

A Combinatorial Optimization Problem in Wireless Communications and Its Analysis

Ralf R. Müller

EE Dept, NTNU

Benjamin Zaidel

Tel Aviv

Dongning Guo

EECS Dept, Northwestern

Rodrigo de Miguel

SINTEF, Trondheim

Aris Moustakas

Physics Dept, NCUA

Vesna Gardašević

EE Dept, NTNU

Finn Knudsen

Math Dept, NTNU

The Problem

Let

$$E := \frac{1}{K} \min_{\mathbf{x} \in \mathcal{X}} \mathbf{x}^\dagger \mathbf{J} \mathbf{x}$$

with $\mathbf{x} \in \mathbb{C}^K$ and $\mathbf{J} \in \mathbb{C}^{K \times K}$.

The Problem

Let

$$E := \frac{1}{K} \min_{\mathbf{x} \in \mathcal{X}} \mathbf{x}^\dagger \mathbf{J} \mathbf{x}$$

with $\mathbf{x} \in \mathbb{C}^K$ and $\mathbf{J} \in \mathbb{C}^{K \times K}$.

Example 1 (sphere):

$$\mathcal{X} = \{\mathbf{x} : \mathbf{x}^\dagger \mathbf{x} = K\} \implies E = \min \lambda(\mathbf{J})$$

The Problem

Let

$$E := \frac{1}{K} \min_{\mathbf{x} \in \mathcal{X}} \mathbf{x}^\dagger \mathbf{J} \mathbf{x}$$

with $\mathbf{x} \in \mathbb{C}^K$ and $\mathbf{J} \in \mathbb{C}^{K \times K}$.

Example 1 (sphere):

$$\mathcal{X} = \{\mathbf{x} : \mathbf{x}^\dagger \mathbf{x} = K\} \implies E = \min \lambda(\mathbf{J})$$

Example 2 (cube):

$$\mathcal{X} = \{+1, -1\}^K \implies ???$$

The Problem

Let

$$E := \frac{1}{K} \min_{\mathbf{x} \in \mathcal{X}} \mathbf{x}^\dagger \mathbf{J} \mathbf{x}$$

with $\mathbf{x} \in \mathbb{C}^K$ and $\mathbf{J} \in \mathbb{C}^{K \times K}$.

Example 1 (sphere):

$$\mathcal{X} = \{\mathbf{x} : \mathbf{x}^\dagger \mathbf{x} = K\} \implies E = \min \lambda(\mathbf{J})$$

Example 2 (cube):

$$\mathcal{X} = \{+1, -1\}^K \implies ???$$

Example 3 (vector precoding):

$$\mathcal{X} = (4\mathbb{Z} + 1)^K \implies ???$$

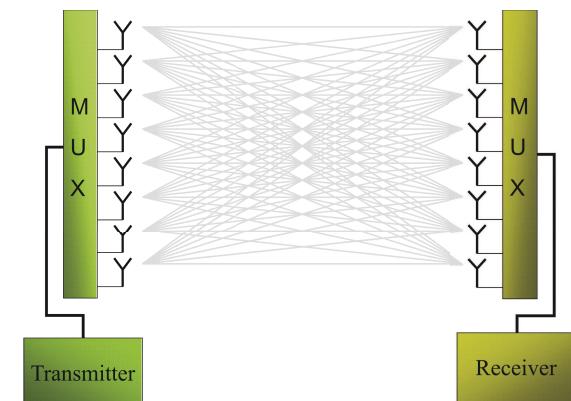
The Gaussian Vector Channel

Let the received vector be given by

$$\mathbf{y} = \mathbf{H}\mathbf{t} + \mathbf{n}$$

where

- \mathbf{t} is the transmitted vector
- \mathbf{n} is uncorrelated (white) Gaussian noise
- \mathbf{H} is a coupling matrix accounting for crosstalk



In many applications, e.g. antenna arrays, code-division multiple-access, the coupling matrix is modelled as a random matrix with independent identically distributed entries (i.i.d. model).

Crosstalk can be processed either at receiver or transmitter



Processing at Transmitter

If the transmitter is a base-station and the receiver is a hand-held device one would prefer to have the complexity at the transmitter.

E.g. let the transmitted vector be

$$\mathbf{t} = \mathbf{H}^\dagger (\mathbf{H} \mathbf{H}^\dagger)^{-1} \mathbf{x}$$

where $\mathbf{x} = \mathbf{s}$ is the data to be sent.

Then,

$$\mathbf{y} = \mathbf{s} + \mathbf{n}.$$

No crosstalk anymore due to channel inversion.

Problems of Simple Channel Inversion

Channel inversion implies a significant power amplification, i.e.

$$\mathbf{x}^\dagger (\mathbf{H}\mathbf{H}^\dagger)^{-1} \mathbf{x} > \mathbf{x}^\dagger \mathbf{x}.$$

In particular, let

- $\alpha = \frac{K}{N} \leq 1$;
- the entries of \mathbf{H} are i.i.d. with variance $1/N$.

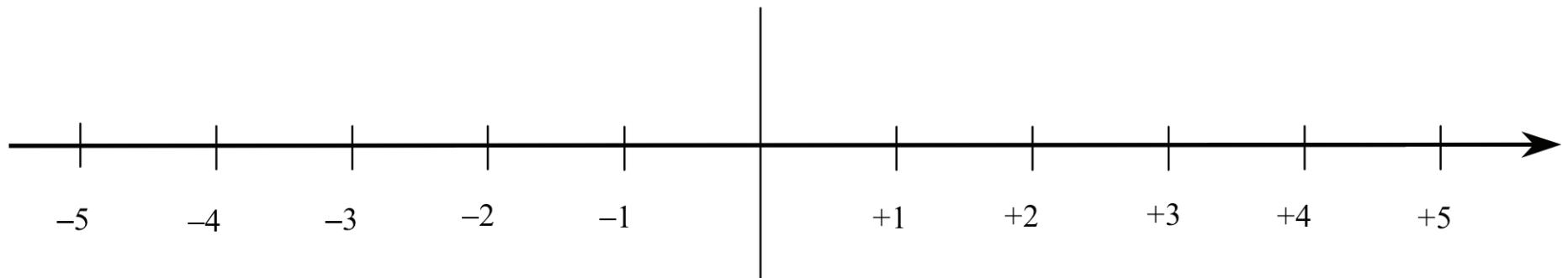
Then, for fixed aspect ratio α

$$\lim_{K \rightarrow \infty} \frac{\mathbf{x}^\dagger (\mathbf{H}\mathbf{H}^\dagger)^{-1} \mathbf{x}}{\mathbf{x}^\dagger \mathbf{x}} = \frac{1}{1 - \alpha}$$

with probability 1.

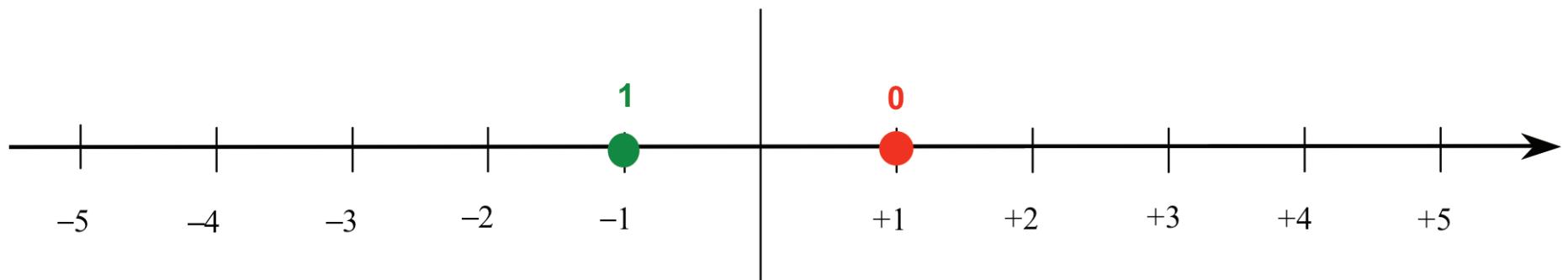
Lattice-Relaxation Precoding

Tomlinson '71, Harashima & Miyakawa '72



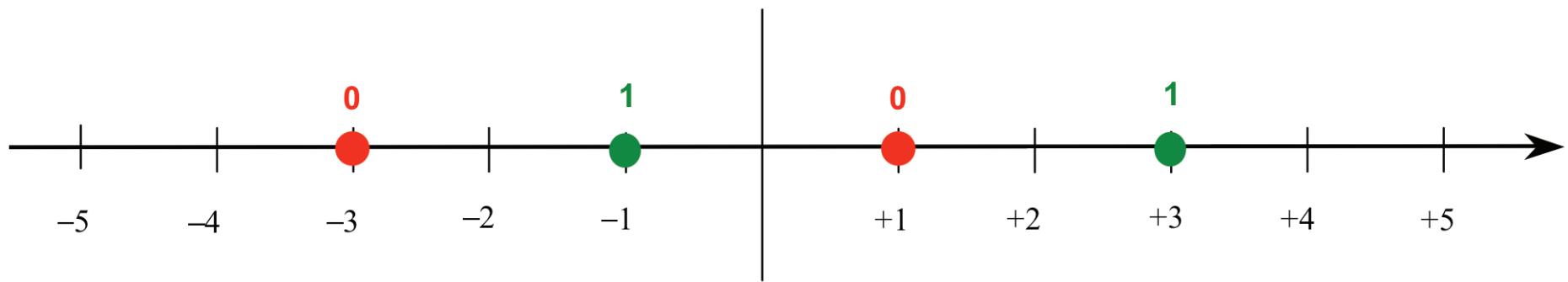
Lattice-Relaxation Precoding

Tomlinson '71, Harashima & Miyakawa '72



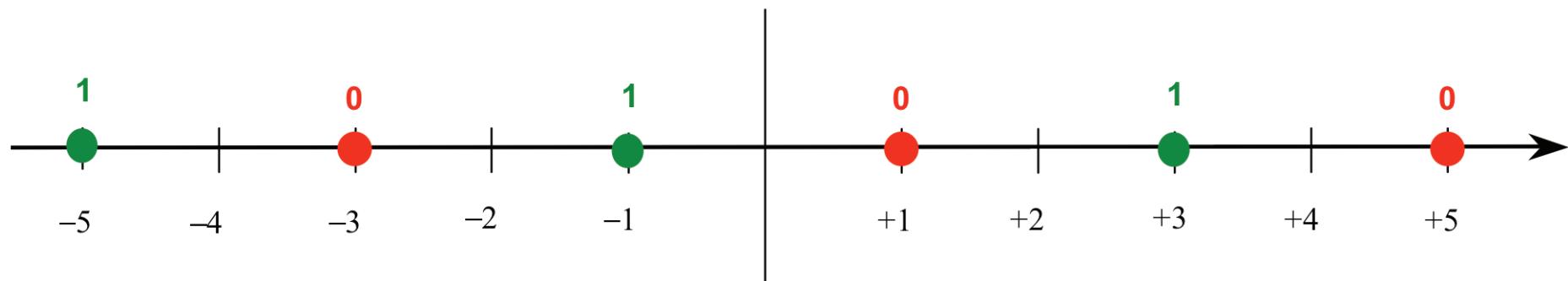
Lattice-Relaxation Precoding

Tomlinson '71; Harashima & Miyakawa '72



Lattice-Relaxation Precoding

Tomlinson '71, Harashima & Miyakawa '72



Instead of representing the logical "0" by +1, we present it by any element of the set $\{\dots, -7, -3, +1, +5, \dots\} = 4\mathbb{Z} + 1$. Correspondingly, the logical "1" is represented by any element of the set $4\mathbb{Z} - 1$.

Choose that representation that gives the smallest transmit power.

General Relaxation Precoding

Let \mathcal{B}_0 and \mathcal{B}_1 denote the sets presenting 0 and 1 , resp.

Let $(s_1, s_2, s_3, \dots, s_K) \in \{0, 1\}^K$ denote the data to be transmitted.

Then, the transmitted energy per data symbol is given by

$$E = \frac{1}{K} \min_{\mathbf{x} \in \mathcal{X}} \mathbf{x}^\dagger \mathbf{J} \mathbf{x}$$

with

$$\mathcal{X} = \mathcal{B}_{s_1} \times \mathcal{B}_{s_2} \times \dots \times \mathcal{B}_{s_K}$$

and

$$\mathbf{J} = (\mathbf{H} \mathbf{H}^\dagger)^{-1}.$$

What is a smart choice for \mathcal{B}_0 and \mathcal{B}_1 ?

Zero Temperature Formulation

Quadratic programming is the problem of finding the zero temperature limit of a quadratic energy potential.

The transmitted power is written as a zero temperature limit

$$E = - \lim_{\beta \rightarrow \infty} \frac{1}{\beta K} \log \sum_{\mathbf{x} \in \mathcal{X}} e^{-\beta \text{tr}(\mathbf{x}^\dagger \mathbf{J} \mathbf{x})}$$

with $\frac{1}{\beta}$ denoting temperature.

Zero Temperature Formulation

Quadratic programming is the problem of finding the zero temperature limit of a quadratic energy potential.

The transmitted power is written as a zero temperature limit

$$\begin{aligned} E &= - \lim_{\beta \rightarrow \infty} \frac{1}{\beta K} \log \sum_{\mathbf{x} \in \mathcal{X}} e^{-\beta \text{tr}(\mathbf{x}^\dagger \mathbf{J} \mathbf{x})} \\ &\xrightarrow{\text{---}} - \lim_{\beta \rightarrow \infty} \lim_{K \rightarrow \infty} \mathbf{E}_{\mathbf{J}} \frac{1}{\beta K} \log \sum_{\mathbf{x} \in \mathcal{X}} e^{-\beta \text{tr}(\mathbf{J} \mathbf{x} \mathbf{x}^\dagger)} \end{aligned}$$

with $\frac{1}{\beta}$ denoting temperature.

The Harish-Chandra Integral

(also called the Itzykson-Zuber integral)

Let \mathbf{P} be any positive semi-definite matrix of bounded rank n and let \mathbf{J} be bi-unitarily invariant. Then,

$$\lim_{K \rightarrow \infty} \frac{1}{K} \log \mathbb{E}_{\mathbf{J}} e^{-K \operatorname{tr} \mathbf{J} \mathbf{P}} = - \sum_{a=1}^n \int_0^{\lambda_a(\mathbf{P})} R_{\mathbf{J}}(-w) dw$$

with λ_a denoting the positive eigenvalues of \mathbf{P} and $R_{\mathbf{J}}(w)$ denoting the R-transform of the spectral measure of \mathbf{J} (Marinari et al. '94; Guionnet & Maïda '05).

The Replica Method

We want

$$\lim_{K \rightarrow \infty} \frac{1}{K} \mathbb{E}_{\mathbf{J}} \log \sum_{\mathbf{x} \in \mathcal{X}} e^{-\beta \text{tr}(\mathbf{J} \mathbf{x} \mathbf{x}^\dagger)} = \lim_{K \rightarrow \infty} \lim_{n \rightarrow 0} \frac{1}{nK} \log \mathbb{E}_{\mathbf{J}} \left(\sum_{\mathbf{x} \in \mathcal{X}} e^{-\beta \text{tr}(\mathbf{J} \mathbf{x} \mathbf{x}^\dagger)} \right)^n$$

The Replica Method

We want

$$\begin{aligned} \lim_{K \rightarrow \infty} \frac{1}{K} \mathbb{E}_{\mathbf{J}} \log \sum_{\mathbf{x} \in \mathcal{X}} e^{-\beta \text{tr}(\mathbf{J} \mathbf{x} \mathbf{x}^\dagger)} &= \lim_{K \rightarrow \infty} \lim_{n \rightarrow 0} \frac{1}{nK} \log \mathbb{E}_{\mathbf{J}} \left(\sum_{\mathbf{x} \in \mathcal{X}} e^{-\beta \text{tr}(\mathbf{J} \mathbf{x} \mathbf{x}^\dagger)} \right)^n \\ &= \lim_{K \rightarrow \infty} \lim_{n \rightarrow 0} \frac{1}{nK} \log \mathbb{E}_{\mathbf{J}} \prod_{a=1}^n \sum_{\mathbf{x}_a \in \mathcal{X}} e^{-\beta \text{tr}(\mathbf{J} \mathbf{x}_a \mathbf{x}_a^\dagger)} \end{aligned}$$

The Replica Method

We want

$$\begin{aligned}
 \lim_{K \rightarrow \infty} \frac{1}{K} \mathbb{E}_{\mathbf{J}} \log \sum_{\mathbf{x} \in \mathcal{X}} e^{-\beta \text{tr}(\mathbf{J} \mathbf{x} \mathbf{x}^\dagger)} &= \lim_{K \rightarrow \infty} \lim_{n \rightarrow 0} \frac{1}{nK} \log \mathbb{E}_{\mathbf{J}} \left(\sum_{\mathbf{x} \in \mathcal{X}} e^{-\beta \text{tr}(\mathbf{J} \mathbf{x} \mathbf{x}^\dagger)} \right)^n \\
 &= \lim_{K \rightarrow \infty} \lim_{n \rightarrow 0} \frac{1}{nK} \log \mathbb{E}_{\mathbf{J}} \prod_{a=1}^n \sum_{\mathbf{x}_a \in \mathcal{X}} e^{-\beta \text{tr}(\mathbf{J} \mathbf{x}_a \mathbf{x}_a^\dagger)} \\
 &= \lim_{K \rightarrow \infty} \lim_{n \rightarrow 0} \frac{1}{nK} \log \mathbb{E}_{\mathbf{J}} \sum_{\mathbf{x}_1 \in \mathcal{X}} \dots \sum_{\mathbf{x}_n \in \mathcal{X}} e^{-\text{tr} \left(\mathbf{J} \beta \sum_{a=1}^n \mathbf{x}_a \mathbf{x}_a^\dagger \right)}
 \end{aligned}$$

The Replica Method

We want

$$\begin{aligned}
 \lim_{K \rightarrow \infty} \frac{1}{K} \mathbb{E}_{\mathbf{J}} \log \sum_{\mathbf{x} \in \mathcal{X}} e^{-\beta \text{tr}(\mathbf{J} \mathbf{x} \mathbf{x}^\dagger)} &= \lim_{K \rightarrow \infty} \lim_{n \rightarrow 0} \frac{1}{nK} \log \mathbb{E}_{\mathbf{J}} \left(\sum_{\mathbf{x} \in \mathcal{X}} e^{-\beta \text{tr}(\mathbf{J} \mathbf{x} \mathbf{x}^\dagger)} \right)^n \\
 &= \lim_{K \rightarrow \infty} \lim_{n \rightarrow 0} \frac{1}{nK} \log \mathbb{E}_{\mathbf{J}} \prod_{a=1}^n \sum_{\mathbf{x}_a \in \mathcal{X}} e^{-\beta \text{tr}(\mathbf{J} \mathbf{x}_a \mathbf{x}_a^\dagger)} \\
 &= \lim_{K \rightarrow \infty} \lim_{n \rightarrow 0} \frac{1}{nK} \log \mathbb{E}_{\mathbf{J}} \sum_{\mathbf{x}_1 \in \mathcal{X}} \dots \sum_{\mathbf{x}_n \in \mathcal{X}} e^{-\text{tr} \left(\mathbf{J} \beta \sum_{a=1}^n \mathbf{x}_a \mathbf{x}_a^\dagger \right)} \\
 &= \lim_{K \rightarrow \infty} \lim_{n \rightarrow 0} \frac{1}{nK} \log \mathbb{E}_{\mathbf{Q}} \exp \left[-K \sum_{a=1}^n \int_0^{\beta \lambda_a(\mathbf{Q})} R_{\mathbf{J}}(-w) dw \right]
 \end{aligned}$$

with

$$Q_{ab} := \frac{1}{K} \mathbf{x}_a^\dagger \mathbf{x}_b.$$

Laplace Integration

We find

$$\lim_{K \rightarrow \infty} \frac{1}{K} \mathbb{E}_{\mathbf{J}} \log \sum_{\mathbf{x} \in \mathcal{X}} e^{-\beta \text{tr}(\mathbf{J} \mathbf{x} \mathbf{x}^\dagger)} = \lim_{K \rightarrow \infty} \lim_{n \rightarrow 0} \frac{1}{nK} \log \mathbb{E}_{\mathbf{Q}} \exp \left[-K \sum_{a=1}^n \int_0^{\beta \lambda_a(\mathbf{Q})} R_{\mathbf{J}}(-w) dw \right]$$

Laplace Integration

We find

$$\begin{aligned}
 \lim_{K \rightarrow \infty} \frac{1}{K} \mathbb{E}_{\mathbf{J}} \log \sum_{\mathbf{x} \in \mathcal{X}} e^{-\beta \text{tr}(\mathbf{J} \mathbf{x} \mathbf{x}^\dagger)} &= \lim_{K \rightarrow \infty} \lim_{n \rightarrow 0} \frac{1}{nK} \log \mathbb{E}_{\mathbf{Q}} \exp \left[-K \sum_{a=1}^n \int_0^{\beta \lambda_a(\mathbf{Q})} R_{\mathbf{J}}(-w) dw \right] \\
 &= \min_{\mathbf{Q}: \text{Pr}(\mathbf{Q}) > 0} \lim_{n \rightarrow 0} \frac{1}{n} \log \left[- \sum_{a=1}^n \int_0^{\beta \lambda_a(\mathbf{Q})} R_{\mathbf{J}}(-w) dw \right]
 \end{aligned}$$

Laplace Integration

We find

$$\begin{aligned}
 \lim_{K \rightarrow \infty} \frac{1}{K} \mathbb{E}_{\mathbf{J}} \log \sum_{\mathbf{x} \in \mathcal{X}} e^{-\beta \text{tr}(\mathbf{J} \mathbf{x} \mathbf{x}^\dagger)} &= \lim_{K \rightarrow \infty} \lim_{n \rightarrow 0} \frac{1}{nK} \log \mathbb{E}_{\mathbf{Q}} \exp \left[-K \sum_{a=1}^n \int_0^{\beta \lambda_a(\mathbf{Q})} R_{\mathbf{J}}(-w) dw \right] \\
 &= \min_{\mathbf{Q}: \text{Pr}(\mathbf{Q}) > 0} \lim_{n \rightarrow 0} \frac{1}{n} \log \left[- \sum_{a=1}^n \int_0^{\beta \lambda_a(\mathbf{Q})} R_{\mathbf{J}}(-w) dw \right] \\
 &\rightsquigarrow \min_{\mathbf{Q}: \text{Pr}(\mathbf{Q}) > 0} \lim_{n \rightarrow 0} \frac{1}{n} \text{tr} [\mathbf{Q} R_{\mathbf{J}}(-\beta \mathbf{Q})].
 \end{aligned}$$

How to optimize over \mathbf{Q} ?

Replica Symmetric (RS) Ansatz

We assume a certain structure for a matrix Q . The easiest try is

$$Q := \begin{bmatrix} q + \frac{\chi}{\beta} & q & \cdots & q & q \\ q & q + \frac{\chi}{\beta} & \ddots & q & q \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ q & q & \ddots & q + \frac{\chi}{\beta} & q \\ q & q & \cdots & q & q + \frac{\chi}{\beta} \end{bmatrix}$$

with some parameters q and χ .

This is a critical step. Sometimes, the structure of Q is more complicated.

RS Solution

Let $P(s)$ denote the limit of the empirical distribution of the information symbols s_1, s_2, \dots, s_K as $K \rightarrow \infty$. Let q and χ be the simultaneous solutions to

$$q = \iint_{x \in \mathcal{B}_s} \operatorname{argmin}^2 \left| z \sqrt{2qR'(-\chi)} - 2xR(-\chi) \right| Dz dP(s)$$

$$\chi = \frac{1}{\sqrt{2qR'(-\chi)}} \iint_{x \in \mathcal{B}_s} \operatorname{argmin} \left| z \sqrt{2qR'(-\chi)} - 2xR(-\chi) \right| z^* Dz dP(s)$$

where $Dz = \exp(-z^2/2)dz/\sqrt{2\pi}$, $R(\cdot)$ is the R-transform of the limiting eigenvalue spectrum of \mathbf{J} , and $0 < \chi < \infty$.

Then, replica symmetry (RS) implies

$$\frac{1}{K} \min_{\mathbf{x} \in \mathcal{X}} \mathbf{x}^\dagger \mathbf{J} \mathbf{x} \rightarrow q \frac{\partial}{\partial \chi} \chi R(-\chi)$$

as $K \rightarrow \infty$.

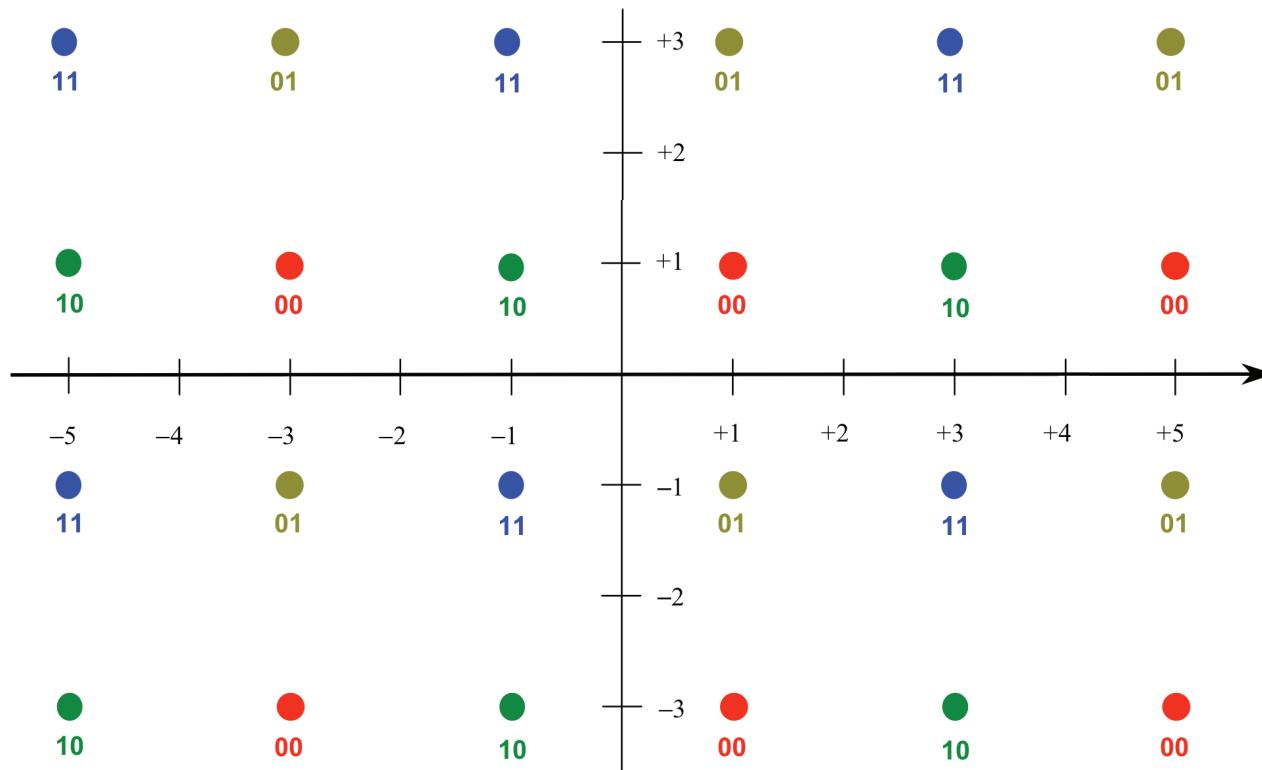
Some R -Transforms

$$\mathbf{I} : R(w) = 1$$

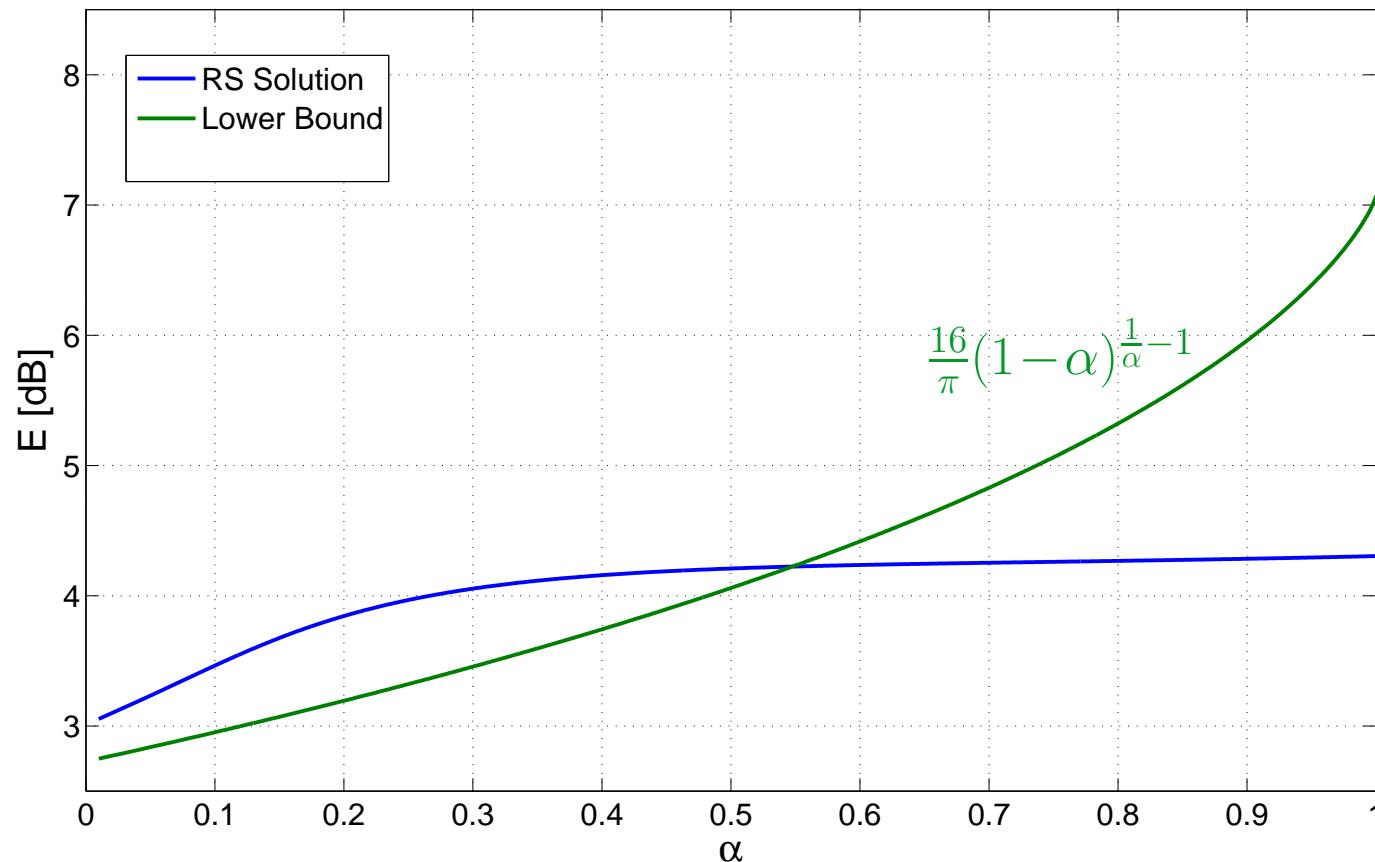
$$\mathbf{H} \mathbf{H}^\dagger : R(w) = \frac{1}{1 - \alpha w} \quad \text{Marchenko-Pastur (MP) law}$$

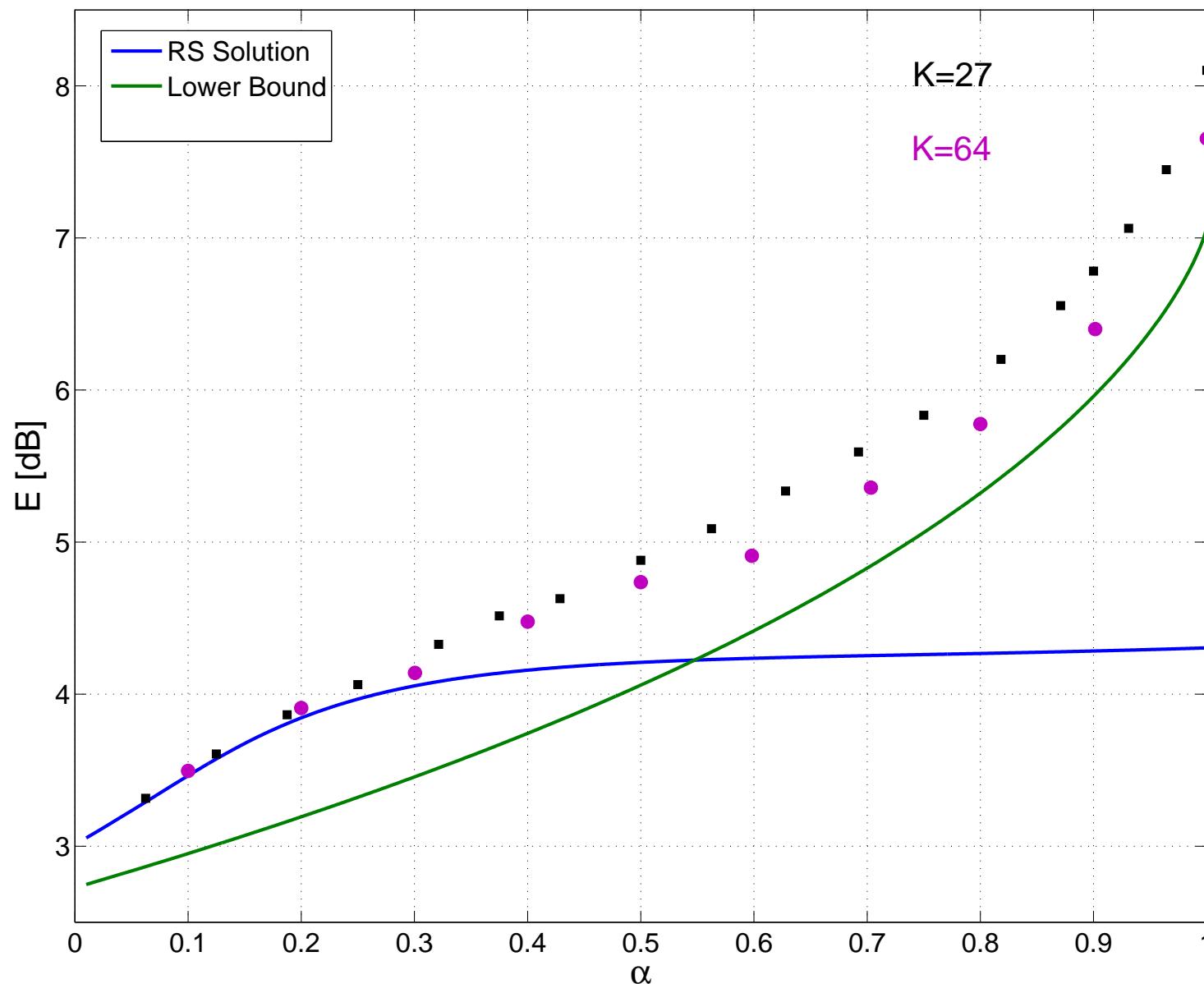
$$(\mathbf{H} \mathbf{H}^\dagger)^{-1} : R(w) = \frac{1 - \alpha - \sqrt{(1 - \alpha)^2 - 4\alpha w}}{2\alpha w} \quad \text{inv. MP}$$

Odd Integer Quadrature Lattice



Complex Lattice Precoding





1-Step Replica Symmetry Breaking

$$Q := \underbrace{\begin{bmatrix} q + p + \frac{\chi}{\beta} & q + p & q & q & \cdots & q & q \\ q + p & q + p + \frac{\chi}{\beta} & q & q & \cdots & q & q \\ q & q & q + p + \frac{\chi}{\beta} & q + p & \ddots & q & q \\ q & q & q + p & q + p + \frac{\chi}{\beta} & & \vdots & \vdots \\ \vdots & \vdots & \ddots & & \ddots & q & q \\ q & q & q & \cdots & q & q + p + \frac{\chi}{\beta} & q + p \\ q & q & q & \cdots & q & q + p & q + p + \frac{\chi}{\beta} \end{bmatrix}}_{\frac{\mu}{\beta} \text{ columns}}$$

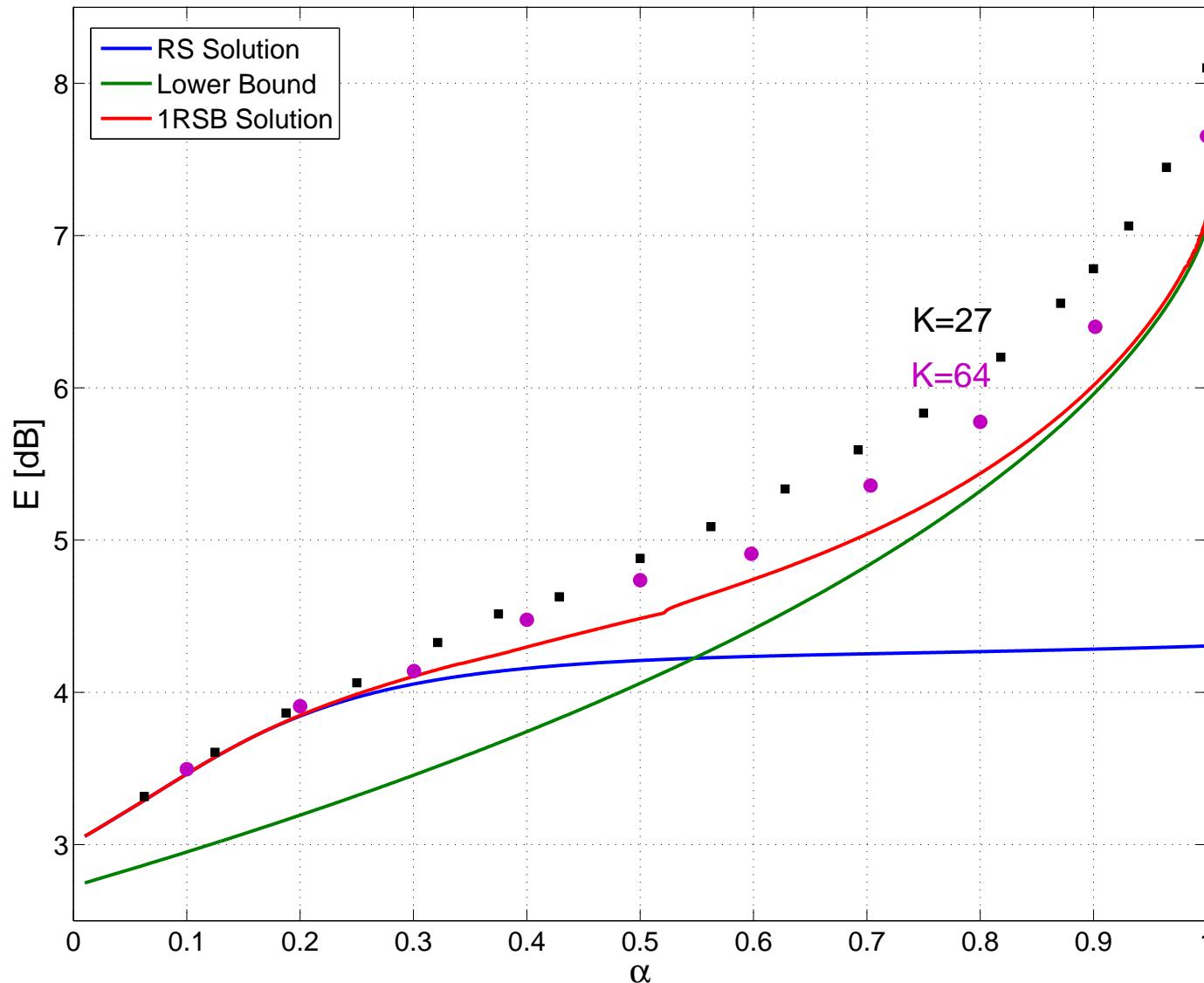
with the macroscopic parameters q, p and χ and the blocksize $\frac{\mu}{\beta}$.

1-Step Replica Symmetry Breaking

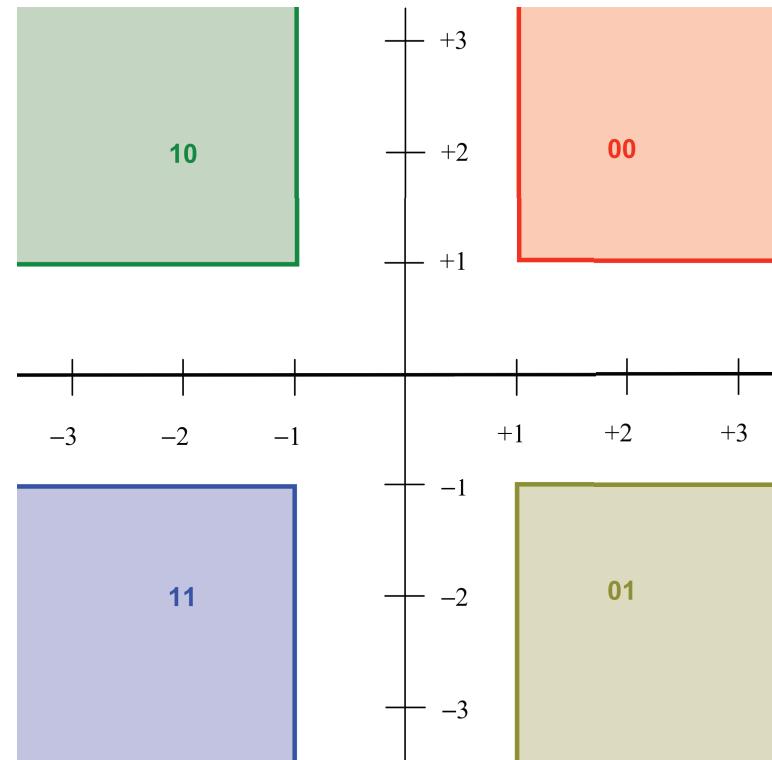
$$E = \left(q + p + \frac{\chi}{\mu} \right) R(-\chi - \mu p) - \frac{\chi}{\mu} R(-\chi) - q(\mu p + \chi) R'(-\chi - \mu p)$$

The macroscopic parameters q, p, χ and μ are given by 4 coupled non-linear equations.

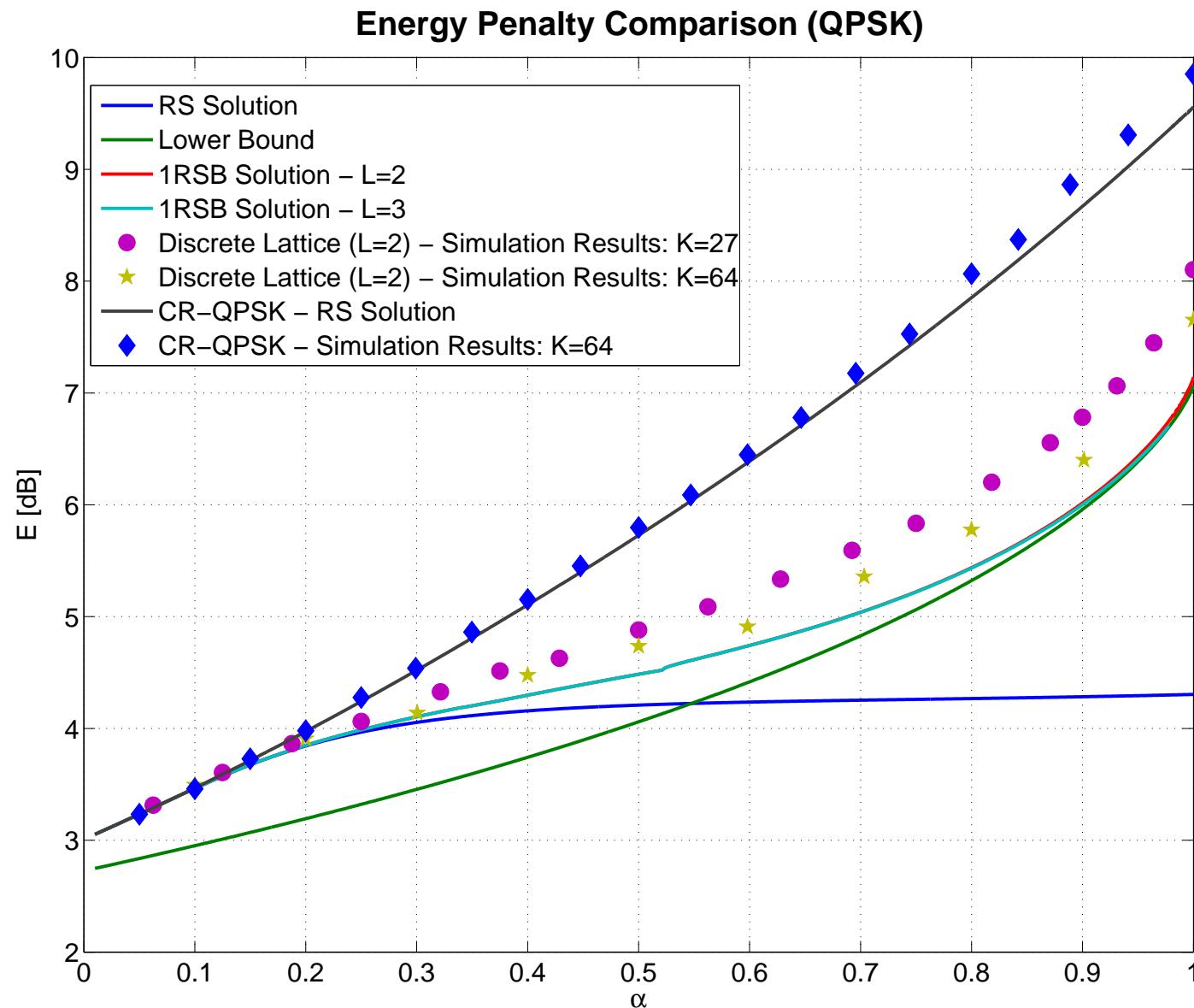
Solving those equations numerically is a tedious and tricky task.



Complex Convex Relaxation



... allows for convex programming (and is replica symmetric).



Inverting Singular Channels

What happens if the MP-law has a mass point at zero ($K > N$)?

Can we precode without interference?

Inverting Singular Channels

What happens if the MP-law has a mass point at zero ($K > N$)?

Can we precode without interference?

The precoder produces

$$\lim_{\epsilon \rightarrow 0} \operatorname{argmin}_{\mathbf{x} \in \mathcal{X}} \frac{\mathbf{x}^\dagger (\mathbf{H} \mathbf{H}^\dagger + \epsilon \mathbf{I})^{-1} \mathbf{x}}{K}$$

The received signal becomes

$$\mathbf{y} = \lim_{\epsilon \rightarrow 0} \mathbf{H} \mathbf{H}^\dagger (\mathbf{H} \mathbf{H}^\dagger + \epsilon \mathbf{I})^{-1} \mathbf{x} + \mathbf{n}.$$

Inverting Singular Channels

What happens if the MP-law has a mass point at zero ($K > N$)?

Can we precode without interference?

The precoder produces

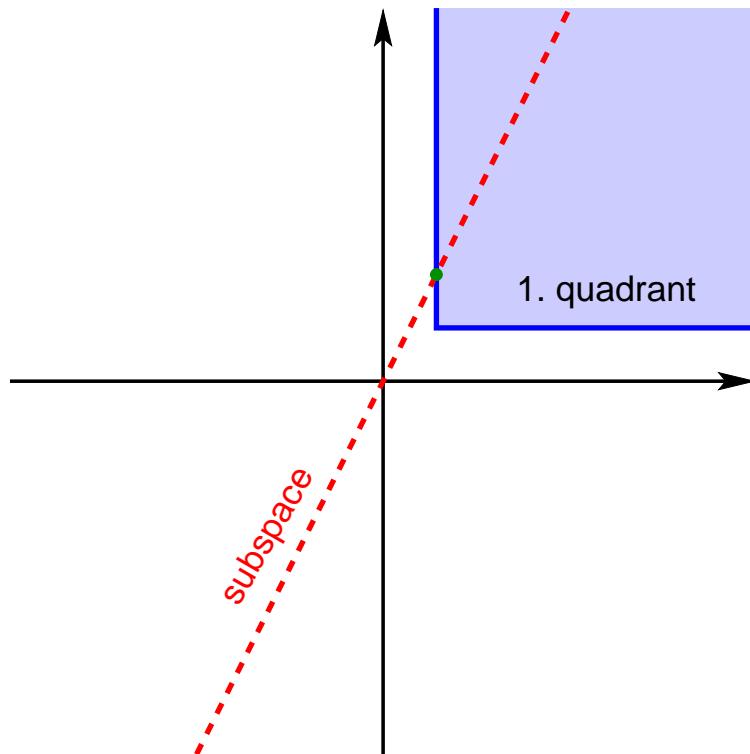
$$\lim_{\epsilon \rightarrow 0} \operatorname{argmin}_{\mathbf{x} \in \mathcal{X}} \frac{\mathbf{x}^\dagger (\mathbf{H} \mathbf{H}^\dagger + \epsilon \mathbf{I})^{-1} \mathbf{x}}{K}$$

The received signal becomes

$$\mathbf{y} = \lim_{\epsilon \rightarrow 0} \mathbf{H} \mathbf{H}^\dagger (\mathbf{H} \mathbf{H}^\dagger + \epsilon \mathbf{I})^{-1} \mathbf{x} + \mathbf{n}.$$

If the energy is finite, there is no interference.

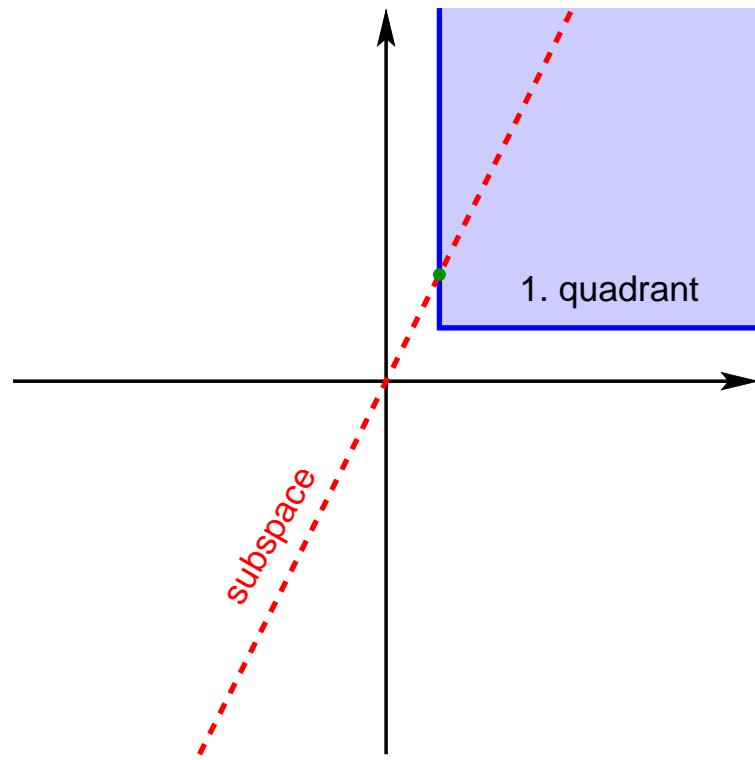
Overloaded Convex Precoding



The probability that a random N dimensional subspace in K real dimensions intersects the 1. K -tant is

$$P(K, N) = 2^{1-K} \sum_{\ell=0}^{N-1} \binom{K-1}{\ell}$$

Overloaded Convex Precoding



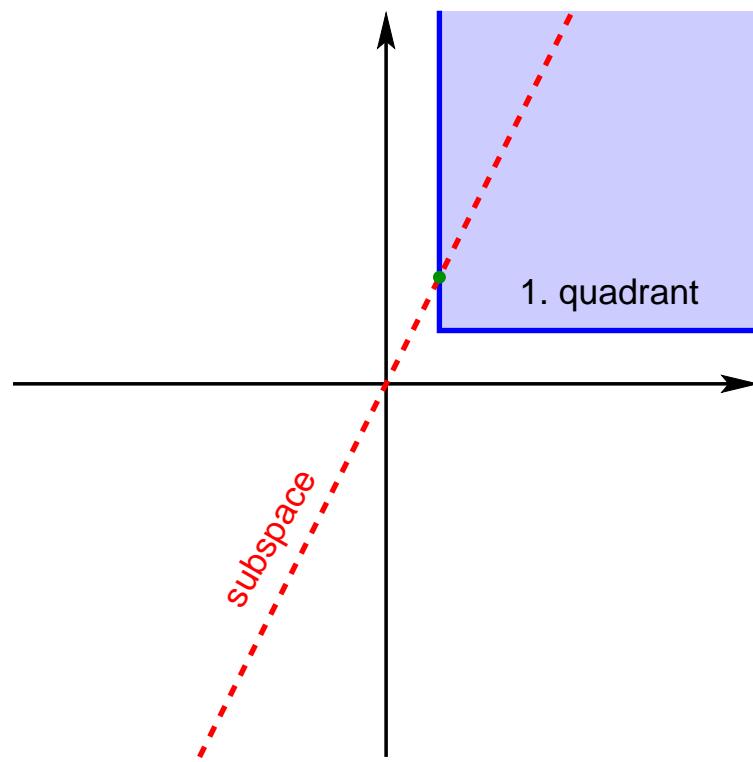
The probability that a random N dimensional subspace in K real dimensions intersects the 1. K -tant is

$$P(K, N) = 2^{1-K} \sum_{\ell=0}^{N-1} \binom{K-1}{\ell}$$

As K, N to infinity, we get

$$P(K, N) = \begin{cases} 1 & K < 2N \\ 1/2 & K = 2N \\ 0 & K > 2N \end{cases}$$

Overloaded Convex Precoding



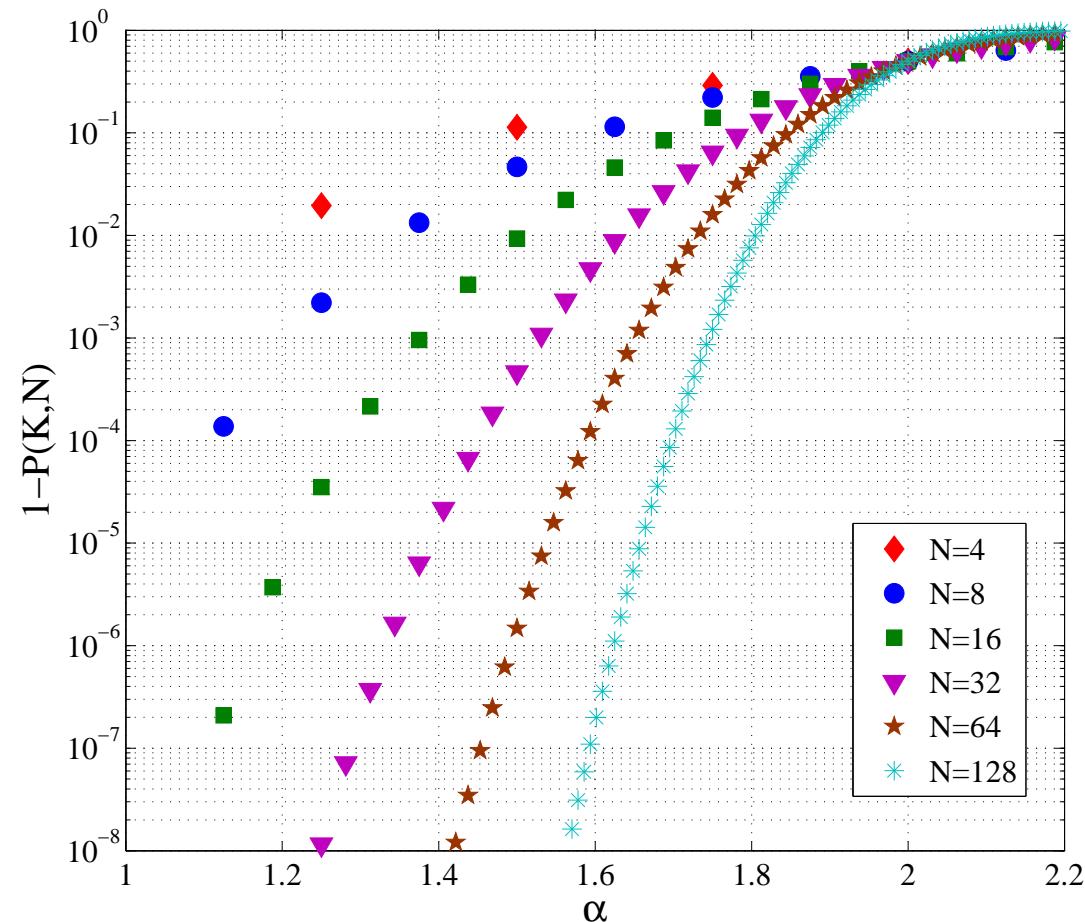
The probability that a random N dimensional subspace in K complex dimensions intersects the 1. K -tant is

$$P(K, N) = 2^{1-2K} \sum_{\ell=0}^{2N-1} \binom{2K-1}{\ell}$$

As K, N to infinity, we get

$$P(K, N) = \begin{cases} 1 & K < 2N \\ 1/2 & K = 2N \\ 0 & K > 2N \end{cases}$$

Overloaded Convex Precoding



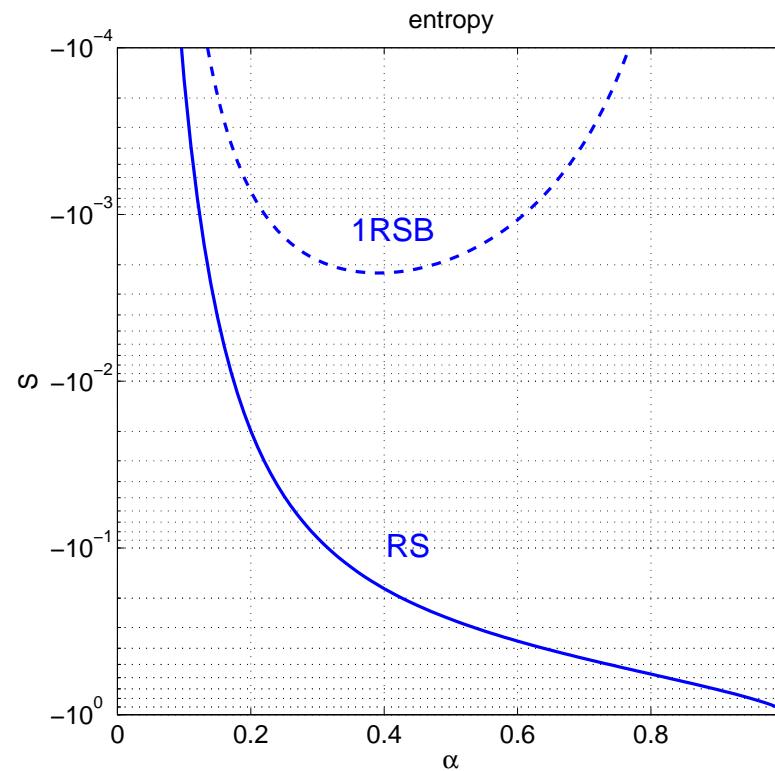
Wanted

$$\lim_{K \rightarrow \infty} \frac{1}{K} \log \mathbb{E}_{\mathbf{A}, \mathbf{B}} e^{-K \operatorname{tr} \mathbf{A} \mathbf{P} \mathbf{B} \mathbf{P}} = f \{R_{\mathbf{A}}(\cdot), R_{\mathbf{B}}(\cdot), \dots\}$$

... or other more complicated exponents.

Negative Entropy

$$S = \chi R(-\chi) - \int_0^\chi R(-w)dw$$



The closer the entropy is to zero, the better the RSB approximation.