A quadratic programming problem arising from vector precoding in wireless communications

Ralf R Müller¹, Dongning Guo² and Aris L Moustakas³
¹ Department of Electronics & Telecommunications, Norwegian University of Science and Technology, Trondheim, Norway
² Department of Electrical Engineering & Computer Science, Northwestern University, Evanston, IL, USA
³ Physics Department, National & Kapodistrian University of Athens, Athens, Greece
E-mail: ralf@iet.ntnu.no, dGuo@northwestern.edu, arislm@phys.uoa.gr

Abstract. A quadratic programming problem is studied in the limit of asymptotically large kernel matrices by means of the replica method. It is found that inverse Wishart kernels are—with within the validity range of the replica symmetric solution—asymptotically invariant to Cartesian relaxations. In the context of vector precoding for wireless communication systems with dual antenna arrays, so-called MIMO systems, this implies that adding more transmit antennas cannot reduce the minimum required transmit energy per bit significantly. By contrast, a new convex relaxation is proposed and shown to be a practical and useful method.

1. Introduction
Non-convex quadratic programming problems arise in many applications in wireless communications. One of them is vector precoding for wireless multiple-input multiple-output (MIMO) systems [1, 2]. Wireless MIMO technology promises to increase data rate over conventional wireless communications without need for more physical radio spectrum by means of multiple antenna elements at both the transmitter side and the receiver side [3, 4]. As a design choice, there can be need for major signal processing at the receiver side, the transmitter side or both of them. Here, we are concerned with systems where sophisticated signal processing is required solely at the transmitter side. This is advantageous for transmitting data to low-cost or battery-driven devices such as cell-phones and personal digital assistants (PDAs).

It is an unavoidable feature of wireless MIMO systems that signals sent from different antenna elements of the transmit array are received with severe crosstalk at the respective antenna elements of the receive array. In order to compensate for the crosstalk, one can use linear joint transmitter processing, also known as linear vector precoding, as suggested in [5, 6]. This comes, however, at the expense of the need for an increased transmit power in order to maintain the distance properties of the received signal constellation. A more sophisticated method for transmitter processing is nonlinear vector precoding, which is simply referred to as vector precoding in this work. It is based on the concept of Tomlinson-Harashima precoding [7, 8] which was originally proposed to combat intersymbol interference. It was proposed for use in the context of MIMO systems in [1, 2]. For a general survey on vector precoding the reader is referred to [9].
In this work, we are mainly concerned with the performance analysis of vector precoding in the limit where the number of antennas goes to infinity. This problem was addressed in [10] employing the replica method which was originally invented for the analysis of spin glasses in statistical physics [11, 12] and has become increasingly powerful to address problems in wireless communications and coding theory [13].

The remainder of the paper consists of five sections. Section 2 shows that inverse Wishart kernels are asymptotically invariant to Cartesian relaxations. Section 3 introduces vector precoding and discusses the implications of section 2. Section 4 proposes a new convex relaxation of the signal set to reduce the complexity of vector precoding and compares its performance to the benchmarks found in [10]. Section 5 concludes the paper.

2. An invariance property of the inverse Wishart kernel

Consider the following quadratic programming problem

\[ E(\rho) = \frac{1}{K} \min_{x \in \mathcal{X}} x^\dagger J x \]

with

\[ \mathcal{X} = \mathcal{X}_1 \times \mathcal{X}_2 \times \cdots \times \mathcal{X}_K \times \mathcal{R}, \]

\[ \mathcal{X}_k \subseteq \mathbb{C}, k = 1, 2, \ldots \text{ and } \mathcal{R} \subseteq \mathbb{C}^{[\rho K]}. \]

In the limit \( K \to \infty \), the problem is well-defined for any \( \rho \geq 0 \) and any random matrix \( J \in \mathbb{C}^{[K+\rho K] \times [K+\rho K]} \) which provides enough randomness to make the quadratic programming problem self-averaging.

We call the problem \( E(\rho) \) for any \( \rho > 0 \) a Cartesian relaxation of the problem \( E(0) \), since the constraint set is relaxed by a Cartesian product between the original constraint set and the relaxation set \( \mathcal{R} \). Since relaxations cannot hurt, we have

\[ \frac{dE(\rho)}{d\rho} \leq 0. \]

In vector precoding for wireless communication systems the kernel takes the form of an inverse Wishart matrix. That is, let \( H \in \mathbb{C}^{K \times K/\alpha}, \alpha \leq 1 \) be composed of independent identically distributed zero-mean complex Gaussian entries, we have

\[ J = (HH^\dagger)^{-1}. \]

Intuitively, one would expect (3) to hold with strict inequality in most non-trivial cases. Thus, when studying such relaxations by means of replica symmetric analysis, it came as a surprise to the authors to discover that, for an inverse Wishart matrix, (3) holds with equality, i.e. the (rather generous) relaxation of the constraint set does not reduce the objective function.

We will devote the remainder of this section to manifest this invariance property of the inverse Wishart kernel within the limitations of the replica-symmetric ansatz. We are aware of the fact that quadratic programming problems, as they can be formulated as ground state energy problems, are vulnerable to replica-symmetry breaking. Nevertheless, the replica-symmetric ansatz still provides strict inequality in (3) for e.g. the standard Wishart kernel and presumably most other random matrices. If after all, it will turn out that the discovered invariance is not a feature of the inverse Wishart kernel but the replica symmetric solution of the quadratic programming problem, we hope that our discovery can still be helpful for a better insight into replica symmetry and its limitations.

We start from a result we recently showed in context of vector precoding for wireless multi-antenna communication systems [10].
Proposition 1 Let the following assumption hold:

(i) The random matrix \( J \) is bi-unitarily invariant, i.e. it can be decomposed into \( J = O D O \)\( ^\dagger \) such that the matrices \( D \) and \( O \) are diagonal and Haar distributed, respectively. Moreover, as \( K \to \infty \), the asymptotic eigenvalue distribution of \( J \) converges to a non-random distribution function which can be uniquely characterized by its R-transform \( R(w) \).

(ii) The quadratic programming problem (1) is self-averaging.

(iii) The parameters \( q \) and \( b \) are solutions to the following pair of coupled fixed-point equations

\[
q = \lim_{K \to \infty} \frac{1}{K} \sum_{k=1}^{K} \left| \argmin_{x \in \chi_k} \left| z \sqrt{\frac{qR'(b)}{2R^2(b)}} - x \right|^2 \right| Dz
\]

\[
b = \lim_{K \to \infty} \frac{1}{K} \sum_{k=1}^{K} \Re \left\{ \argmin_{x \in \chi_k} \left| z \sqrt{\frac{qR'(b)}{2R^2(b)}} - x \right| z^* \right\} \frac{Dz}{\sqrt{2qR'(b)}}.
\]

with \( Dz = \exp(-z^2/2)/(2\pi)dz \) being the complex Gaussian measure.

Then, if \( 0 < b < \infty \), the replica symmetric ansatz yields

\[
E(0) \to q \frac{\partial}{\partial b} bR(-b)
\]

as \( K \to \infty \).

In order to fit (1) into the framework of Proposition 1, we first let \( R = C^{(aK)} \) without loss of generality. Thus, we find

\[
E(\rho) \to (1 + \rho) \tilde{q} \frac{\partial}{\partial b} bR(-\tilde{b})
\]

with

\[
\tilde{q} = \frac{1}{1 + \rho} \left[ \rho \frac{\tilde{q}R'(-\tilde{b})}{R^2(-\tilde{b})} + \lim_{K \to \infty} \frac{1}{K} \sum_{k=1}^{K} \left| \argmin_{x \in \chi_k} \left| z \sqrt{\frac{\tilde{q}R'(\tilde{b})}{2R^2(\tilde{b})}} - x \right|^2 \right| Dz \right]
\]

\[
\tilde{b} = \frac{1}{1 + \rho} \left[ \frac{\rho}{R(-\tilde{b})} + \lim_{K \to \infty} \frac{1}{K} \sum_{k=1}^{K} \Re \left\{ \argmin_{x \in \chi_k} \left| z \sqrt{\frac{\tilde{q}R'(\tilde{b})}{2R^2(\tilde{b})}} - x \right| z^* \right\} \frac{Dz}{\sqrt{2\tilde{q}R'(\tilde{b})}} \right]
\]

For the inverse Wishart matrix we find from [10]

\[
R(-\tilde{b}) = \frac{\tilde{\alpha} - 1 + \sqrt{(1 - \tilde{\alpha})^2 + 4\tilde{\alpha}\tilde{b}}}{2\tilde{\alpha}\tilde{b}}
\]

\[
R'(-\tilde{b}) = \frac{\left( \frac{\tilde{\alpha} - 1 + \sqrt{(1 - \tilde{\alpha})^2 + 4\tilde{\alpha}\tilde{b}}}{2\tilde{\alpha}\tilde{b}} \right)^2}{4\tilde{\alpha}\tilde{b}^2(1 - \tilde{\alpha})^2 + 4\tilde{\alpha}\tilde{b}}
\]

with \( \tilde{\alpha} = (1 + \rho)\alpha \) and \( \alpha \) being a free parameter of the Wishart distribution. It also turns out helpful to recognize that

\[
\frac{R^2(-\tilde{b})}{R(-\tilde{b})} = \frac{(1 - \tilde{\alpha})^2 + 4\tilde{\alpha}\tilde{b}}{\tilde{\alpha}}
\]

1 The R-transform of a measure \( \mu(X) \) is given as \( R(w) = G^{-1}(-w) - w^{-1} \) where \( G^{-1}(\cdot) \) is the functional inverse (inverse with respect to composition) of \( G(\cdot) \) and \( G(w) = f(X - w)^{-1}\mu(X) \) [14, 15].
and it is convenient to replace the parameter $\tilde{b}$ by the substitution

$$
\tilde{p} = \sqrt{(1 - \tilde{\alpha})^2 + 4\tilde{\alpha}\tilde{b}}.
$$

Thus, we get

$$
\tilde{q} = \frac{1}{(1 + \rho)} \left[ \rho \frac{\tilde{q}_\alpha}{\tilde{p}} + \lim_{K \to \infty} \frac{1}{K} \sum_{k=1}^{K} \left( \arg\min_{x \in \mathcal{X}_k} \left| z \sqrt{\frac{\tilde{q}_\alpha}{2\tilde{p}}} - x \right|^2 \right) \right]
$$

$$
\tilde{b} = \frac{1}{(1 + \rho)R(-\tilde{b})} \left[ \rho + \sqrt{\frac{\tilde{p}}{2\tilde{q}_\alpha}} \lim_{K \to \infty} \frac{1}{K} \sum_{k=1}^{K} \int \mathcal{R} \left\{ \arg\min_{x \in \mathcal{X}_k} \left| z \sqrt{\frac{\tilde{q}_\alpha}{2\tilde{p}}} - x \right|^2 \right\} \right]
$$

$$
E(\rho) = \frac{1}{(1 + \rho) R(-\tilde{b})} \left[ \rho + \sqrt{\frac{\tilde{p}}{2\tilde{q}_\alpha}} \lim_{K \to \infty} \frac{1}{K} \sum_{k=1}^{K} \int \mathcal{R} \left\{ \arg\min_{x \in \mathcal{X}_k} \left| z \sqrt{\frac{\tilde{q}_\alpha}{2\tilde{p}}} - x \right|^2 \right\} \right]
$$

Plugging (11) into (16), and re-arranging terms, we find

$$
\tilde{p} = 1 - \tilde{\alpha} + \frac{2\tilde{\alpha}}{1 + \rho} \left[ \rho + \sqrt{\frac{\tilde{p}}{2\tilde{q}_\alpha}} \lim_{K \to \infty} \frac{1}{K} \sum_{k=1}^{K} \int \mathcal{R} \left\{ \arg\min_{x \in \mathcal{X}_k} \left| z \sqrt{\frac{\tilde{q}_\alpha}{2\tilde{p}}} - x \right|^2 \right\} \right].
$$

Thus, we find from (17)

$$
E(\rho) = \frac{\rho E(\rho)\alpha + \lim_{K \to \infty} \frac{1}{K} \sum_{k=1}^{K} \int \mathcal{R} \left\{ \arg\min_{x \in \mathcal{X}_k} \left| z \sqrt{\frac{E(\rho)\alpha}{2}} - x \right|^2 \right\} \right]}{1 - \tilde{\alpha} + 2\alpha \tilde{\rho} + \sqrt{\frac{2\tilde{\alpha}}{E(\rho)\alpha}} \lim_{K \to \infty} \frac{1}{K} \sum_{k=1}^{K} \int \mathcal{R} \left\{ \arg\min_{x \in \mathcal{X}_k} \left| z \sqrt{\frac{E(\rho)\alpha}{2}} - x \right|^2 \right\} \right]}
$$

and after some trivial algebra

$$
E(\rho) = \frac{\lim_{K \to \infty} \frac{1}{K} \sum_{k=1}^{K} \int \mathcal{R} \left\{ \arg\min_{x \in \mathcal{X}_k} \left| z \sqrt{\frac{E(\rho)\alpha}{2}} - x \right|^2 \right\} \right]}{1 - \alpha + \sqrt{\frac{2\tilde{\alpha}}{E(\rho)\alpha}} \lim_{K \to \infty} \frac{1}{K} \sum_{k=1}^{K} \int \mathcal{R} \left\{ \arg\min_{x \in \mathcal{X}_k} \left| z \sqrt{\frac{E(\rho)\alpha}{2}} - x \right|^2 \right\} \right]}
$$

Clearly, $E(\rho)$ does not depend on $\rho$ and the proof is complete.

### 3. Vector Precoding

Following [10], vector precoding is conveniently introduced in a general way as follows:

Vector precoding aims to minimize the transmitted power that is associated with the transmission of a certain data vector $s \in S^K$ of length $K$. For that purpose, the original symbol alphabet $S$ is relaxed into the alphabet $\mathcal{B}$. The data representation in the relaxed alphabet is redundant. That means that several symbols in the relaxed alphabet represent the same data. Due to the redundant representation, we can now choose that representation of our data which requires the least power for transmission. This way of saving transmit power is what we call vector precoding.

That means, for any $s \in S$ there is a set $\mathcal{B}_s \subset \mathcal{B}$ such that all elements of $\mathcal{B}_s$ represent the data $s$. Take quaternary phase-shift keying (QPSK) as an example, i.e., $S = \{00, 01, 10, 11\}$. Let $j = \sqrt{-1}$. Without vector precoding, it is most common (Gray mapping) to choose $\mathcal{B}_{00} = \{+1 + j\}$, $\mathcal{B}_{01} = \{+1 - j\}$, $\mathcal{B}_{10} = \{-1 + j\}$, and $\mathcal{B}_{11} = \{-1 - j\}$. For QPSK modulation, vector precoding
Figure 1. Four equally spaced integer lattices representing the four QPSK states 00, 01, 10, and 11, respectively.

is the idea to have $B_{00} \supset \{+1 + j\}$, $B_{01} \supset \{+1 - j\}$, $B_{10} \supset \{-1 + j\}$ and $B_{11} \supset \{-1 - j\}$, i.e. to allow for supersets of the quaternary constellation. A popular choice for those supersets is due to Tomlinson, Harashima, and Miyakawa [7, 8] where

$B_{00} = (4Z + 1) \times (4jZ + j)$,
$B_{01} = (4Z + 1) \times (4jZ - j)$,
$B_{10} = (4Z - 1) \times (4jZ + j)$, and
$B_{11} = (4Z - 1) \times (4jZ - j)$ are disjoint subsets of the odd integer lattice points (see also figure 1).

In order to avoid ambiguities, we should have

$B_i \cap B_j = \emptyset \quad \forall i \neq j$. \hspace{1cm} (21)

In addition, one would like to design the sets $B_i$ such that the distance properties of the received signals are preserved. This is for instance easily achieved by letting the sets $B_i$ to be distinct sub-lattices of $B$ as in figure 1. However, the choice of non-convex sets $B_i$ in general leads to NP-hardness of the vector precoding problem. Therefore, it is advantageous to use convex sets $B_i$ even if that increases the transmitted energy per symbol for sake of feasibility.

Let $s = [s_1, \ldots, s_K]^T$ denote the information to be encoded. Let $t = T \mathbf{x}$ be the $K/\alpha \times 1$ vector being sent where $\alpha$ denotes the ratio between data streams to transmit antennas and the $K \times 1$ vector $\mathbf{x}$ be the redundant representation of the data. Then, the precoding problem can be written as the minimization of the following quadratic form

$\min_{\mathbf{x} \in \mathcal{X}} ||T\mathbf{x}||_2^2 = \min_{\mathbf{x} \in \mathcal{X}} \mathbf{x}^\dagger J\mathbf{x}$ \hspace{1cm} (22)

over the discrete set

$\mathcal{X} = B_{s_1} \times B_{s_2} \times \cdots \times B_{s_K}$ \hspace{1cm} (23)

with $J = T^\dagger T$.

Consider a vector-valued communication system. Let the received vector be given as

$\mathbf{r} = Ht + \mathbf{n}$ \hspace{1cm} (24)
where \( n \) is additive noise. Let the components of the transmitted and received vectors be signals sent and received at different antenna elements, respectively.

We want to ensure that the received signal is (up to additive noise) identical to the data vector. This design criteria leads us to choose the precoding matrix

\[
T = H' \left( HH' \right)^{-1}
\]

This means that we invert the channel and get \( r = x + n \) if the matrix inverse exists. This allows to keep the signal processing at the receiver at a minimum. This is advantageous if the receiver shall be a low-cost or battery-powered device.

To model the statistics of the entries of \( H \) is a non-trivial task and a topic of ongoing research, see e.g. [16] and references therein. For sake of convenience, we choose in this first order approach that the entries of the channel matrix \( H \) are independent and identically distributed complex Gaussian random variables with zero mean and variance \( \alpha/K \). Thus, the kernel matrix \( J \) is an inverse Wishart matrix.

For inverse Wishart matrices, the invariance property found in section 2 applies. But what does it mean in context of vector precoding? A Cartesian relaxation in a MIMO system with vector precoding would be to add a couple of receive antennas, precisely speaking \( \rho K \) of them to the \( K \) receive antennas already in use and design dummy data-streams for those receive antennas. The dummy data does not bear any useful information, but intends to reduce the radiated power at the transmit antennas. The invariance property for inverse Wishart matrices means in this context that, regardless of how many additional receive antennas we provide and regardless how we design the dummy data, the radiated power at the transmitter will not decay.

The situation does not become better for Cartesian relaxations if we are willing to compromise on the data rate. That is, we do not add receive antennas, but reduce the total data rate in order to free some of the existing receive antennas for the dummy data. In such a case we would reduce the radiated power at the transmit antennas. However, we could achieve the same power saving by simply shutting down those receive antennas that are meant to receive the dummy data.

### 4. A new convex relaxation

We aim to perform vector precoding at feasible complexity. Lattice precoding performs very well [10], but is NP-hard. We have seen in section 2 that Cartesian relaxations do not help. In order to allow for polynomial complexity, we should allow for convex sets. Since we should not reduce the minimum distance between the signal constellation, it seems intuitive to try the following convex relaxation:

\[
\begin{align*}
S &= \{00, 01, 10, 11\} \\
B_{00} &= \{z \in \mathbb{C} : \Re z \leq -1 \land \Im z \leq -1\} \\
B_{01} &= \{z \in \mathbb{C} : \Re z \leq -1 \land \Im z \geq +1\} \\
B_{10} &= \{z \in \mathbb{C} : \Re z \geq +1 \land \Im z \leq -1\} \\
B_{11} &= \{z \in \mathbb{C} : \Re z \geq +1 \land \Im z \geq +1\}
\end{align*}
\]

These signal sets are depicted in figure 2. All these signal sets are convex. The kernel matrix \( J \) is positive-semidefinite. Therefore, the corresponding quadratic programming problems can be solved in polynomial time. Furthermore, the mapping is independent of each other in both quadrature components.

From Proposition 1 and (14), we get

\[
q = \int_{z \in B_{11}} \sqrt{\frac{q R'(-b)}{2 R^2(-b)} - x^2}^2 \, Dz
\]

\[6\]
Figure 2. Complex convex relaxation representing the four quaternary states 00, 01, 10, and 11, respectively.

\[
\begin{align*}
\mathbb{R} \left\{ \argmin_{x \in B_{11}} \left| z \right| \sqrt{\frac{q R''(-b)}{R^2(-b)}} - x \left| z^* \right| \right\} & \left( \frac{2p}{2q} \right) \exp \left( -\frac{z^2}{2} \right) \frac{Dz}{\sqrt{2q R''(-b)}} \\
\int_{-\infty}^{+\infty} \left\{ 1 \right. \\
\int_{-\infty}^{+\infty} \left\{ z^2 \sqrt{\frac{q R''(-b)}{R^2(-b)}} - \frac{p}{2q} \right\} \exp \left(-\frac{z^2}{2}\right) dz \\
\int_{-\infty}^{+\infty} \left\{ z \sqrt{\frac{q R''(-b)}{R^2(-b)}} - \frac{p}{2q} \right\} \exp \left(-\frac{z^2}{2}\right) dz \\
\int_{-\infty}^{+\infty} \left\{ z \sqrt{\frac{q R''(-b)}{R^2(-b)}} - \frac{p}{2q} \right\} \exp \left(-\frac{z^2}{2}\right) dz \\
\int_{-\infty}^{+\infty} \left\{ z \sqrt{\frac{q R''(-b)}{R^2(-b)}} - \frac{p}{2q} \right\} \exp \left(-\frac{z^2}{2}\right) dz \\
\int_{-\infty}^{+\infty} \left\{ z \sqrt{\frac{q R''(-b)}{R^2(-b)}} - \frac{p}{2q} \right\} \exp \left(-\frac{z^2}{2}\right) dz \\
\int_{-\infty}^{+\infty} \left\{ z \sqrt{\frac{q R''(-b)}{R^2(-b)}} - \frac{p}{2q} \right\} \exp \left(-\frac{z^2}{2}\right) dz
\end{align*}
\]

with \( Q(x) = \int_{-\infty}^{+\infty} \exp(-x^2/2) dx / \sqrt{2\pi} \). Finally, we can use (36) to simplify (33)

\[
\begin{align*}
q & = 2 + \frac{2 - 2p}{\alpha} - q \left( 1 - \alpha \right) + 2 \sqrt{\frac{q R''(-b)}{R^2(-b)}} \exp \left(-\frac{p}{q R''(-b)}\right). \\
(37)
\end{align*}
\]

The solutions to these fixed point equations are shown in figure 3. For unit load, the energy per symbol is found to be approximately 3 times higher than with lattice precoding. This is the price to be paid for precoding with polynomial complexity. In comparison to sole channel inversion, the proposed convex relaxation allows to increase the load by about 30% without need for more transmitted energy. This means that the transmitter can achieve the same performance with at least 30% fewer transmit antennas.
The considerations in the above paragraph concerning the performance of the proposed convex relaxation are somewhat pessimistic. Admittedly, the energy per symbol is about 3 times larger than with lattice precoding. However, this is not the full story. Lattice precoding preserves the average minimum distance to signal points of other sets. The proposed convex relaxation, however, increases the average minimum distance and, therefore, makes the signal more robust against channel impairments such as additive noise and fading.

The increase in average minimum distance is directly quantified by the parameter

\[ q = \lim_{K \to \infty} \frac{1}{K} x^\top x. \]  

(38)

This parameter is shown in figure 4. For unit load, the parameter \( q \) takes the value of approximately 5.8 which suggests an increase of the average squared distance by a factor of approximately 2.9. Thus, tripling the transmit power in comparison to lattice precoding, comes almost with a tripling of average squared distance.

5. Conclusions

We found that Cartesian relaxations do not help to reduce the transmitted power for inversion of vector channels with independent identically distributed channel matrices. Instead the convex relaxation proposed in this work gives noticeable gains over pure channel inversion for loads close to unity. The convex relaxation also comes with an increase in the average minimum distance between signal points which provides additional robustness against additive and multiplicative channel noise.
Acknowledgments
This research was supported by the Research Council of Norway and the European Commission under grants 171133/V30 and MIRG-CT-2005-030833, resp. This work was initiated while R. Müller and D. Guo were visiting the Institute for Mathematical Sciences at the National University of Singapore in 2006.

References
[1] Windpassinger C, Fischer R F H, Vencel T and Huber J B 2004 IEEE Transactions on Wireless Communications 3 1305–1316
[2] Hochwald B M, Ped C and Swindlehurst A 2005 IEEE Transactions on Communications 53 537–544
[3] Foschini G and Gans M 1998 Wireless Personal Communications 6 311–335
[4] Telatar I E 1999 European Transactions on Telecommunications 10 585–595
[5] Vojčić B R and Jang W M 1998 IEEE Transactions on Communications 46 1346–1355
[6] Peel C B, Hochwald B M and Swindlehurst A L 2005 IEEE Transactions on Communications 53 195–202
[7] Tomlinson M 1971 IEE Electronics Letters 7 138–139
[8] Harashima H and Miyakawa H 1972 IEEE Transactions on Communications COM-20 774–780
[9] Fischer R F 2002 Predecoding and Signal Shaping for Digital Transmission (John Wiley & Sons)
[10] Müller R R, Guo D and Moustakas A L 2007 Submitted to IEEE Journal on Selected Areas in Communications ArXiv:0706.1169v1 [cs.IT]
[11] Mezard M, Parisi G and Virasoro M A 1987 Spin Glass Theory and Beyond (Singapore: World Scientific)
[12] Fischer K and Hertz J 1991 Spin Glasses (Cambridge, U.K.: Cambridge University Press)
[13] Nishimori H 2001 Statistical Physics of Spin Glasses and Information Processing (Oxford, U.K.: Oxford University Press)
[14] Voiculescu D V, Dykema K J and Nica A 1992 Free Random Variables (Providence, RI: American Mathematical Society)
[15] Tulino A M and Verdú S 2004 Foundations and Trends in Communications and Information Theory 1
[16] Debbah M and Müller R 2005 IEEE Transactions on Information Theory 51 1667–1690