Vector Precoding in High Dimensions: A Replica Analysis

Ralf R. Müller
Dept Elec. & Telecommun.
NTNU
Trondheim, Norway
ralf@iet.ntnu.no

Dongning Guo
Dept EECS
Northwestern University
Evanston, IL, U.S.A.
dGuo@northwestern.edu

Aris L. Moustakas
Physics Dept
National & Capodistrian Univ. of Athens
Athens, Greece
arislm@phys.uoa.gr

***Abstract*— We apply the replica method to analyze vector pre-coding, a method to reduce transmit power in antenna array communications, in the limit of an infinite number of dimensions of the signal vector. The analysis applies to a very general class of channel matrices. The statistics of the channel matrix enter the transmitted energy per symbol via its R-transform.**

We specialize our result to inversion of an i.i.d. channel and two cases of signal point optimization i) 2-point lattice pre-coding and ii) compact relaxation. In the two cases the replica symmetric transmitted energy is found to be 4.3 dB and 9.6 dB above the orthogonal case for a square channel matrix, respectively.

I. Introduction

Vector precoding aims to minimize the transmitted power that is associated with the transmission of a certain data vector of length K. For that purpose, the original symbol alphabet $\mathcal{S}$ is relaxed into the alphabet $\mathcal{B} \supset \mathcal{S}$. The representation in the relaxed alphabet is redundant. That means that several symbols in the relaxed alphabet represent the same information. Due to the redundant representation, we can choose that representation of our information which requires the least power to be transmitted. This way of saving transmit power is called vector pre-coding.[1]

That means, for any $s \in \mathcal{S}$, there is a set $\mathcal{B}_s \subset \mathcal{B}$ such that all elements of $\mathcal{B}_s$ represent the same information as s. Often, $s \in \mathcal{B}_s$ though this is not necessary, in general. In order to avoid ambiguities, we should have

$$\mathcal{B}_s \cap \mathcal{B}_{s'} = \emptyset \qquad \forall s \neq s' \wedge s, s' \in \mathcal{B}. \tag{1}$$

In addition, one would like to design the sets $\mathcal{B}_s$ for all $s \in \mathcal{S}$ such that the distance properties between the presented information are preserved. This is easily achieved by letting $\mathcal{B}$ be a lattice expansion of $\mathcal{S}$ and $\mathcal{B}_s$ be a judicious choice of disjoint sub-lattices. However, we are not concerned with these design issues here. We aim to analyze the power saving achieved by a particular choice of the sets $\mathcal{B}_s$. This goal is achieved using the replica method invented in statistical physics.

[1]This research was supported by the Research Council of Norway, the National Science Foundation, DARPA, and the European Commission under grants 171133/V30, CCF-0644344, W911NF-07-1-0028, and MIRG-CT-2005-030833, resp.

The replica method was introduced into multiuser communications by the landmark paper of Tanaka [1] for the purpose of studying the performance of the maximum a-posteriori detector. Additionally, the replica method has also been successfully used for the design and analysis of error correction codes. Vector pre-coding has been discussed in the non-asymptotic regime by several authors, e.g. [2]. For more information on vector pre-coding, see also [3].

II. Problem Statement

Let $\boldsymbol{s} = [s_1, s_2, \ldots, s_K]^{\mathrm{T}}$ denote the information to be encoded. Then, the pre-coding problem can be written as the minimization of the quadratic form

$$\min_{\boldsymbol{x} \in \mathcal{X}} \boldsymbol{x}^\dagger \boldsymbol{R} \boldsymbol{x} \tag{2}$$

over the discrete set

$$\mathcal{X} = \mathcal{B}_{s_1} \times \mathcal{B}_{s_2} \times \cdots \times \mathcal{B}_{s_K} \tag{3}$$

for a given $K \times K$ matrix $\boldsymbol{R}$.

In order to allow for analytical tractability, we need a few assumptions:

Assumption 1 (self-averaging property): We have

$$\lim_{K\to\infty} \Pr\left(\frac{1}{K} \left| \min_{\boldsymbol{x}\in\mathcal{X}} \boldsymbol{x}^\dagger \boldsymbol{R}\boldsymbol{x} - \mathop{\mathrm{E}}_{\boldsymbol{R}} \min_{\boldsymbol{x}\in\mathcal{X}} \boldsymbol{x}^\dagger \boldsymbol{R}\boldsymbol{x} \right| > \epsilon \right) = 0 \tag{4}$$

for all $\epsilon > 0$, i.e. convergence in probability.

Assumption 2 (replica continuity): For all $\beta > 0$, the continuation of the function

$$f(n) = \prod_{a=1}^{n} \sum_{\boldsymbol{x}_a \in \mathcal{X}} \mathrm{e}^{-\beta \boldsymbol{x}_a^\dagger \boldsymbol{R} \boldsymbol{x}_a} \tag{5}$$

onto the positive real line is equal to

$$\left(\sum_{\boldsymbol{x}\in\mathcal{X}} \mathrm{e}^{-\beta \boldsymbol{x}^\dagger \boldsymbol{R}\boldsymbol{x}} \right)^n$$

in the vicinity of $n = 0$.

Assumption 3 (unitary invariance): The random matrix $\boldsymbol{R}$, can be decomposed into

$$\boldsymbol{R} = \boldsymbol{O}\boldsymbol{D}\boldsymbol{O}^\dagger \tag{6}$$

such that the matrices $\boldsymbol{D}$ and $\boldsymbol{O}$ are diagonal and Haar distributed, respectively. Moreover, as $K \to \infty$, the asymptotic eigenvalue distribution of $\boldsymbol{R}$ converges to a non-random distribution function which can be uniquely characterized by its R-transform[2] $R(w)$.

III. General Result

With Assumption 1, $\mathrm{E} \log x \overset{0 \leftarrow n}{\longleftarrow} \partial/\partial n \log \mathrm{E}\, x^n$, postulating the exchange of limits and derivation, and Assumption 2, we find for the average transmitted energy per symbol in the large system limit

$$E_\mathrm{s} = \lim_{K\to\infty} \frac{1}{K} \mathop{\mathrm{E}}_{\boldsymbol{R}} \min_{\boldsymbol{x}\in\mathcal{X}} \boldsymbol{x}^\dagger \boldsymbol{R} \boldsymbol{x} \tag{7}$$

$$= \lim_{K\to\infty} \lim_{\beta\to\infty} \frac{-1}{\beta K} \mathop{\mathrm{E}}_{\boldsymbol{R}} \log \sum_{\boldsymbol{x}\in\mathcal{X}} \mathrm{e}^{-\beta \boldsymbol{x}^\dagger \boldsymbol{R} \boldsymbol{x}} \tag{8}$$

$$= \lim_{\beta\to\infty} \frac{-1}{\beta} \lim_{n\to 0} \frac{\partial}{\partial n} \underbrace{\lim_{K\to\infty} \frac{1}{K} \log \mathop{\mathrm{E}}_{\boldsymbol{R}} \prod_{a=1}^{n} \sum_{\boldsymbol{x}_a\in\mathcal{X}} \mathrm{e}^{-\beta \boldsymbol{x}_a^\dagger \boldsymbol{R} \boldsymbol{x}_a}}_{\triangleq\, \Xi_n} . \tag{9}$$

Since traces commute, we have

$$\Xi_n = \lim_{K\to\infty} \frac{1}{K} \log \mathop{\mathrm{E}}_{\boldsymbol{R}} \sum_{\boldsymbol{x}_1\in\mathcal{X}} \cdots \sum_{\boldsymbol{x}_n\in\mathcal{X}} \mathrm{e}^{\mathrm{tr}\left(-\beta \boldsymbol{R} \sum\limits_{a=1}^{n} \boldsymbol{x}_a \boldsymbol{x}_a^\dagger\right)} . \tag{10}$$

Using Assumption 3, we can integrate over the Haar distributed eigenvectors of $\boldsymbol{R}$. Let $R(w)$ denote the R-transform [4], [5], [6] of the asymptotic eigenvalue distribution of $\boldsymbol{R}$. Then, we have from [7], [8][3]

$$\Xi_n = \lim_{K\to\infty} \frac{1}{K} \log \tag{11}$$

$$\sum_{\boldsymbol{x}_1\in\mathcal{X}} \cdots \sum_{\boldsymbol{x}_n\in\mathcal{X}} \exp\left[-K \sum_{a=1}^{n} \int_0^1 \lambda_a R(-\lambda_a w) dw \right]$$

with λ_i denoting the n positive eigenvalues of

$$\beta \sum_{a=1}^{n} \boldsymbol{x}_a \boldsymbol{x}_a^\dagger . \tag{12}$$

The eigenvalues λ_i are completely determined by the inner products

$$K Q_{ab} = \boldsymbol{x}_a^\dagger \boldsymbol{x}_b . \tag{13}$$

In order to perform the summation in (12), the Kn-dimensional space spanned by the replicas is split into subshells $\left\{ \boldsymbol{x}_1, \ldots, \boldsymbol{x}_n \,\middle|\, \boldsymbol{x}_a^\dagger \boldsymbol{x}_b = K Q_{ab} \right\}$ where the inner product of two different vectors $\boldsymbol{x}_a$ and $\boldsymbol{x}_b$ is constant.[4] With this splitting of the space, we find[5]

$$\Xi_n = \lim_{K\to\infty} \frac{1}{K} \log \int\limits_{\mathbb{R}^{n(n+1)/2}} \mathrm{e}^{K\mathcal{I}\{Q\}} \mathrm{e}^{-K\mathcal{G}\{Q\}} \prod_{a\le b} \mathrm{d}Q_{ab}, \tag{14}$$

where

$$\mathrm{e}^{K\mathcal{I}\{Q\}} = \int \left[\prod_{a\le b} \delta\left(\boldsymbol{x}_a^\dagger \boldsymbol{x}_b - K Q_{ab} \right) \right] \prod_{a=1}^{n} \mathrm{dP}_{\boldsymbol{x}}(\boldsymbol{x}_a) \tag{15}$$

denotes the probability weight of the subshell and

$$\mathcal{G}\{Q\} = \sum_{a=1}^{n} \int_0^{\lambda_a\{Q\}} R(-w)\,\mathrm{d}w \tag{16}$$

This procedure is a change of integration variables in multiple dimensions where the integration of an exponential function over the replicas has been replaced by integration over the variables $\{Q\}$. In the following the two exponential terms in (14) are evaluated separately.

First, we turn to the evaluation of the measure $\mathrm{e}^{K\mathcal{I}\{Q\}}$. For some $t \in \mathbb{R}$ and $\mathcal{J} = (t - \mathrm{j}\infty; t + \mathrm{j}\infty)$, the measure $\mathrm{e}^{K\mathcal{I}\{Q\}}$ can be expressed with the Fourier expansion of the Dirac measure as

$$\mathrm{e}^{K\mathcal{I}\{Q\}} = \int\limits_{\mathcal{J}^{n(n+1)/2}} \mathrm{e}^{\log \prod\limits_{k=1}^{K} M_k\{\tilde{Q}\} - K \sum\limits_{a\le b} \tilde{Q}_{ab} Q_{ab}} \prod_{a\le b} \frac{\mathrm{d}\tilde{Q}_{ab}}{2\pi\mathrm{j}} \tag{17}$$

with

$$M_k\left\{\tilde{Q}\right\} = \int \exp\left(\sum_{a\le b} \tilde{Q}_{ab} x_a x_b \right) \prod_{a=1}^{n} \mathrm{dP}_{x_k}(x_a). \tag{18}$$

In the limit of $K \to \infty$ one of the exponential terms in (14) will dominate over all others. Thus, only that value of the correlation Q_{ab} which maximizes the exponent is relevant for calculation of the integral.

At this point, we assume replica symmetry. This means, that in order to find the maximum of the objective function, we consider only a subset of the potential possibilities that the variables Q_{ab} could take. Here, we restrict them to the following two different possibilities $Q_{ab} = q, \forall a \neq b$ and $Q_{aa} = q + b/\beta, \forall a$. One case distinction has been made to distinguish correlations Q_{ab} which correspond to correlations between different and identical replica indices. We apply the same idea to the correlation variables in the Fourier domain and set with a modest amount of forsight $\tilde{Q}_{ab} = 2\beta^2 f, \forall a \neq b$ and $\tilde{Q}_{aa} = \beta^2 f - \beta e, \forall a$.

At this point the crucial benefit of the replica method becomes obvious. Assuming replica continuity, we have managed to reduce the evaluation of a continuous function to sampling it at integer points. Assuming replica symmetry we have reduced the task of evaluating infinitely many integer points to calculating four different correlations (two in the original and two in the Fourier domain).

The assumption of replica symmetry leads to

$$\sum_{a\le b} \tilde{Q}_{ab} Q_{ab} = n(n-1)\beta^2 f q + n(\beta f - e)(\beta q + b) \tag{19}$$

[2]See [4], [5], [6] for the definition of the R-transform.

[3]In their original work [7], Marinari et al. do not formulate their result in terms of the R-transform from free probability theory, but in terms of what they call the *generating function*.

[4]The notation $f\{Q\}$ expresses dependency of the function $f(\cdot)$ on $Q_{ab}, 1 \le a \le b \le n$.

[5]The notation $\prod_{a\le b}$ is used as shortcut for $\prod_{a=1}^{n} \prod_{b=a}^{n}$.

Proposition 1 (main result): Given Assumptions 1 to 3 and replica symmetry, we have

$$E_{\mathrm{s}} = \lim_{K\to\infty} \frac{1}{K} \min_{\boldsymbol{x}\in\mathcal{X}} \boldsymbol{x}^\dagger \boldsymbol{R}\boldsymbol{x} = \lim_{\beta\to\infty} R(-b)\left(q+\frac{b}{\beta}\right) - qbR'(-b)$$

in probability and the the parameters b and q being determined by the following system of fixed-point equations

$$q = \iint \left[\frac{\int x\mathrm{e}^{\beta\sqrt{2qR'(-b)}zx-\beta R(-b)x^2}\mathrm{dP}_{x_k}(x)}{\int \mathrm{e}^{\beta\sqrt{2qR'(-b)}zx-\beta R(-b)x^2}\mathrm{dP}_{x_k}(x)}\right]^2 \mathrm{D}z\,\mathrm{dP}_s(x_k)$$

$$b = \beta\iint \frac{\int x^2\mathrm{e}^{\beta\sqrt{2qR'(-b)}zx-\beta R(-b)x^2}\mathrm{dP}_{x_k}(x)}{\int \mathrm{e}^{\beta\sqrt{2qR'(-b)}zx-\beta R(-b)x^2}\mathrm{dP}_{x_k}(x)} \mathrm{D}z\,\mathrm{dP}_s(x_k) - \beta q$$

with $\mathrm{D}z$ being the Gaussian measure defined in (25).

---

and

$$M_k = \int \mathrm{e}^{\beta\sum\limits_{a=1}^{n}(\beta f-e)x_a^2+\sum\limits_{b=a+1}^{n}\beta f x_a x_b} \prod_{a=1}^{n} \mathrm{dP}_{x_k}(x_a) \tag{20}$$

Note that the prior distribution enters the transmitted energy only via (20). We will focus on this later on after having finished with the other terms.

For the evaluation of $\mathcal{G}\{Q\}$ in (14), we can use the replica symmetry to explicitly calculate the eigenvalues λ_i. Considerations of linear algebra lead to the conclusion that the eigenvalues b and $b+\beta nq$ occur with multiplicities $n-1$ and 1, respectively. Thus we get

$$\mathcal{G}(q,b) = (n-1)\int_0^b R(-w)\,\mathrm{d}w + \int_0^{b+\beta nq} R(-w)\,\mathrm{d}w. \tag{21}$$

Since the integral in (14) is dominated by the maximum argument of the exponential function, the derivatives of

$$\mathcal{G}\{Q\} + \sum_{a\le b} \tilde{Q}_{ab} Q_{ab} \tag{22}$$

with respect to q and b must vanish as $K\to\infty$. Taking derivatives after plugging (19) and (21) into (22), gives

$$R(-b-\beta nq) + (n-1)\beta f + \beta f - e = 0$$
$$(n-1)R(-b) + R(-b-\beta nq) + n(\beta f - e) = 0$$

solving for e and f gives

$$e = R(-b) \tag{23}$$
$$f \overset{n\to 0}{\longrightarrow} qR'(-b). \tag{24}$$

Consider now the integration over the prior distribution in the moment-generating function. Consider (20) giving the only term that involves the prior distribution and apply the Hubbard-Stratonovich transform

$$\mathrm{e}^{\frac{x}{2}} = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} \mathrm{e}^{\pm\sqrt{x}z-\frac{z^2}{2}}\mathrm{d}z =: \int \mathrm{e}^{\pm\sqrt{x}z}\mathrm{D}z. \tag{25}$$

Then, we find

$$M_k = \int\left(\int \mathrm{e}^{\beta\sqrt{2f}zx-\beta ex^2}\mathrm{dP}_{x_k}(x)\right)^n \mathrm{D}z \tag{26}$$

Moreover, for $K\to\infty$, we have by the law of large numbers

$$\log M = \frac{1}{K}\log\prod_{k=1}^{K} M_k \tag{27}$$
$$= \int \log\int\left(\int \frac{\mathrm{e}^{\beta\sqrt{2f}zx}}{\mathrm{e}^{\beta ex^2}}\,\mathrm{dP}_{x_k}(x)\right)^n \mathrm{D}z\mathrm{dP}_s(x_k)$$

In the large system limit, the integral in (17) is dominated by that value of the integration variable which maximizes the argument of the exponential function. Thus, partial derivatives of

$$\log M - n(n-1)\beta^2 fq - n(\beta f-e)(b+\beta q) \tag{28}$$

with respect to f and e must vanish as $K\to\infty$. An explicit calculation of the two derivatives gives the expressions for the macroscopic parameters q and b as $n\to 0$ shown on top of the page.

Returning to our initial goal, the evaluation of the average transmitted energy per symbol, and collecting our previous results, we find Proposition 1 shown on top of the page. Note that in Proposition 1, the parameter b depends on β and may diverge as $\beta\to\infty$ for certain R-transforms. In those cases, b needs to be re-normalized to allow for meaningful solutions of the system of fixed-point equations. The current normalization was chosen, as it leads to well-defined solutions in the pre-coding examples to be discussed next.

IV. Channel Inversion

We now give an example for vector-precoding in context of antenna array communications. Consider the vector channel $\boldsymbol{y} = \boldsymbol{H}\boldsymbol{t}$.[6] In order to save complexity, we want to make symbol-by-symbol decisions on the K components of the observable $\boldsymbol{y}$. Thus, we choose the $N\times 1$ transmitted vector as

$$\boldsymbol{t} = \boldsymbol{H}^\dagger(\boldsymbol{H}\boldsymbol{H}^\dagger)^{-1}\boldsymbol{x}. \tag{29}$$

Thus, we find that $\boldsymbol{R} = (\boldsymbol{H}\boldsymbol{H}^\dagger)^{-1}$. In that case, we find after some algebra which is omitted due to space limitations that

$$R(w) = \frac{1-\alpha-\sqrt{(1-\alpha)^2-4\alpha w}}{2\alpha w} \tag{30}$$

[6] The red script corrects typos in the published paper.

with $\alpha = \frac{K}{N}$. Note that the inverse only exists for $\alpha \leq 1$.

A. One-Dimensional Integer Lattice

Consider now binary phase-shify keying, i.e.

$$\mathcal{S} = \{+1, -1\} \tag{31}$$

and a symmetric lattice extension, i.e.

$$\mathrm{p}_{+1}(x) = \mathrm{p}_{-1}(-x)\,. \tag{32}$$

Then, we find

$$q = \int \left[\frac{\int x \mathrm{e}^{\beta\sqrt{2qR'(-b)}zx - \beta R(-b)x^2} \mathrm{dP}_1(x)}{\int \mathrm{e}^{\beta\sqrt{2qR'(-b)}zx - \beta R(-b)x^2} \mathrm{dP}_1(x)} \right]^2 \mathrm{D}z \tag{33}$$

and

$$b = \beta \int \frac{\int x^2 \mathrm{e}^{\beta\sqrt{2qR'(-b)}zx - \beta R(-b)x^2} \mathrm{dP}_1(x)}{\int \mathrm{e}^{\beta\sqrt{2qR'(-b)}zx - \beta R(-b)x^2} \mathrm{dP}_1(x)} \mathrm{D}z - \beta q. \tag{34}$$

Moreover, let

$$\mathrm{p}_1(x) = \frac{1}{L} \sum_{i=1}^{L} \delta\left(x + (2i-1)(-1)^i\right). \tag{35}$$

This case describes Tomlinson-Harashima precoding [9], [10] with optimization over L different representations for each information bit. Note that $L = 1$ leaves only a single choice for each information symbol and corresponds to no pre-coding at all. As expected, this implies $b = 0, q = 1$ and $E_\mathrm{s} = R(0)$ which serves as an trivial upper bound on the transmitted energy with pre-coding. In the following, we restrict ourselves to the case of $L = 2$ without making explicit reference to it.

1) Square Channel Matrix: The R-transform simplifies to $R(w) = (-w)^{-\frac{1}{2}}$. This implies

$$b \overset{\beta\to\infty}{\longrightarrow} \frac{q^2}{16} \mathrm{W}^2\left(\frac{32}{\pi q^2}\right) \tag{36}$$

$$q \overset{\beta\to\infty}{\longrightarrow} 1 + 8\mathrm{Q}\left(\sqrt{\mathrm{W}\left(\frac{32}{\pi q^2}\right)}\right) \approx 2.5557 \tag{37}$$

and with replica symmetry

$$E_\mathrm{s} = \frac{q}{2\sqrt{b}} = \frac{2}{\mathrm{W}\left(\frac{32}{\pi q^2}\right)} \approx 2.6942 \tag{38}$$

where $\mathrm{W}(\cdot)$ is Lambert's W-function [11] and $\mathrm{Q}(x) = \int_x^\infty \mathrm{D}z$. Thus, pre-coding with just two additional symbols $\{+3; -3\}$ reduces the power penalty resulting from channel inversion from infinity to 4.3 dB.

2) Rectangular Channel Matrix: After some algebra, we find

$$E_\mathrm{s} = qR(-b) - qbR'(-b) \tag{39}$$

$$b = \frac{2}{\sqrt{\pi q R'(-b)}} \exp\left(-\frac{R^2(-b)}{qR'(-b)}\right) \tag{40}$$

$$q = 1 + 8\mathrm{Q}\left(R(-b)\sqrt{\frac{2}{qR'(-b)}}\right) \tag{41}$$

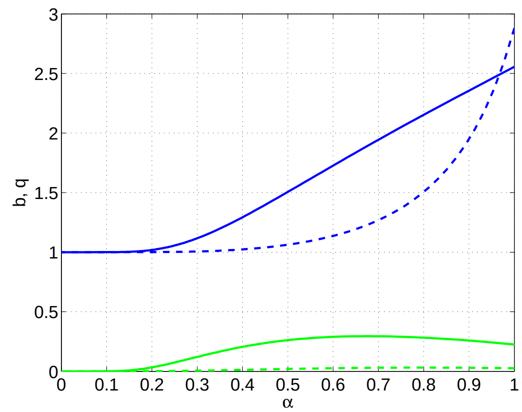

Fig. 1. The macroscopic parameters q (upper lines) and b (lower lines) versus the load α.

With the R-transform of the channel matrix and the new parameter

$$c = \sqrt{(1-\alpha)^2 + 4\alpha b}, \tag{42}$$

this gives

$$c \overset{\beta\to\infty}{\longrightarrow} \frac{4\sqrt{\alpha c}\,\mathrm{e}^{-\frac{c}{\alpha q}}}{\sqrt{\pi q}} + 1 - \alpha \tag{43}$$

$$q \overset{\beta\to\infty}{\longrightarrow} 1 + 8\mathrm{Q}\left(\sqrt{\frac{2c}{\alpha q}}\right) \tag{44}$$

and with replica symmetry

$$E_\mathrm{s} = \frac{q}{c}. \tag{45}$$

Thus, we find

$$E_\mathrm{s} = \frac{1 + 8\mathrm{Q}\left(\sqrt{\frac{2}{\alpha E_\mathrm{s}}}\right)}{\frac{4\sqrt{\alpha}\mathrm{e}^{-\frac{1}{\alpha E_\mathrm{s}}}}{\sqrt{\pi E_\mathrm{s}}} + 1 - \alpha} \tag{46}$$

which is a parameter-free fixed point equation for the average transmitted energy per symbol.

The solutions of these fixed-point equations are shown by the solid lines in Fig. 1. Clearly for small load, the parameter q tends to 1, as in that case, no gain due to pre-coding is possible and the symbol $+1$ saves power compared to -3. The minimum of the transmit power is shown by the solid line in Fig. 2. The lower dash-dotted line shows the lower bound on the power with precoding given by 1 which would be achieved (even without vector pre-coding) in the case of an orthogonal channel. The upper dash-dotted line shows the power required without precoding given by

$$R(0) = (1-\alpha)^{-1}. \tag{47}$$

We notice that the gain due to vector precoding becomes negligibly small for $\alpha < \frac{1}{3}$.

In addition to having found an analytical result for the transmitted energy per symbol, we can also use our result to reduce the complexity of a sphere decoder by

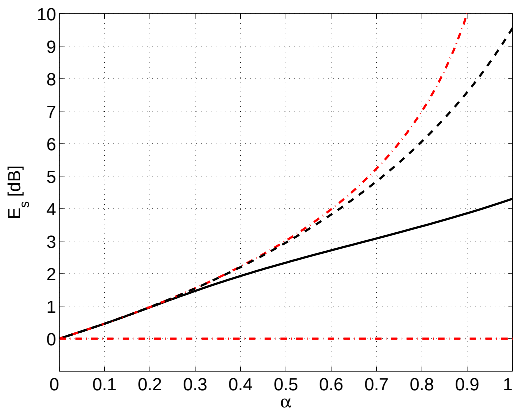

Fig. 2. The transmitted energy per symbol versus the load. The dash-dotted lines show the lower and upper bounds on the free energy given by 1 and $R(0)$, respectively.

searching for the optimum precoding vector only in the typical set. From $\lim_{K\to\infty}\frac{1}{K}\boldsymbol{x}^\dagger\boldsymbol{x}=q$, we find

$$\Pr(|x_k|=1)=\frac{9-q}{8}\geq 0.8 \tag{48}$$

which reduces the complexity of the exhaustive search for large K from 2^K to $2^{K\mathrm{h}\left(\frac{q-1}{8}\right)}$ where $\mathrm{h}(\cdot)$ is the binary entropy function.

B. Compact Relaxation

Consider now again a binary alphabet $\mathcal{S}=\{+1,-1\}$, but the compact relaxed alphabets

$$\mathcal{B}_{+1}=[+1;+\infty) \quad\wedge\quad \mathcal{B}_{-1}=(-\infty;-1]. \tag{49}$$

This choice of alphabet relaxation has the advantage of allowing for convex programming to find the optimum transmitted signal. Note that the minimum distance between signal points is preserved, as signal points may drift apart from each other, but may not get closer.

1) Square Channel Matrix: For the square channel matrix, we find for $\beta\to\infty$

$$q = 2+\sqrt{\frac{q}{2\pi\sqrt{b}}}\,\mathrm{e}^{-\frac{2\sqrt{b}}{q}}-4\sqrt{b}\approx 2.8796 \tag{50}$$

$$b = \frac{1}{4}\mathrm{Q}^2\left(2b^{\frac14}q^{-\frac12}\right)\approx 0.0254. \tag{51}$$

and with replica symmetry

$$E_\mathrm{s}=\frac{q}{2\sqrt{b}}\approx 9.0287 \tag{52}$$

Compact relaxation reduces the power loss due to channel inversion from infinity to about 9.6 dB. This is about 5.3 dB worse than 2-point lattice pre-coding. The latter, however, is NP-complete while compact relaxation is a polynomial-time algorithm. Moreover, compact relaxation slightly reduces the bit error rate, while lattice pre-coding slightly increases it.

2) Rectangular Channel Matrix: Introducing the new parameter c defined in (42), we find in the limit of $\beta\to\infty$

$$q = \frac{2}{\alpha}-\frac{2c}{\alpha}-\frac{q}{c}(1-\alpha)+\sqrt{\frac{\alpha q}{\pi c}}\,\mathrm{e}^{-\frac{c}{\alpha q}} \tag{53}$$

$$c = \alpha\mathrm{Q}\left(\sqrt{\frac{2c}{\alpha q}}\right)+1-\alpha \tag{54}$$

and with replica symmetry (45). The solutions of these fixed point equations are shown by the dashed lines in Figs. 1 and 2, resp. We notice that compact relaxation falls far behind lattice precoding and noticeably improves over linear precoding only for loads $\alpha>\frac12$.

V. Conclusions

We have shown that vector-precoding allows to reduce the power penalty incurred by inversion of the covariance matrix of the channel. Even with a polynomial-time pre-coding algorithm, an infinite power saving is possible. With lattice pre-coding—an NP-complete search algorithm—even greater power savings are possible. Provided that replica symmetry holds, we have found compact analytical expressions for the required energy per transmitted symbol in several special cases.

In future work, we will consider larger lattices than the 4-point integer lattice discussed in this work, complex-valued pre-coding, and regularized pre-coding matrices of the form $(\boldsymbol{H}\boldsymbol{H}^\dagger+\gamma\mathbf{I})^{-1}\boldsymbol{H}\boldsymbol{H}^\dagger(\boldsymbol{H}\boldsymbol{H}^\dagger+\gamma\mathbf{I})^{-1}$ where γ is a free design parameter to be optimized.

References

[1] T. Tanaka. A statistical mechanics approach to large-system analysis of CDMA multiuser detectors. *IEEE Trans. Inform. Theory*, 48(11):2888–2910, Nov. 2002.

[2] B. M. Hochwald, C. B. Peel, and A. L. Swindlehurst. A vector-perturbation technique for near-capacity multiantenna multiuser communication-Part II: Perturbation. *IEEE Trans. Commun.*, 53(3):537–544, Mar. 2005.

[3] R. F. H. Fischer. *Precoding and Signal Shaping for Digital Transmission*. John Wiley & Sons, 2002.

[4] D. V. Voiculescu, K. J. Dykema, and A. Nica. *Free Random Variables*. Am. Math. Soc., Providence, RI, 1992.

[5] R. R. Müller. Applications of large random matrices in communications engineering. In *Proc. Int'l Conf. Adv. Internet, Processing, Syst. & Interdisciplinary Res. (IPSI)*, Sveti Stefan, Montenegro, Oct. 2003.

[6] A. M. Tulino and S. Verdú. Random matrix theory and wireless communications. *Foundations & Trends in Commun. & Inform. Theory*, 1(1), Jun. 2004.

[7] E. Marinari, G. Parisi, and F. Ritort. Replica field theory for deterministic models (II): A non-random spin glass with glassy behavior. *J. Phys. A: Math. & Gen*, 27:7647–7668, 1994.

[8] C. Itzykson and J. B. Zuber. The planar approximation(II). *J. Math. Phys.*, 21(3):411–421, Mar. 1980.

[9] M. Tomlinson. New automatic equaliser employing modulo arithmetic. *IEE Elec. Lett.*, 7:138–139, Mar. 1971.

[10] H. Harashima and H. Miyakawa. Matched-transmission technique for channels with intersymbol interference. *IEEE Trans. Commun.*, COM-20:774–780, Aug. 1972.

[11] R. M. Corless, G. H. Gonnet, D. E. G. Hare, D. J. Jeffrey, and D. E. Knuth. On the Lambert W function. *Adv. Comput. Math.*, 5:329–359, 1996.