

Minimization of Quadratic Forms in Wireless Communications

Ralf R. Müller

Department of Electronics & Telecommunications
Norwegian University of Science & Technology, Trondheim, Norway

mueller@iet.ntnu.no

Dongning Guo

Department of Electrical Engineering & Computer Science
Northwestern University, Evanston, IL, U.S.A.

dguo@northwestern.edu

Aris L. Moustakas

Physics Department
National & Capodistrian University of Athens, Greece

arism@phys.uoa.gr

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:

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

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:

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

for Wishart matrix $\longrightarrow [1 - \sqrt{\alpha}]_+^2$

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:

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

for Wishart matrix $\longrightarrow [1 - \sqrt{\alpha}]_+^2$

Example 2:

$$\mathcal{X} = \{x : x^2 = 1\}^K \implies ???$$

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:

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

for Wishart matrix $\longrightarrow [1 - \sqrt{\alpha}]_+^2$

Example 2:

$$\mathcal{X} = \{x : x^2 = 1\}^K \implies ???$$

Example 3:

$$\mathcal{X} = \{x : |x|^2 = 1\}^K \implies ???$$

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:

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

for Wishart matrix $\longrightarrow [1 - \sqrt{\alpha}]_+^2$

Example 2:

$$\mathcal{X} = \{x : x^2 = 1\}^K \implies ???$$

for Wishart matrix $\longrightarrow \approx \left[1 - \frac{\alpha}{\sqrt{\pi}}\right]_+^2$

Example 3:

$$\mathcal{X} = \{x : |x|^2 = 1\}^K \implies ???$$

for Wishart matrix $\longrightarrow \approx \left[1 - \frac{\sqrt{\pi\alpha}}{2}\right]_+^2$

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:

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

for Wigner matrix $\longrightarrow -2$

Example 2:

$$\mathcal{X} = \{x : x^2 = 1\}^K \implies ???$$

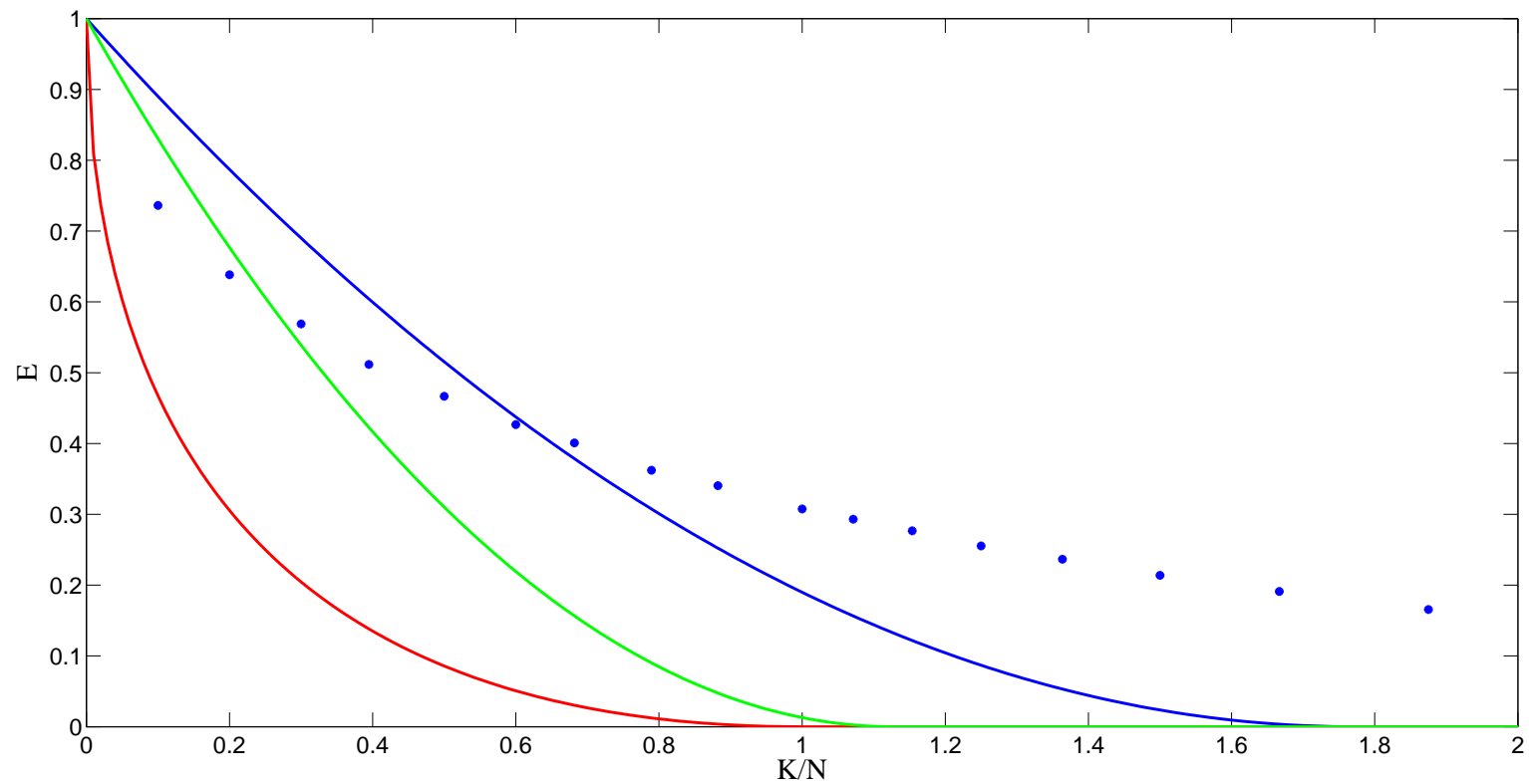
for Wigner matrix $\longrightarrow \approx -\frac{2}{\sqrt{\pi}}$

Example 3:

$$\mathcal{X} = \{x : |x|^2 = 1\}^K \implies ???$$

for Wigner matrix $\longrightarrow \approx -\sqrt{\pi}$

Wishart Matrix



$$\mathbf{x}^\dagger \mathbf{x} = K \quad |x|^2 = 1 \quad x^2 = 1 : K = 15, \infty$$

The Gaussian Vector Channel

Let the received vector be given by

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

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} is the data to be sent.

Then,

$$\mathbf{r} = \mathbf{x} + \mathbf{n}.$$

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 $1/N$.

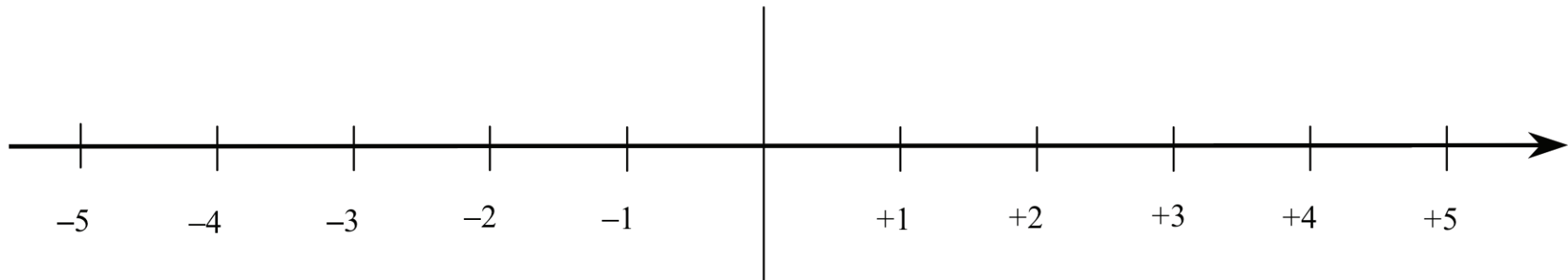
Then, for fixed aspect ratio α

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

with probability 1.

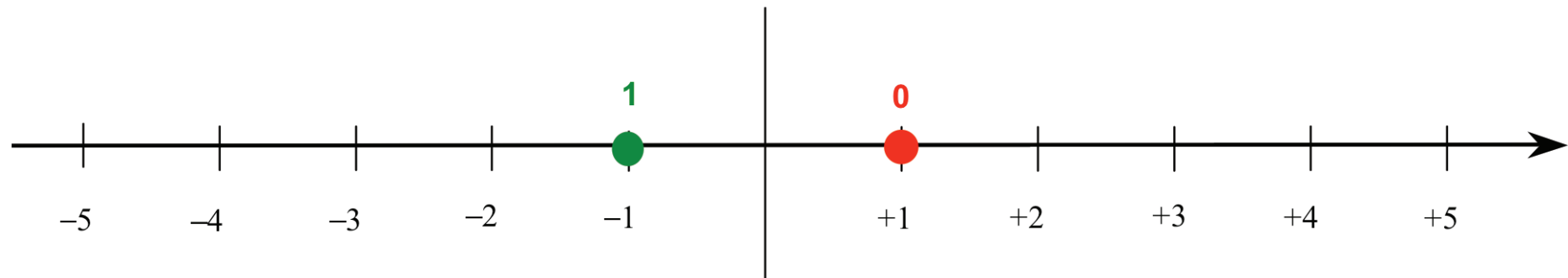
Tomlinson-Harashima Precoding

Tomlinson '71, Harashima & Miyakawa '72



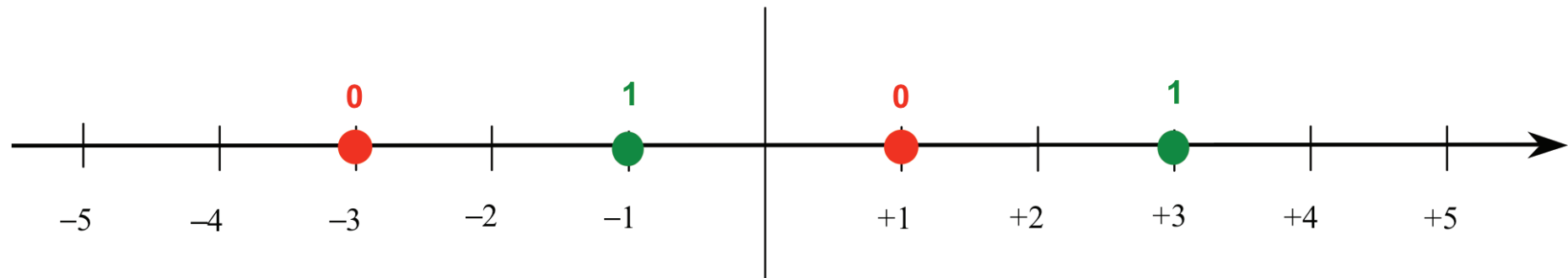
Tomlinson-Harashima Precoding

Tomlinson '71, Harashima & Miyakawa '72



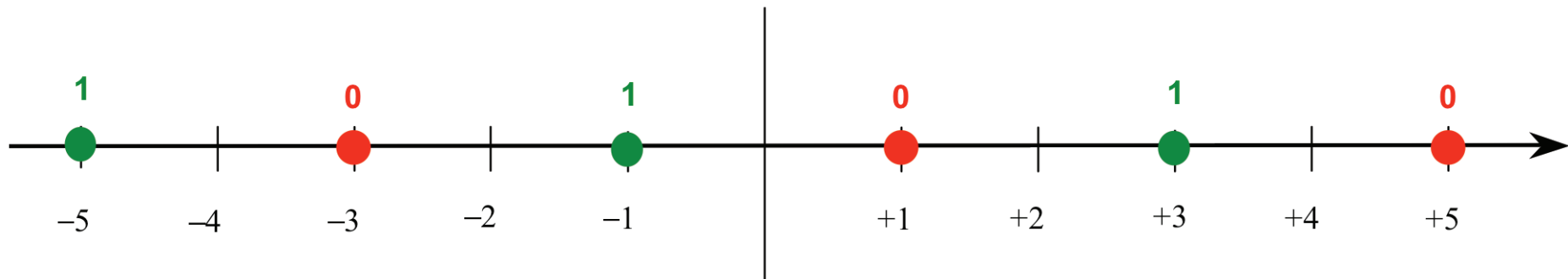
Tomlinson-Harashima Precoding

Tomlinson '71; Harashima & Miyakawa '72



Tomlinson-Harashima 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.

Generalized TH Precoding

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

Let $(s_1, s_2, s_3, \dots, s_K)$ 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

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

Zero Temperature Formulation

Quadratic programming is the problem of finding the zero temperature limit (ground state energy) of a quadratic Hamiltonian.

The transmitted power is written as a zero temperature limit

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

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

Zero Temperature Formulation

Quadratic programming is the problem of finding the zero temperature limit (ground state energy) of a quadratic Hamiltonian.

The transmitted power is written as a zero temperature limit

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

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

Free Fourier Transform

We want

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

Free Fourier Transform

We want

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

We know

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

Free Fourier Transform

We want

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

We know

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

We would like to exchange expectation and logarithm:

$$\mathbb{E}_{\mathbf{X}} \log X = \lim_{n \rightarrow 0} \frac{1}{n} \log \mathbb{E}_{\mathbf{X}} X^n.$$

Replica Continuity

We want

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

Replica Continuity

We want

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

Replica Continuity

We want

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

Replica Continuity

We want

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

with

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

Replica Symmetry

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

with some macroscopic parameters q and χ .

This is the most critical step. In general, the structure of \mathbf{Q} is more complicated. Generalizations are called replica symmetry breaking (RSB).

Main Result

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

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

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

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 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

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

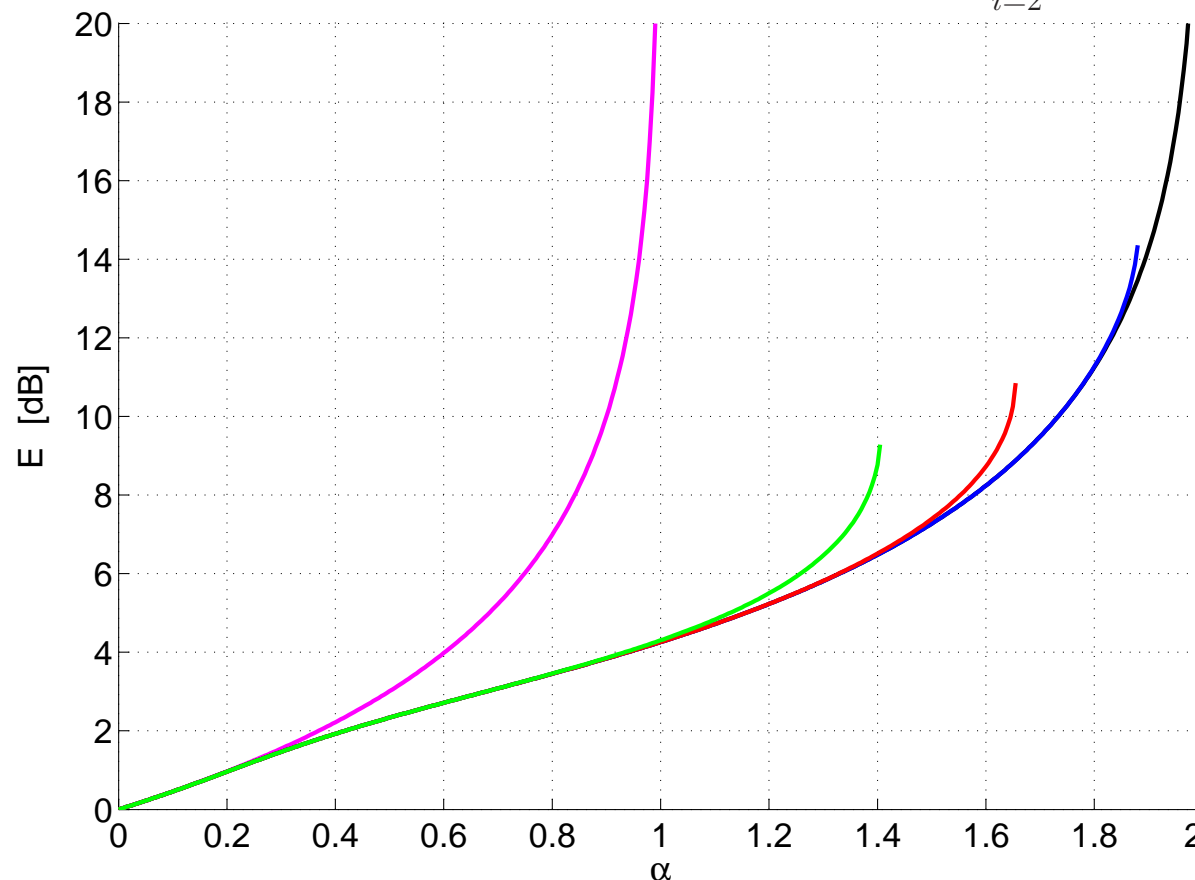
$$\mathbf{H}\mathbf{H}^\dagger : R(w) = \frac{1}{1 - \alpha w} \quad \text{Marchenko-Pastur (MP) law}$$

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

$$\mathbf{U} + \mathbf{U}^\dagger : R(w) = \frac{-1 + \sqrt{1 + 4w^2}}{w}$$

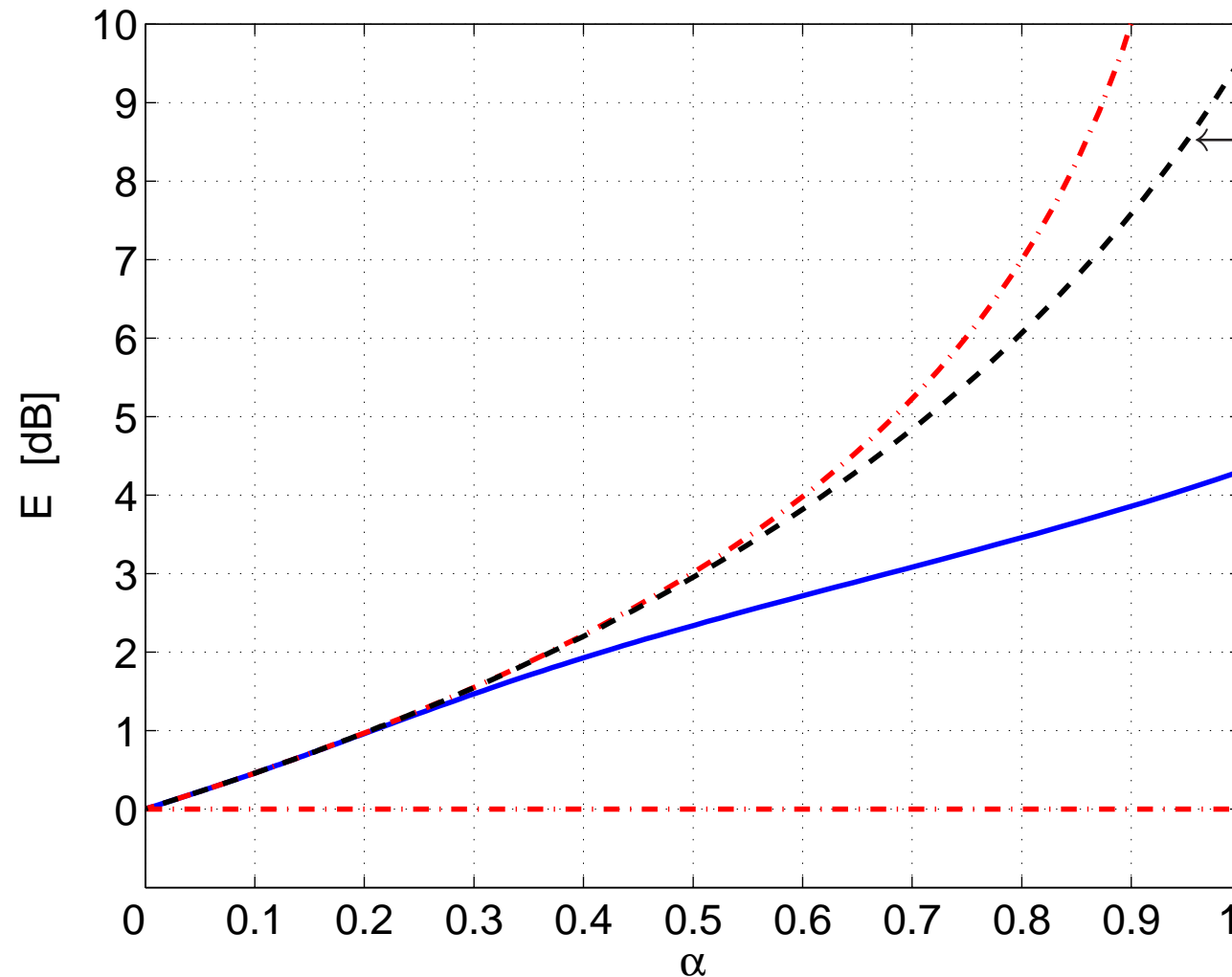
Inv. MP with Odd Integer Lattice (TH Precoding)

Let $\mathbf{J} = (\mathbf{H}\mathbf{H}^\dagger)^{-1}$ and $\chi < \infty$:
$$E = \frac{c_1^2 + \sum_{i=2}^L (c_i^2 - c_{i-1}^2) Q\left(\frac{c_i + c_{i-1}}{\sqrt{2\alpha E}}\right)}{1 - \alpha + \sqrt{\frac{\alpha}{\pi E}} \sum_{i=2}^L (c_i - c_{i-1}) \exp\left(-\frac{(c_i + c_{i-1})^2}{4\alpha E}\right)}$$



$$L = 1, 2, 3, 6, 100$$

Convex Relaxation



conv. relaxation

$$\mathcal{B}_0 = [+1; +\infty)$$

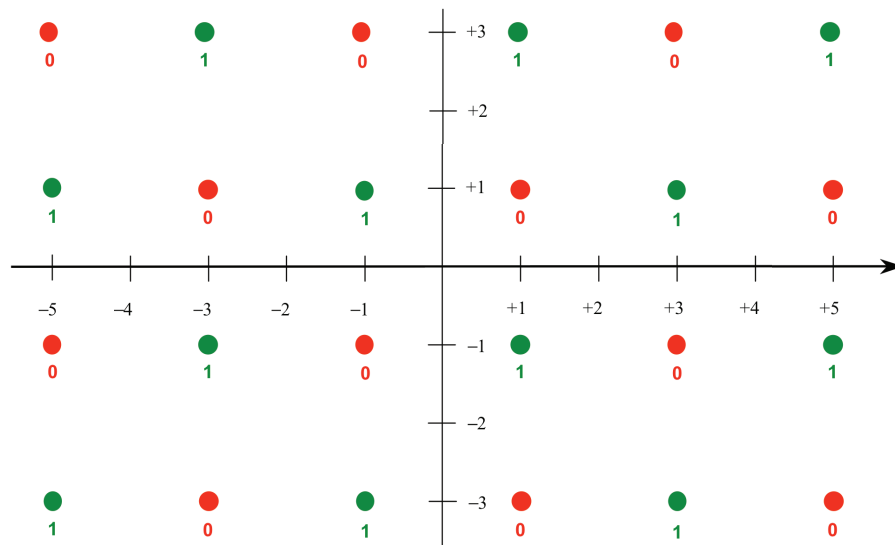
$$\mathcal{B}_1 = (-\infty; -1]$$

Both sets are convex.

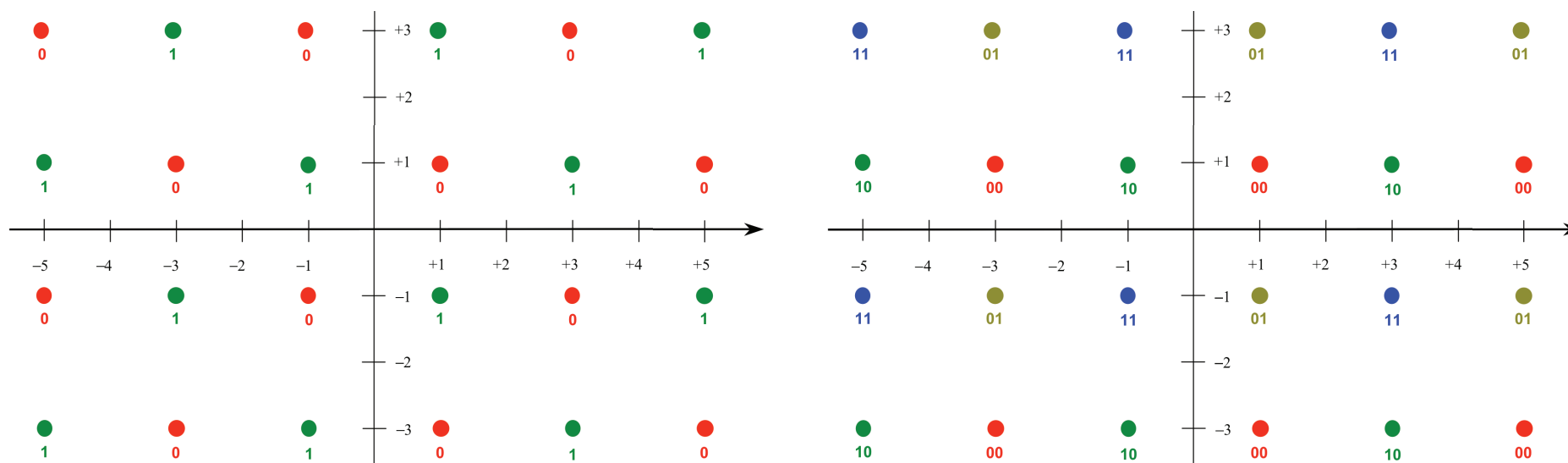


Convex optimization,
but small gains.

Odd Integer Quadrature Lattice

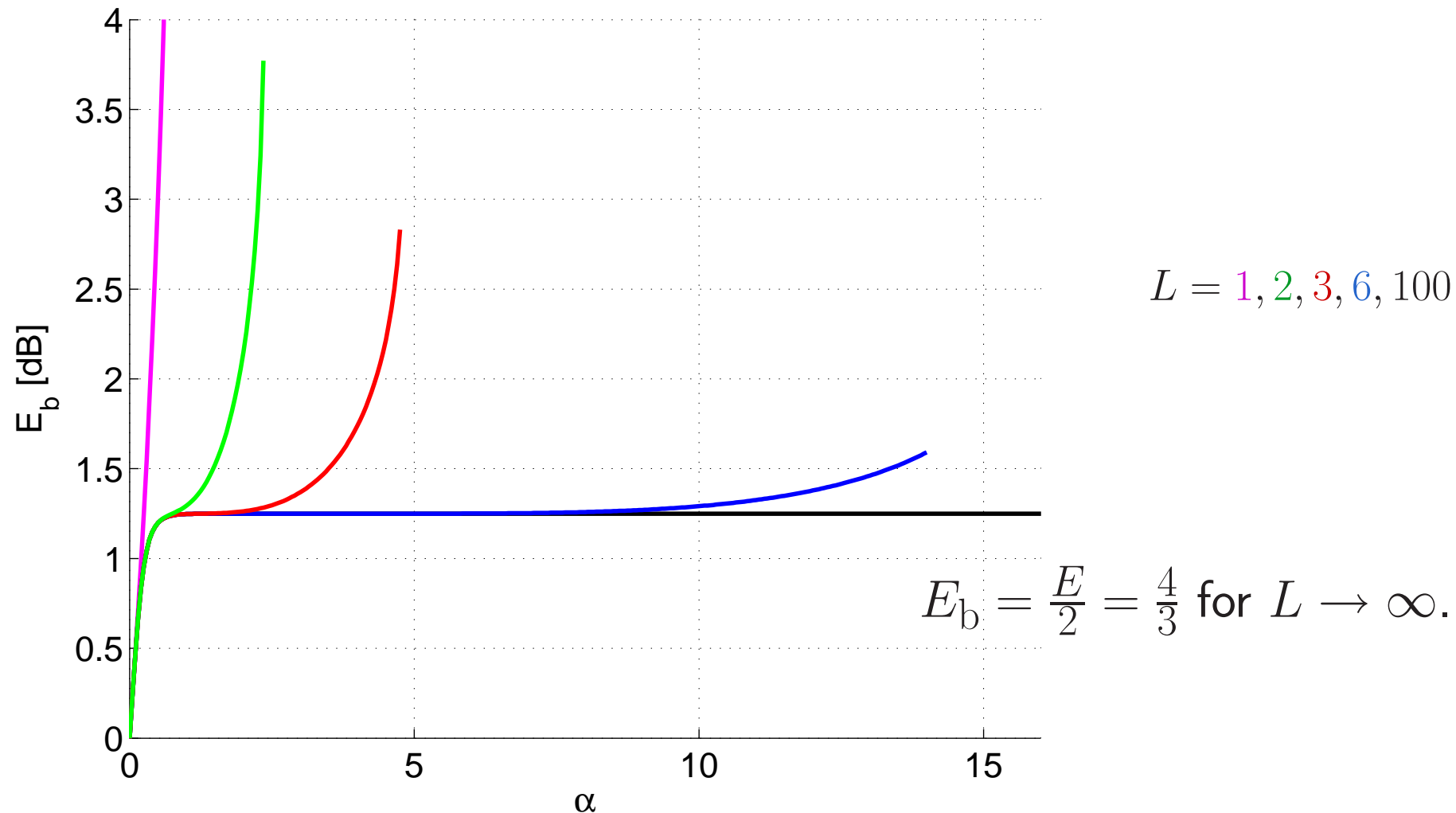


Odd Integer Quadrature Lattice

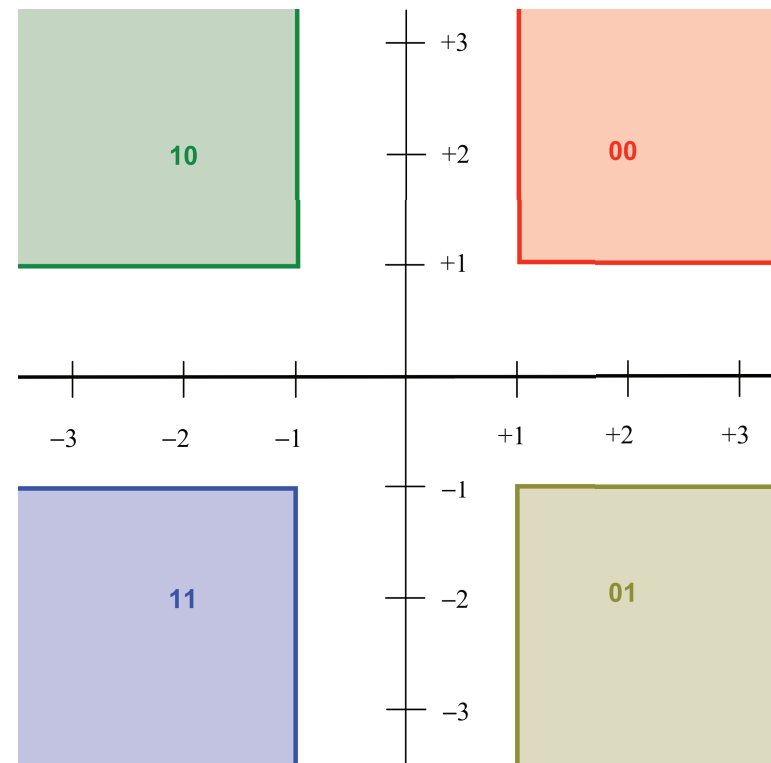


Same energy per bit $E_b = \frac{E}{\log_2 |\mathcal{S}|}$ in both cases.

Complex TH Precoding

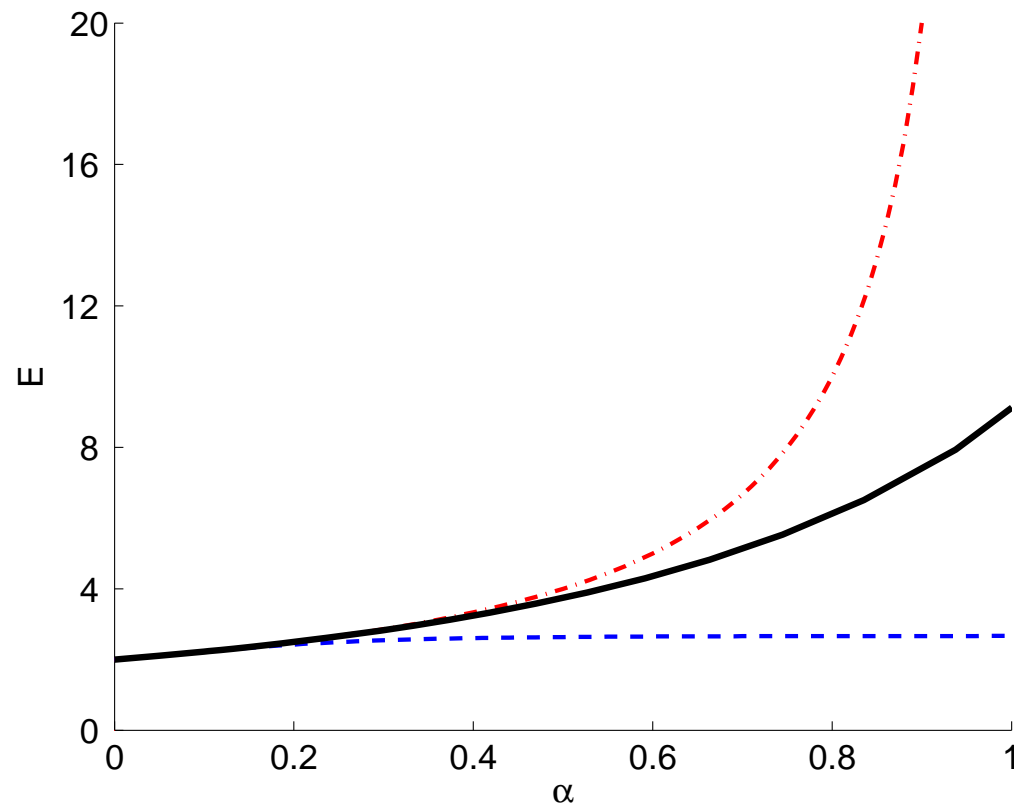


Complex Convex Relaxation



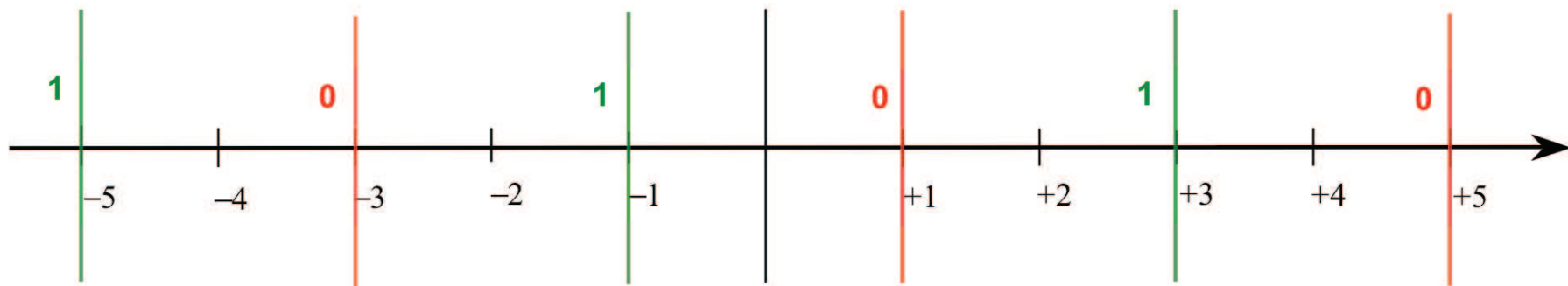
... allows for convex programming.

Complex Convex Relaxation (cont'd)



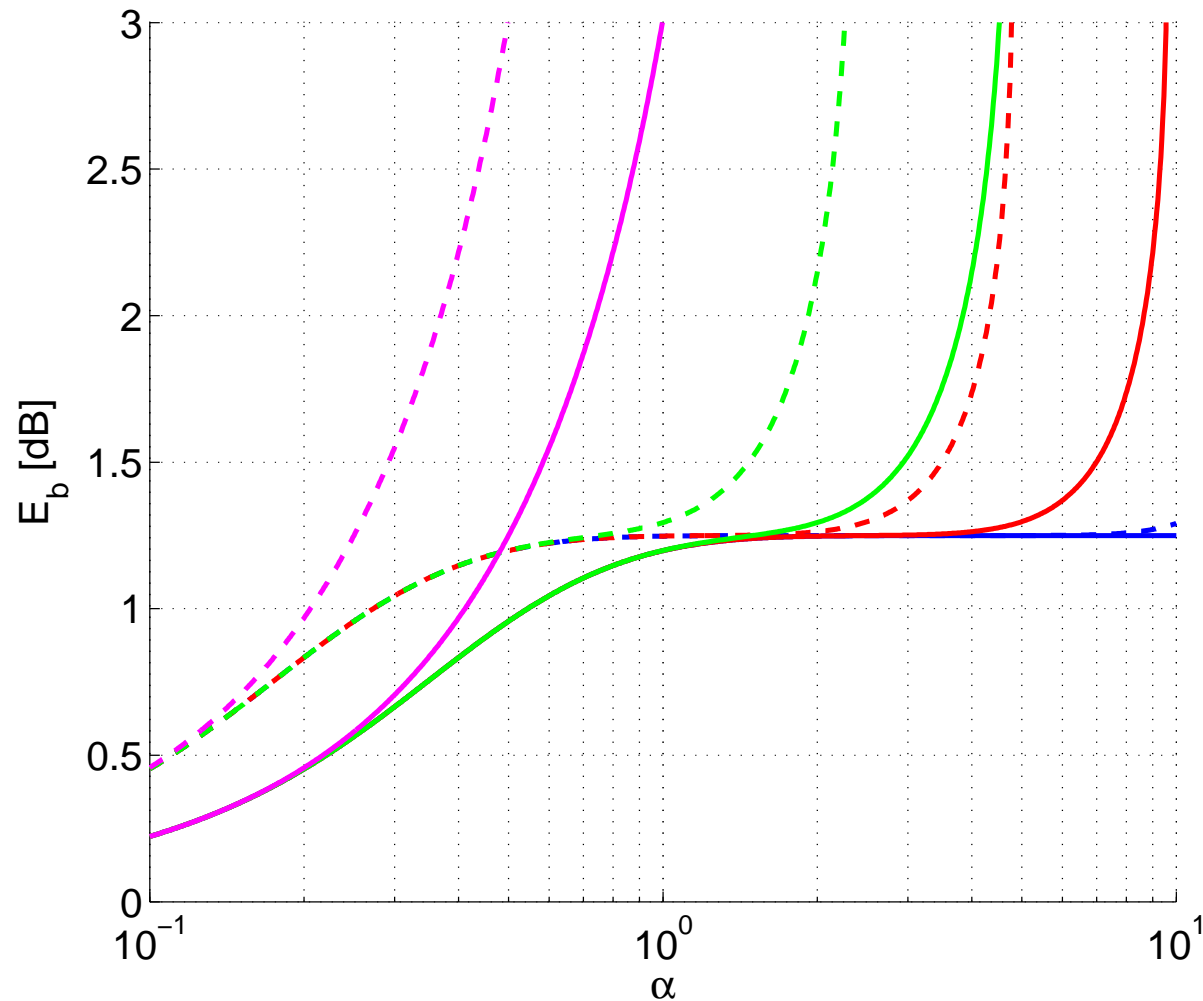
... achieves part of the gain of TH precoding.

Complex Semi-Discrete Set



The imaginary part is purely used to reduce transmit energy.

Complex Semi-Discrete Set (solid lines)



$$L = 1, 2, 3, 6$$

$$E_b = \frac{4}{3} \text{ for } L \rightarrow \infty.$$

Convex opt. for $L = 1$.

A Fake Gain

The inverse MP kernel has the following property:

Let \mathbf{H} and \mathbf{H}' be random matrices of size $K \times N$ and $K' \times N$ respectively, with $K' > K$ and with i.i.d. entries of zero mean and variance $1/N$. Then,

$$\min_{\mathbf{x} \in \mathcal{X}} \frac{\mathbf{x}^\dagger (\mathbf{H}\mathbf{H}^\dagger)^{-1} \mathbf{x}}{K} - \min_{\mathbf{x} \in \mathcal{X} \times \mathbb{C}^{K'-K}} \frac{\mathbf{x}^\dagger (\mathbf{H}'\mathbf{H}'^\dagger)^{-1} \mathbf{x}}{K} \longrightarrow 0.$$

The redundant symbols serve no purpose.

Wanted

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

Rigorous or Hand-Waving

Discovering Antimatter

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

Can we precode without interference?

Discovering Antimatter

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

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

Discovering Antimatter

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

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

If the energy is finite, there is no interference.