is the behavior for $\alpha > \alpha_{\rm s}$. In Fig. 3 we show the data for $\alpha = 1.9$. The BP error rate at $l = 4$ is close to the MAP one with dense signatures. However it worsens at l grows (and seems to approach the natural guess for BP behavior with dense signatures). Sparse signatures are the crucial ingredient allowing for low complexity detection and close-to-optimal performances.

References

[1] T. Richardson, and R. Urbanke, “Modern Coding Theory”, draft available at `http://lthcwww.epfl.ch/index.php`

[2] S. Ciliberti, M. Mézard, and R. Zecchina, “Lossy Data Compression with Random Gates”, *Phys. Rev. Lett.* 95, 038701 (2005)

[3] C. Méasson, A. Montanari, T. Richardson, and R. Urbanke, “Life Above Threshold: From List Decoding to Area Theorem and MSE”, *Proc. of the IEEE Inform. Theory Workshop*, San Antonio, Texas, October 24-29 2004. Available online at `http://arxiv.org/abs/cs.IT/0410028`.

[4] C. Méasson, A. Montanari, and R. Urbanke, “Maxwell's Construction: The Hidden Bridge between Iterative and Maximum A Posteriori Decoding”. Submitted to *IEEE Trans. Inform. Theory*. Available online at `http://arxiv.org/abs/cs.IT/0506083`.

[5] C. Méasson, A. Montanari, T. Richardson, and R. Urbanke. “The Generalized Area Theorem and Some of its Consequences”. Submitted to *IEEE Trans. Inform. Theory*. Available online at `http://arxiv.org/abs/cs.IT/0506083`.

[6] A. Ashikhmin, G. Kramer and S. ten Brink, “Extrinsic Information Transfer Functions: model and erasure channel property”, *IEEE Trans. Inform. Theory*, vol. 50, pp. 2657–2673, 2004.

[7] D. Guo, S. Shamai, and S. Verdú, “Mutual information and minimum mean-square error in Gaussian channels”, *IEEE Trans. Inform. Theory*, vol. 51, pp. 1261–1282, 2005.

[8] D. Guo, S. Shamai, and S. Verdú, “Mutual information and conditional mean estimation in Poisson channels”, *Proc. of the IEEE Inform. Theory Workshop*, San Antonio, Texas, October 24-29 2004.

[9] S. Verdú, “Multiuser Detection”, Cambridge University Press, Cambridge, 1998

[10] D. Tse and S. V. Hanly, “Linear Multiuser Receivers: Effective Interference, Effective Bandwidth and User Capacity”, *IEEE Trans. Inform. Theory*, vol. 45, pp. 641-657, 1999

[11] S. Verdu and S. Shamai, “Spectral Efficiency of CDMA with Random Spreading”, *IEEE Trans. Inform. Theory*, vol. 45, pp. 622-640, 1999

[12] A. Montanari, B. Prabhakar, and D. Tse, “Belief Propagation-Based Multi-User Detection”, Proc. of the Forty-Third Allerton Conference on Communications, Control and Computing, Monticello, Illinois, October 2005. Available online at `http://arxiv.org/abs/cs.IT/0510044`.

[13] T. Tanaka, “A Statistical–Mechanics Approach to Large–Systems Analysis of CDMA Multiuser Detectors”, *IEEE Trans. Inform. Theory*, vol. 48, pp. 2888-2910, 2002

[14] Y. Kabashima, “A CDMA Multiuser Detection Algorithm on the Basis of Belief Propagation”, *J. Phys. A: Math. Gen.*, vol. 36, pp. 11111–11121, 2003

[15] T. Tanaka and M. Okada, “Approximate Belief Propagation, Density Evolution, and Neurodynamics for CDMA Multiuser Detection”, *IEEE Trans. Inform. Theory*, vol. 51, pp. 700-706, 2005

[16] J. P. Neirotti and D. Saad. “Improved message passing for inference in densely connected systems”. *Europhys. Lett.* vol. 71, 866-872, (2005).

[17] M. Luby, “LT-codes”, *Proceedings of FOCS-43*, 271, 2002