

On the concentration of the capacity for a code division multiple access system

Satish Babu Korada and Nicolas Macris
 School of computer and communication science,
 Ecole Polytechnique Fédérale de Lausanne
 EPFL - LTHC - IC - Station 14
 Lausanne CH-1015, Switzerland
 Email: {satish.korada, nicolas.macris}@epfl.ch

Abstract— We prove the concentration of the capacity, in the large system limit, for a code division multiple access system over an additive white Gaussian noise channel, with Gaussian signature sequences and binary input symbols. The probabilistic tools that are used are quite powerful and could have applications in many other similar situations.

I. INTRODUCTION AND MAIN RESULTS

We consider the following setting for randomly spread code division multiple access (CDMA) communication on an additive white Gaussian noise (AWGN) channel. K users send binary information symbols $\underline{x} = (x_1, \dots, x_K)^t$, $x_k \in \{\pm 1\}$, through a single AWGN channel. Each user has a signature sequence $\underline{s}_k = (s_{1k}, \dots, s_{Nk})^t$ assumed to be known to the receiver. At each time interval $i = 1, \dots, N$ the received signal is $\frac{1}{\sqrt{N}} \sum_{k=1}^K s_{ik} x_k + n_i$ where $\underline{n} = (n_1, \dots, n_N)^t$ are independent identically distributed Gaussian variables $\mathcal{N}(0, \sigma^2)$ and σ is the noise amplitude. The scaling factor $1/\sqrt{N}$ is introduced so that the energy of each user per information bit is normalized to 1. Here we take s_{ik} as generated randomly from independent identically distributed standard Gaussians $\mathcal{N}(0, 1)$ and denote the corresponding $K \times N$ random matrix as \mathbf{s} . The received vector $\underline{y} = (y_1, \dots, y_N)^t$ can be expressed as

$$\underline{y} = \frac{1}{\sqrt{N}} \mathbf{s} \underline{x} + \underline{n}$$

If all users communicate at the same rate R the capacity is the maximum achievable such rate

$$C(\mathbf{s}) = \frac{1}{K} \max_{p(\underline{x})} I(\underline{X}; \underline{Y} | \mathbf{s})$$

A priori this quantity depends on the realisation of the random matrix \mathbf{s} . The main result in this contribution is a proof that in the large system limit (K and $N \rightarrow +\infty$ with $K/N = \beta$ fixed) $C(\mathbf{s})$ concentrates on its mean $\mathbb{E}_{\mathbf{s}}[C(\mathbf{s})]$.

Tanaka [1] has derived analytical expressions, in the large system limit, for $\mathbb{E}_{\mathbf{s}}[C(\mathbf{s})]$ and the bit error rate, thanks to the formal “replica trick” of statistical mechanics. Though the replica method is not rigorous it is generally conjectured to yield an exact result. In the case of Gaussian inputs this can be explicitly verified by comparing the expressions resulting from the replica method and the rigorous random matrix calculations [4],[3]. In the present case with binary inputs

the usual random matrix techniques do not work but the conjecture is again supported by the recent work of Montanari and Tse [2] who proved that Tanaka’s formula is rigorous at least for regimes of parameters (β and σ) without phase transitions (the bit error rate has no jumps). The extension of such proofs to regimes with phase transitions (the bit error rate is discontinuous) is to the best of our knowledge still open.

However, for the case of binary inputs, these works do not address the underlying question of the concentration of relevant quantities such as the capacity and the bit error rate. Let us mention that in the case of a Gaussian input signal, the concentration can be deduced from general theorems on the concentration of the spectral density for random matrices, but this approach breaks down for binary inputs.

In this contribution we provide all the details of a proof of concentration for the capacity thanks to powerful probabilistic tools developed by Talagrand [5], [6]. The mathematical underpinning of this result is in fact the concentration of a “free energy” which we also prove. We also indicate why the considerations about the free energy could conceivably allow to attack the more difficult problem of the concentration of the bit error rate.

In the sequel $\|\cdot\|$ always denotes an appropriate Euclidean norm.

From general arguments the maximum of the mutual information is attained for a uniform input distribution $p(\underline{x}) = 2^{-K}$,

$$C(\mathbf{s}) = \ln 2 - \frac{1}{K} H(\underline{X} | \underline{Y}, \mathbf{s})$$

where the conditional entropy is the average (over \underline{Y} only) Shannon entropy of the posterior distribution

$$p(\underline{x} | \underline{y}, \mathbf{s}) = \frac{e^{-\frac{1}{2\sigma^2} \|\mathbf{s} \underline{x} - \underline{y}\|^2}}{Z(\underline{y}, \mathbf{s})} \quad (1)$$

with the normalization factor

$$Z(\underline{y}, \mathbf{s}) = \sum_{\underline{x}} e^{-\frac{1}{2\sigma^2} \|\mathbf{s} \underline{x} - \underline{y}\|^2} \quad (2)$$

The average over \underline{Y} is carried out with the distribution

$$\begin{aligned} p(\underline{y} | \mathbf{s}) &= \sum_{\underline{x}^0} p(\underline{y} | \underline{x}^0, \mathbf{s}) p(\underline{x}^0) \\ &= \frac{1}{2^K} \sum_{\underline{x}^0} \frac{e^{-\frac{1}{2\sigma^2} \|\mathbf{N}^{-\frac{1}{2}} \mathbf{s} \underline{x}^0 - \underline{y}\|^2}}{(\sqrt{2\pi}\sigma^2)^N} \end{aligned}$$

where \underline{x}^0 is to be interpreted as an input signal.

Although in this contribution we do not use any specific notion from statistical mechanics, it is insightful to view (1) as the Gibbs measure of a “spin glass” where $x_k \in \{-1, +1\}$ are the spin variables, and (given) realizations $\underline{y}, \mathbf{s}$ the “quenched disorder”. The normalization factor (2) is called the “partition function” and $\ln Z(\underline{y}, \mathbf{s})$ is the “free energy”. In the appendix we deduce a relationship between the capacity and the average free energy of the spin glass, namely

$$C(\mathbf{s}) = \ln 2 - \frac{1}{2\beta} - \frac{1}{K} \mathbb{E}_{\underline{Y}}[\ln Z(\underline{y}, \mathbf{s})] \quad (3)$$

Let us mention that the “underlying graph” of the spin glass is the *complete graph* with K vertices, because if the norm in (1) is expanded all pairs $x_k x_l$ are present. The situation is thus similar to the classic Sherrington-Kirkpatrick model [9] for which the self averaging property of the free energy was first shown by Shcherbina and Pastur in [10]. The exponential concentration results for the free energy have appeared only recently and this has been achieved by Guerra and Toninelli by an “interpolation method” [8]. However, in the present case, the situation is more complicated and it is not clear how to apply Guerra and Toninelli’s method¹. The concentration properties stated in the following theorems are not expected to be optimal.

Theorem 1: [concentration of the capacity.] There exists an integer N_1 and a strictly positive constant α_1 such that for all $N > N_1$ and all $t > 0$,

$$\mathbb{P}[|C(\mathbf{s}) - \mathbb{E}_{\mathbf{s}}[C(\mathbf{s})]| \geq t] \leq 3e^{-\alpha_1 t^2 N}$$

One can take $\alpha_1 = \sigma^4 / (16\beta(64\beta + 32 + \sigma^2))$.

We also prove the slightly more general, and obviously related, concentration result for the free energy itself,

Theorem 2: [concentration of free energy.] There exists an integer N_2 and a strictly positive constant α_2 such that for all $N \geq N_2$ and all $t > 0$,

$$\mathbb{P}[K^{-1} |\ln Z(\underline{y}, \mathbf{s}) - \mathbb{E}_{\underline{Y}, \mathbf{s}}[\ln Z(\underline{y}, \mathbf{s})]| \geq t] \leq 3e^{-\alpha_2 t^2 \sqrt{N}}$$

One can take $\alpha_2 = \sigma^4 \beta^2 / (32(2\sqrt{\beta} + \sigma)^2)$.

We wish to argue here that theorem 2 suggests a method for proving the concentration of the bit error rate. Given the observed output signal, the optimal bit estimate is defined through the marginal $p(x_k | \underline{y}, \mathbf{s})$ of (1)

$$\hat{x}_k = \operatorname{argmax}_{x_k \in \{\pm 1\}} p(x_k | \underline{y}, \mathbf{s})$$

¹As already stated we use other tools which by the way can also be applied to the Sherrington-Kirkpatrick model.

The bit error rate is the overlap with the input \underline{x}_0 ,

$$\frac{1}{2} \left(1 - \frac{1}{K} \sum_{k=1}^K x_{0,k} \hat{x}_k \right) \quad (4)$$

The connection with statistical mechanics comes from the remark that $\hat{x}_k = \operatorname{sign}\langle x_k \rangle$ where $\langle - \rangle$ is the average with respect to the Gibbs measure (1). The “magnetization” $\langle x_k \rangle$ can be obtained from the free energy by adding first an infinitesimal “external magnetic field” to the exponent in (1), namely $h \sum_{k=1}^K x_k^0 x_k$, and then differentiating the free energy,

$$\sum_{k=1}^K x_k^0 \langle x_k \rangle = \lim_{h \rightarrow 0} \frac{d}{dh} \ln Z(\underline{y}, \mathbf{s})$$

Similarly one can use “real replicas” (copies of the Gibbs measure) to obtain by differentiation of the replicated free energy $\sum_{k=1}^K (x_k^0 \langle x_k \rangle)^n$ and eventually reconstruct (4). Then by rather standard arguments the (easily verifiable) convexity of the free energy with respect to h should enable one to deduce concentration of (4) from that of the free energy. However the completion of this last point requires also a uniform (with respect to system size) control of the derivative of the free energy precisely at $h = 0$, which at present is still lacking. We hope to come back to these considerations in the future.

II. PROBABILISTIC TOOLS

Our proofs rely on a general concentration theorem for suitable *Lipschitz functions of many Gaussian random variables* [5], [6] and this is why we need Gaussian signature sequences. In the version that we use here we need functions that are *Lipschitz with respect to the Euclidean distance*. More precisely we say that a function $f : \mathbb{R}^M \rightarrow \mathbb{R}$ is a Lipschitz function with constant L_M if for all $(\underline{u}, \underline{v}) \in \mathbb{R}^M \times \mathbb{R}^M$

$$|f(\underline{u}) - f(\underline{v})| \leq L_M \|\underline{u} - \underline{v}\|$$

When another distance is used the function will still be Lipschitz but one has to carefully keep track of the possibly qualitatively different M dependence.

Theorem 3: [5] Let U_1, \dots, U_M be M independent identically distributed Gaussian random variables with distribution $\mathcal{N}(0, v^2)$ and let $f : \mathbb{R}^M \rightarrow \mathbb{R}$ be Lipschitz with respect to the Euclidean distance, with constant L_M . Then f satisfies

$$\mathbb{P}[|f(u_1, \dots, u_M) - \mathbb{E}[f(u_1, \dots, u_M)]| \geq t] \leq 2e^{-\frac{t^2}{2v^2 L_M^2}}$$

In our application it will not be possible to apply directly this theorem because the relevant function is Lipschitz only on a subset $G \subset \mathbb{R}^M$. It turns out that the measure of the complement G^c is negligible as $M \rightarrow +\infty$. For the “good part” of the function supported on G we will use the following result of McShane and Whitney

Theorem 4: [7] Let $f : G \rightarrow \mathbb{R}$, be Lipschitz over $G \subset \mathbb{R}^M$ with constant L_M . Then there exists an extension $g : \mathbb{R}^M \rightarrow \mathbb{R}$

such that $g|_G = f$ which is Lipschitz with the same constant over the whole of \mathbb{R}^M .

From these two theorems we can prove the

Lemma 1: Let f and g be as in theorem 4. Assume $0 \in G$ and $\mathbb{E}[f(\underline{u})^2] \leq C^2$, $f(0)^2 \leq C^2$ for some positive number C . Then for

$$\frac{t}{2} \geq 3(C + v\sqrt{M})\sqrt{\mathbb{P}(G^c)}$$

we have

$$\mathbb{P}[|f(\underline{u}) - \mathbb{E}[f(\underline{u})]| \geq t] \leq 2e^{-\frac{t^2}{8v^2L_M^2}} + \mathbb{P}[G^c]$$

Proof: We drop the \underline{u} dependence to lighten the notation. Notice that $0 \in G$ implies $f(0) = g(0)$. Thus $g(0)^2 \leq C^2$. Also, since g is Lipschitz on the whole of \mathbb{R}^M

$$\begin{aligned} \mathbb{E}[g^2] &\leq 2(g(0)^2 + \mathbb{E}[(g - g(0))^2]) \\ &\leq 2(C^2 + L_M \mathbb{E}[\|\underline{u}\|^2]) \\ &= 2(C^2 + Mv^2L_M) \end{aligned}$$

Furthermore on G we have $g = f$, so by the Cauchy-Schwartz inequality

$$\begin{aligned} |\mathbb{E}[g - f]| &= |\mathbb{E}[(g - f)1_{G^c}]| \\ &\leq (\mathbb{E}[g^2]^{1/2} + \mathbb{E}[f^2]^{1/2})\sqrt{\mathbb{P}[G^c]} \\ &\leq (C + \sqrt{2}(C^2 + Mv^2L_M)^{1/2})\sqrt{\mathbb{P}[G^c]} \\ &\leq 3(C + v\sqrt{M}L_M)\sqrt{\mathbb{P}[G^c]} \leq \frac{t}{2} \end{aligned}$$

Moreover

$$\begin{aligned} \mathbb{P}[|f - \mathbb{E}f| \geq t] &= \mathbb{P}[|g - \mathbb{E}f| \geq t \mid \underline{U} \in G]\mathbb{P}[G] \\ &\quad + \mathbb{P}[|f - \mathbb{E}f| \geq t \mid \underline{U} \in G^c]\mathbb{P}[G^c] \\ &\leq \mathbb{P}[|g - \mathbb{E}g| \geq t - |\mathbb{E}g - \mathbb{E}f|] + \mathbb{P}[G^c] \end{aligned}$$

The result of the lemma then follows from

$$\mathbb{P}[|g - \mathbb{E}g| \geq t - |\mathbb{E}g - \mathbb{E}f|] \leq \mathbb{P}[|g - \mathbb{E}g| \geq \frac{t}{2}]$$

and the application of theorem 3. \blacksquare

In order to prove theorems 1 and 2 it will be sufficient to find suitable sets G with measure nearly equal to one (as $M \rightarrow +\infty$), on which the capacity and free energy have a Lipschitz constant $L_M \rightarrow 0$.

III. PROOFS OF THEOREMS 1 AND 2

For the proofs, it is convenient to reformulate the statements of the theorems as follows. Let $\underline{1}$ be the K dimensional vector $(1, \dots, 1)$, \mathbf{s}^0 be the $K \times N$ matrix with elements $s_{ik}x_k^0$, and consider the partition function

$$Z'(\underline{n}, \mathbf{s}^0) = \sum_{\underline{x}} e^{-\frac{1}{2\sigma^2}\|N^{-1/2}\mathbf{s}^0(\underline{x}-\underline{1})-\sigma\underline{n}\|^2} \quad (5)$$

where we recall that $\underline{n} = (n_1, \dots, n_N)$ are independent Gaussian variables $\mathcal{N}(0, 1)$. Notice that due to the invariance of the distribution of s_{ik} under the transformation $s_{ik} \rightarrow x_k^0 s_{ik}$,

$$\mathbb{E}_{\underline{N}, \mathbf{s}}[\ln Z'(\underline{n}, \mathbf{s}^0)] = \mathbb{E}_{\underline{N}, \mathbf{s}}[\ln Z'(\underline{n}, \mathbf{s})]$$

The statement of theorems 1 and 2 are equivalent to

$$\begin{aligned} \mathbb{P}[K^{-1}|2^{-K} \sum_{\underline{x}^0} \mathbb{E}_{\underline{N}}[\ln Z'(\underline{n}, \mathbf{s}^0)] - \mathbb{E}_{\underline{N}, \mathbf{s}}[\ln Z'(\underline{n}, \mathbf{s})]| \geq t] \\ \leq 3e^{-\alpha_1 t^2 N} \end{aligned} \quad (6)$$

and

$$\mathbb{P}[K^{-1}|\ln Z'(\underline{n}, \mathbf{s}^0) - \mathbb{E}_{\underline{N}, \mathbf{s}}[\ln Z'(\underline{n}, \mathbf{s})]| \geq t] \leq 3e^{-\alpha_2 t^2 \sqrt{N}} \quad (7)$$

To see this use the change of variable $\underline{y} = N^{-1/2}\mathbf{s}\underline{x}_0 + \sigma\underline{n}$ followed by $x_k \rightarrow x_k x_k^0$ in the partition function summation (2).

A. Proof of (6)

Let B be a positive constant to be chosen later and define

$$G = \{\mathbf{s} \mid \text{for all } \underline{x}, \mathbf{s}^0, \|\mathbf{s}^0(\underline{x} - \underline{1})\|^2 \leq BN\}$$

Lemma 2: We have the following estimate for the measure of G^c ,

$$\mathbb{P}(G^c) \leq 3^K 2^{\frac{N}{2}} e^{-\frac{B}{16\beta}}$$

Proof: First notice that for any given \underline{x} ,

$$\frac{1}{\sqrt{K}} \sum_{k=1}^K s_{ik}^0 (x_k - 1), \quad i = 1, \dots, N$$

are independent Gaussian random variables with zero mean and variance smaller than 4. Thus the identity

$$\int dx \frac{e^{-\frac{x^2}{2a^2}}}{\sqrt{2\pi a^2}} e^{\frac{x^2}{16}} = \left(1 - \frac{a^2}{8}\right)^{-\frac{1}{2}}$$

implies (because $a \leq 4$)

$$\mathbb{E}[e^{\frac{1}{16K}\|\mathbf{s}^0(\underline{x}-\underline{1})\|^2}] \leq 2^{\frac{N}{2}}$$

Then from the Markov inequality, for any \underline{x}

$$\mathbb{P}(\|\mathbf{s}^0(\underline{x} - \underline{1})\|^2 \geq BN) \leq 2^{\frac{N}{2}} e^{-\frac{BN}{16K}} = 2^{\frac{N}{2}} e^{-\frac{B}{16\beta}}$$

The result of the lemma then follows from the union bound. \blacksquare

We will apply lemma 2 to

$$f(\mathbf{s}) = \frac{1}{K2^{-K}} \sum_{\underline{x}^0} \mathbb{E}_{\underline{N}}[\ln Z'(\underline{n}, \mathbf{s}^0)]$$

for a suitable choice of B . In the application the matrix \mathbf{s} is to be thought as a vector with KN components and norm

$$\|\mathbf{s}\| = \left(\sum_{i=1}^N \sum_{k=1}^K s_{ik}^2 \right)^{\frac{1}{2}}$$

Clearly $0 \in G$ and $f(0)^2 = (\frac{1}{K} \mathbb{E}_{\underline{N}}[\frac{1}{2}\|\underline{n}\|^2])^2 = 1/4\beta^2$. Also it is evident that $\ln Z'(\underline{n}, \mathbf{s}^0) \leq K \ln 2$. On the other hand restricting the sum in the partition function to $\underline{x} = \underline{1}$ we have

$$\mathbb{E}_{\underline{N}}[\ln Z'(\underline{n}, \mathbf{s}^0)] \geq -\frac{1}{2\sigma^2} \mathbb{E}_{\underline{N}}[\sigma^2 \|\underline{n}\|^2] = -\frac{N}{2}$$

Therefore we have

$$\mathbb{E}_{\mathbf{S}}[f(\mathbf{S})^2] \leq \max(\frac{1}{4\beta^2}, (\ln 2)^2) = C^2$$

Let us now compute the Lipschitz constant.

Lemma 3: $K^{-1}\mathbb{E}_{\underline{N}}[2^{-K} \sum_{\underline{x}^0} \ln Z'(\underline{n}, \mathbf{s}^0)]$ is Lipschitz on G , with constant

$$L_N = \sigma^{-2} \sqrt{\beta} K^{-1} (\sqrt{B} + \sqrt{N} \sigma)$$

Proof: The exponent of the partition function is²

$$H(\underline{n}, \mathbf{s}^0, \underline{x}) = \frac{1}{2\sigma^2} \|N^{-1/2} \mathbf{s}^0(\underline{x} - \underline{1}) - \sigma \underline{n}\|^2 \quad (8)$$

In the appendix we show that

$$|H(\underline{n}, \mathbf{s}^0, \underline{x}) - H(\underline{n}, \mathbf{t}^0, \underline{x})| \leq \sigma^{-2} 2\sqrt{\beta}(\sqrt{B} + \|\underline{n}\|) \|\mathbf{s} - \mathbf{t}\| \quad (9)$$

Using this inequality together with

$$H(\underline{n}, \mathbf{s}^0, \underline{x}) \geq H(\underline{n}, \mathbf{t}^0, \underline{x}) - |H(\underline{n}, \mathbf{s}^0, \underline{x}) - H(\underline{n}, \mathbf{t}^0, \underline{x})|$$

we have for $(\mathbf{s}, \mathbf{t}) \in G \times G$

$$\begin{aligned} & \ln \frac{\sum_{\underline{x}} \exp(-H(\underline{n}, \mathbf{s}^0, \underline{x}))}{\sum_{\underline{x}} \exp(-H(\underline{n}, \mathbf{t}^0, \underline{x}))} \\ & \leq \ln \frac{\sum_{\underline{x}} \exp(|H(\underline{n}, \mathbf{s}^0, \underline{x}) - H(\underline{n}, \mathbf{t}^0, \underline{x})| - H(\underline{n}, \mathbf{t}^0, \underline{x}))}{\sum_{\underline{x}} \exp(-H(\underline{n}, \mathbf{t}^0, \underline{x}))} \\ & \leq \sigma^{-2} 2\sqrt{\beta}(\sqrt{B} + \|\underline{n}\|) \|\mathbf{s} - \mathbf{t}\| \end{aligned}$$

Therefore taking the expectation over the noise, we get

$$\begin{aligned} & \left| \sum_{\underline{x}^0} 2^{-K} \mathbb{E}_{\underline{N}}[\ln Z'(\underline{n}, \mathbf{s}^0)] - \sum_{\underline{x}^0} 2^{-K} \mathbb{E}_{\underline{N}}[\ln Z'(\underline{n}, \mathbf{t}^0)] \right| \\ & \leq \sigma^{-2} 2\sqrt{\beta}(\sqrt{B} + \sigma \mathbb{E}[\|\underline{n}\|]) \|\mathbf{s} - \mathbf{t}\| \\ & \leq \sigma^{-2} 2\sqrt{\beta}(\sqrt{B} + \sigma \mathbb{E}[\|\underline{n}\|^2]^{1/2}) \|\mathbf{s} - \mathbf{t}\| \end{aligned}$$

which yields the Lipschitz constant of the lemma. \blacksquare

Finally (6) follows from lemmas 1, 2 and 3 with the choice $B = 32\beta(2K + N)$. We obtain $\alpha_1 = 1/(8KL_N^2) \geq \sigma^4/(16\beta(64\beta + 32 + \sigma^2))$.

B. Proof of (7)

This case is more cumbersome but the ideas are the same. We choose the set G as

$$G = \left\{ \mathbf{s}, \underline{n} \mid \max_i |n_i| \leq \sqrt{A} \text{ and for all } \underline{x}, \|\mathbf{s}^0(\underline{x} - \underline{1})\|^2 \leq BN \right\}$$

where as before A and B will be chosen appropriately later on. For Gaussian noise $\mathbb{P}[|n_i| \geq \sqrt{A}] \leq 4e^{-\frac{A}{4}}$ therefore from the union bound $\mathbb{P}(\max_i |n_i| \geq \sqrt{A}) \leq 4Ne^{-\frac{A}{4}}$. Using lemma 2 we obtain an estimate for the measure of G^c ,

$$\mathbb{P}[G^c] \leq 4Ne^{-\frac{A}{4}} + 2^{K+\frac{N}{2}} e^{-\frac{B}{16\beta}}$$

The goal is to apply lemma 1 to $f(\underline{n}, \mathbf{s}) = \ln Z'(\underline{n}, \mathbf{s}^0)$ defined on $\mathbb{R}^K \times \mathbb{R}^{NK}$.

²a Hamiltonian

Clearly $(0, 0) \in G$, $f(0, 0) = \ln 2$ and by the same argument as before we have $\mathbb{E}[f(\underline{n}, \mathbf{s})^2] \leq \max(\frac{1}{4\beta^2}, (\ln 2)^2) = C^2$. It remains to compute the Lipschitz constant.

Lemma 4: The free energy $K^{-1} \ln Z'(\underline{n}, \mathbf{s}^0)$ is Lipschitz on G with constant

$$L_N = \sigma^{-2} (2\sqrt{\beta} + \sigma) K^{-1} (\sigma\sqrt{NA} + \sqrt{B})$$

Proof: For the same Hamiltonian (8) we show in the appendix,

$$\begin{aligned} & |H(\underline{n}, \mathbf{s}^0, \underline{x}) - H(\underline{n}, \mathbf{t}^0, \underline{x})| \\ & \leq \sigma^{-2} 2(2\sqrt{\beta} + \sigma)(\sigma\sqrt{NA} + \sqrt{B}) \|(\underline{n}, \mathbf{s}) - (\underline{n}, \mathbf{t})\| \quad (10) \end{aligned}$$

Then proceeding in the same way as in the proof of lemma 3 we get

$$\begin{aligned} & |\ln Z'(\underline{n}, \mathbf{s}^0) - \ln Z'(\underline{n}, \mathbf{t}^0)| \\ & \leq \sigma^{-2} (2\sqrt{\beta} + \sigma)(\sigma\sqrt{NA} + \sqrt{B}) \|(\underline{n}, \mathbf{s}) - (\underline{n}, \mathbf{t})\| \end{aligned}$$

We can now conclude the proof of (7) by collecting the previous results and choosing $A = \sqrt{N}/\sigma^2$ and $B = 32\beta(K + N)$. This gives $\alpha_2 = 1/(8\sqrt{N}L_N^2) \geq \sigma^4\beta^2/(32(2\sqrt{\beta} + \sigma)^2)$. \blacksquare

IV. CONCLUDING REMARKS

In this contribution we have shown that the capacity of the CDMA system with Gaussian signature sequences and binary inputs satisfies a concentration property. We expect that the powerful probabilistic tools used here have applications for other similar situations in communication systems. The main open problem is the analogous property for the bit error rate.

Essentially the same technique applies to a class of other signature distributions satisfying a log-Sobolev inequality but unfortunately we have not yet been able to cover the important case of binary signature sequences. Interestingly if we change the sign of the exponent in (1) and (2), the problem becomes very much related to the Hopfield model for neural networks and in that case we can prove concentration results even for binary \mathbf{S} because the corresponding distributions have convex level sets [6].

After submission of this paper we have been able to prove the self averaging for the case of binary signature sequences, namely that $\mathbb{P}[|C(\mathbf{s}) - \mathbb{E}_{\mathbf{S}}[C(\mathbf{s})]| \geq t] = O(\frac{1}{Nt^2})$ and similarly for the free energy. The proofs will be presented elsewhere.

V. APPENDIX

A. Relation between capacity and free energy

Replacing (1) in the conditional entropy

$$\begin{aligned} H(\underline{X}|\underline{Y}, \mathbf{s}) &= -\mathbb{E}_{\underline{Y}} \left[\sum_{\underline{x}} p(\underline{x}|\underline{y}, \mathbf{s}) \ln p(\underline{x}|\underline{y}, \mathbf{s}) \right] \\ &= \mathbb{E}_{\underline{Y}} \left[\sum_{\underline{x}} p(\underline{x}|\underline{y}, \mathbf{s}) \ln Z(\underline{y}, \mathbf{s}) \right] \\ &\quad + \mathbb{E}_{\underline{Y}} \left[\sum_{\underline{x}} p(\underline{x}|\underline{y}, \mathbf{s}) \frac{1}{2\sigma^2} \|N^{-\frac{1}{2}} \mathbf{s}\underline{x} - \underline{y}\|^2 \right] \end{aligned}$$

The last term on the right hand side can be computed exactly. Indeed,

$$\begin{aligned}
& \mathbb{E}_{\underline{Y}} \left[\sum_{\underline{x}} p(\underline{x}|\underline{y}, \mathbf{s}) \frac{1}{2\sigma^2} \|N^{-\frac{1}{2}} \mathbf{s}\underline{x} - \underline{y}\|^2 \right] \\
&= \frac{1}{2K} \int d\underline{y} \frac{Z(\underline{y}, \mathbf{s})}{(\sqrt{2\pi\sigma^2})^N} \sum_{\underline{x}} p(\underline{x}|\underline{y}, \mathbf{s}) \frac{1}{2\sigma^2} \|N^{-\frac{1}{2}} \mathbf{s}\underline{x} - \underline{y}\|^2 \\
&= \frac{1}{2K} \sum_{\underline{x}} \int d\underline{y} \frac{1}{(\sqrt{2\pi\sigma^2})^N} e^{-\frac{1}{2\sigma^2} \|N^{-\frac{1}{2}} \mathbf{s}\underline{x} - \underline{y}\|^2} \\
&\quad \times \frac{1}{2\sigma^2} \|N^{-\frac{1}{2}} \mathbf{s}\underline{x} - \underline{y}\|^2 \\
&= N \frac{1}{2} = K \frac{1}{2\beta}
\end{aligned}$$

Therefore the relation between Shannon's conditional entropy and the free energy is

$$H(\underline{X} | \underline{Y}, \mathbf{s}) = \mathbb{E}_{\underline{Y}}[\ln Z(\underline{y}, \mathbf{s})] + \frac{K}{2\beta}$$

This is equivalent to the announced relation (3).

B. Proof of (10)

Let $\underline{n}, \underline{m}$ be two noise realizations and \mathbf{s}, \mathbf{t} two spreading sequences all belonging to the appropriate set G . Let $\underline{y} = \underline{x} - \underline{1}$. First we expand the Euclidean norms

$$\begin{aligned}
& \|N^{-\frac{1}{2}} \mathbf{s}^0 \underline{y} - \sigma \underline{n}\|^2 - \|N^{-\frac{1}{2}} \mathbf{t}^0 \underline{y} - \sigma \underline{m}\|^2 \\
&= \sigma^2 \|\underline{n}\|^2 - \sigma^2 \|\underline{m}\|^2 + N^{-1} (\|\mathbf{s}^0 \underline{y}\|^2 - \|\mathbf{t}^0 \underline{y}\|^2) \\
&\quad + \sigma 2N^{-\frac{1}{2}} (\underline{n}^t \cdot \mathbf{s}^0 \underline{y} - \underline{m}^t \cdot \mathbf{t}^0 \underline{y}) \\
&= \sigma^2 (\underline{n} - \underline{m})^t \cdot (\underline{n} + \underline{m}) + N^{-1} (\mathbf{s}^0 \underline{y} - \mathbf{t}^0 \underline{y})^t \cdot (\mathbf{s}^0 \underline{y} + \mathbf{t}^0 \underline{y}) \\
&\quad + 2\sigma N^{-\frac{1}{2}} (\underline{n} - \underline{m})^t \cdot \mathbf{s}^0 \underline{y} + 2\sigma N^{-\frac{1}{2}} \underline{m}^t \cdot (\mathbf{s}^0 \underline{y} - \mathbf{t}^0 \underline{y})
\end{aligned}$$

We estimate each of the four terms on the right hand side of the last equality. By Cauchy-Schwartz the first term is bounded by

$$\begin{aligned}
\|\underline{n} - \underline{m}\| \|\underline{n} + \underline{m}\| &\leq \sqrt{N} \max_i (|n_i| + |m_i|) \|\underline{n} - \underline{m}\| \\
&\leq 2\sqrt{NA} \|\underline{n} - \underline{m}\|
\end{aligned}$$

Using Cauchy-Schwartz and $\|(\mathbf{s}^0 - \mathbf{t}^0) \underline{y}\| \leq \|\mathbf{s}^0 - \mathbf{t}^0\| \|\underline{y}\|$ where $\|\mathbf{s}^0 - \mathbf{t}^0\| = \|\mathbf{s} - \mathbf{t}\|$ is the (Hilbert-Schmidt) norm,

$$\|\mathbf{s} - \mathbf{t}\| = \left(\sum_{i=1}^N \sum_{l=1}^K (s_{il} - t_{il})^2 \right)^{1/2}$$

we obtain for the second term the estimate

$$\begin{aligned}
N^{-1} \|\mathbf{s} - \mathbf{t}\| \|\underline{y}\| (\|\mathbf{s}^0 \underline{y}\| + \|\mathbf{t}^0 \underline{y}\|) &\leq N^{-1} \|\mathbf{s} - \mathbf{t}\| 2\sqrt{K} 2\sqrt{BN} \\
&= 4\sqrt{\beta B} \|\mathbf{s} - \mathbf{t}\|
\end{aligned}$$

Similarly the third term is bounded by,

$$\begin{aligned}
2N^{-\frac{1}{2}} \|\underline{n} - \underline{m}\| \|\mathbf{s}^0 \underline{y}\| &\leq 2N^{-\frac{1}{2}} \|\underline{n} - \underline{m}\| \sqrt{BN} \\
&= 2\sqrt{B} \|\underline{n} - \underline{m}\|
\end{aligned}$$

and the fourth one by

$$\begin{aligned}
2N^{-\frac{1}{2}} \|\underline{m}\| \|\mathbf{s} - \mathbf{t}\| \|\underline{y}\| &\leq 2N^{-\frac{1}{2}} \sqrt{NA} \|\mathbf{s} - \mathbf{t}\| 2\sqrt{K} \\
&= 4\sqrt{\beta NA} \|\mathbf{s} - \mathbf{t}\|
\end{aligned}$$

Collecting all four estimates we obtain

$$\begin{aligned}
& \|N^{-\frac{1}{2}} \mathbf{s}^0 (\underline{x} - \underline{1}) - \sigma \underline{n}\|^2 - \|N^{-\frac{1}{2}} \mathbf{t}^0 (\underline{x} - \underline{1}) - \sigma \underline{m}\|^2 \\
&\leq 2\sigma (\sigma \sqrt{NA} + \sqrt{B}) \|\underline{n} - \underline{m}\| + 4\sqrt{\beta} (\sigma \sqrt{NA} + \sqrt{B}) \|\mathbf{s} - \mathbf{t}\| \\
&\leq 2(2\sqrt{\beta} + \sigma) (\sigma \sqrt{NA} + \sqrt{B}) \|(\underline{n}, \mathbf{s}) - (\underline{m}, \mathbf{t})\|
\end{aligned}$$

where the last norm is the Euclidean norm in $\mathbb{R}^N \times \mathbb{R}^{NK}$.

C. Proof of (9)

Let \mathbf{s} and \mathbf{t} be two spreading sequences both belonging to the appropriate G . Let $\underline{y} = \underline{x} - \underline{1}$. Following similar steps as in the previous paragraph with $\underline{n} = \underline{m}$ the result can be read off

$$\begin{aligned}
& \|N^{-\frac{1}{2}} \mathbf{s}^0 \underline{y} - \sigma \underline{n}\|^2 - \|N^{-\frac{1}{2}} \mathbf{t}^0 \underline{y} - \sigma \underline{n}\|^2 \\
&\leq 4\sqrt{\beta} (\sqrt{B} + \sigma \|\underline{n}\|) \|\mathbf{s} - \mathbf{t}\|
\end{aligned}$$

ACKNOWLEDGEMENT

We would like to thank Olivier Lévêque and Rüdiger Urbanke for discussions. The work presented in this paper was partially supported by the National Competence Center in Research on Mobile Information and Communication Systems (NCCR-MICS), a center supported by the Swiss National Science Foundation under grant number 5005-67322.

REFERENCES

- [1] T. Tanaka, "A statistical mechanics approach to large-system analysis of cdma multiuser detectors," *IEEE Transactions on Information theory* vol. 48, Issue 11, (2002) pp. 2888–2910.
- [2] A. Montanari, D. Tse, "Analysis of Belief Propagation for Non-Linear Problems: The Example of CDMA (or: How to Prove Tanaka's Formula)," *IEEE Information Theory Workshop Punta del Este, Uruguay*, March 13–17, 2006
- [3] D. Tse, S. V. Hanly, "Linear Multiuser Receivers: Effective Interference, Effective Bandwidth and User Capacity", *IEEE Transactions on Information theory* vol. 45, (1999) pp. 641–657.
- [4] S. Verdú, S. Shamai, "Spectral Efficiency of CDMA with Random Spreading", *IEEE Transactions on Information theory* vol. 45, (1999) pp. 622–640.
- [5] M. Talagrand, "A New look at independence", *The Annals of Probability* vol. 24, Issue. 1, (1996) pp. 1–34.
- [6] M. Talagrand, "Spin Glasses: A Challenge for Mathematicians" Springer-Verlag (2003)
- [7] J. Heinonen, "Lectures on Lipschitz analysis", Report University of Jyväskylä. Department of mathematics and statistics (2005)
- [8] F. Guerra, F.L. Toninelli, "The infinite volume limit in generalized mean field disordered models", *Markov Proc. Rel. Fields* vol. 9 No. 2, pp. 195–207 (2003).
- [9] M. Mézard, G. Parisi, M. Virasoro, *Spin glass theory and beyond*, World Scientific (1987)
- [10] L.A. Pastur, M.V. Shcherbina, "Absence of Self-Averaging of the Order Parameter in the Sherrington-Kirkpatrick Model", *Journal of Statistical Physics*, vol. 62, (1991)