MMSE Interference in Gaussian Channels[1]

Shlomo Shamai

Department of Electrical Engineering Technion - Israel Institute of Technology

2012 Information Theory and Applications Workshop
5-10 February, San Diego

[1] The talk is based on recent studies done jointly with Ronit Bustin

The Framework

The Scalar Additive Gaussian Channel

Our framework is the scalar additive Gaussian channel:

$$Y = \sqrt{\mathsf{snr}} X + N$$

where $N \sim \mathcal{N}(0,1)$. Through which we transmit length n codewords.

Constraint

We limit our investigation to power constrained codes:

$$\forall \mathbf{x} \in \mathsf{C}_n \quad \frac{1}{n}\sum_{i=1}^{n} \mathbf{x}_i^2 \leq 1$$

The Framework

The Scalar Additive Gaussian Channel

Our framework is the scalar additive Gaussian channel:

$$Y = \sqrt{\mathsf{snr}} X + N$$

where $N \sim \mathcal{N}(0,1)$. Through which we transmit length n codewords.

Constraint

We limit our investigation to power constrained codes:

$$\forall \mathbf{x} \in \mathsf{C}_n \quad \frac{1}{n}\sum_{i=1}^{n} \mathbf{x}_i^2 \leq 1$$

Optimal Point-to-Point Codes

Theorem [Peleg, Sanderovich and Shamai, ETT 2007]

For every capacity achieving code-sequence, C_n, over the Gaussian channel, the mutual information, when $n \to \infty$, is as follows:

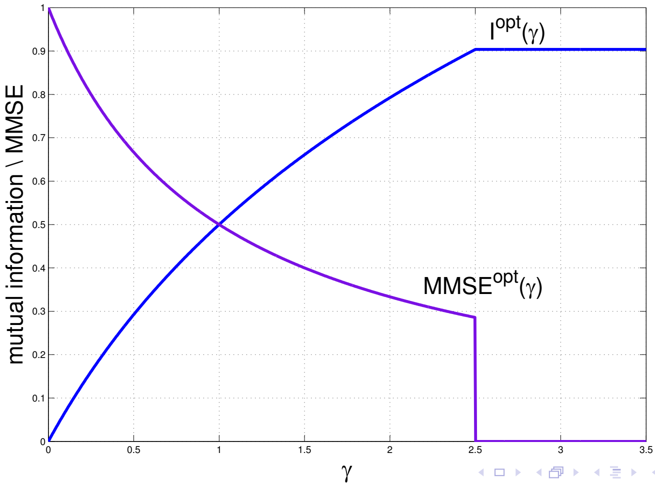
$$I(X; \sqrt{\gamma}X + N) = \begin{cases} \frac{1}{2}\log(1+\gamma), & \gamma \leq \mathsf{snr} \\ \frac{1}{2}\log(1+\mathsf{snr}), & \mathrm{o/w} \end{cases}$$

and the MMSE is:

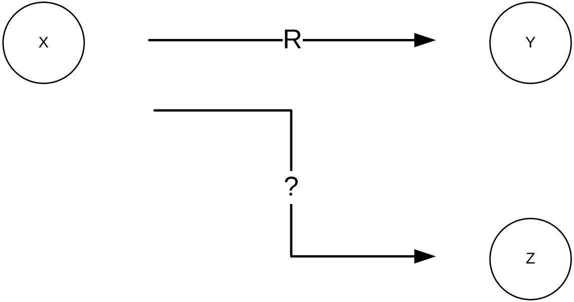
$$\mathsf{MMSE}^{c}(\gamma) = \begin{cases} \frac{1}{1+\gamma}, & \gamma \leq \mathsf{snr} \\ 0, & \mathrm{o/w} \end{cases}$$

Optimal Point-to-Point Codes - Cont.

The mutual information (and MMSE) of optimal point-to-point codes follow the behavior of an i.i.d. Gaussian input up to snr.



What is the effect on an unintended receiver?



Assumption: the unintended receiver, Z, has smaller snr, that is, $\text{snr}_z < \text{snr}$.

How should we measure the effect (disturbance)?

For optimal point-to-point codes both the mutual information and MMSE are completely known.

But what about non-optimal code (that do not attain capacity)?

How to measure the disturbance?

- Bandemer and El Gamal, 2011: measure the disturbance at the unintended receiver using the mutual information at Z. That is, assuming this mutual information is at most R_d what is the maximum possible rate to the intended receiver, Y.
- In this work we measure the disturbance at the unintended receiver using the MMSE of the input, X, at Z. That is, assuming the MMSE is constrained to be at most $\frac{\beta}{1+\beta \mathsf{snr}_z}$ what is the maximum possible rate to the intended receiver, Y.

How to measure the disturbance?

- Bandemer and El Gamal, 2011: measure the disturbance at the unintended receiver using the mutual information at Z. That is, assuming this mutual information is at most R_d what is the maximum possible rate to the intended receiver, Y.
- In this work we measure the disturbance at the unintended receiver using the MMSE of the input, X, at Z. That is, assuming the MMSE is constrained to be at most $\frac{\beta}{1+\beta \mathsf{snr}_z}$ what is the maximum possible rate to the intended receiver, Y.

Definitions

$$I_n(\gamma) = \frac{1}{n} I\left(\mathbf{X}; \mathbf{Y}(\gamma)\right)$$

$$I(\gamma) = \lim_{n\to\infty} \frac{1}{n} I\left(\mathbf{X}; \mathbf{Y}(\gamma)\right)$$

$$\mathrm{MMSE}^{c_n}(\gamma) = \frac{1}{n} \mathrm{Tr}(\mathbf{E}_{\mathbf{X}}(\gamma))$$

$$\mathrm{MMSE}^{c}(\gamma) = \lim_{n\to\infty} \mathrm{MMSE}^{c_n}(\gamma)$$

where $\mathbf{E}_{\mathbf{X}}(\gamma)$ is the MMSE matrix of estimating $\mathbf{X}$ from $\mathbf{Y}(\gamma) = \sqrt{\gamma}\mathbf{X} + \mathbf{N}$.

The I-MMSE approach

1. The I-MMSE relationship [Guo, Shamai and Verdú, IT 2005]

A fundamental relationship between the mutual information and the MMSE in the Gaussian channel:

$$I_n(\gamma) = \frac{1}{2}\int_0^{\mathsf{snr}} \mathsf{MMSE}^{c_n}(\gamma)\mathrm{d}\gamma$$

Taking the limit of $n \to \infty$:

$$I(\gamma) = \frac{1}{2}\int_0^{\mathsf{snr}} \mathsf{MMSE}^{c}(\gamma)\mathrm{d}\gamma$$

2. The “single crossing point” property

The property was originally derived for the scalar case in [Guo, Wu, Shamai and Verdú, IT 2011].
Several MIMO extensions are given in [Bustin, Payaró, Palomar and Shamai <arXiv >].
We require the simplest extension.

Define the following function for an arbitrary random vector $\mathbf{X}$:

$$q_{\mathbf{A}}(\mathbf{X}, \sigma^2, \gamma) = \frac{\sigma^2}{1+\sigma^2\gamma}\mathrm{Tr}\left(\mathbf{A}\right) - \mathrm{Tr}\left(\mathbf{A}\mathbf{E}_{\mathbf{X}}(\gamma)\right)$$

where $\mathbf{A}$ is some $n \times n$ general weighting matrix.

The I-MMSE approach

2. The "single crossing point" property

The property was originally derived for the scalar case in [Guo, Wu, Shamai and Verdú, IT 2011].

Several MIMO extensions are given in [Bustin, Payaró, Palomar and Shamai <arXiv >].

We require the simplest extension.

Define the following function for an arbitrary random vector $\mathbf{X}$:

$$q_{\mathbf{A}}(\mathbf{X}, \sigma^2, \gamma) = \frac{\sigma^2}{1+\sigma^2\gamma} \mathrm{Tr}\left(\mathbf{A}\right) - \mathrm{Tr}\left(\mathbf{A}\mathbf{E}_{\mathbf{X}}(\gamma)\right)$$

where $\mathbf{A}$ is some $n \times n$ general weighting matrix.

The I-MMSE approach - Cont.

Theorem [Bustin, Payaró, Palomar and Shamai <arXiv >]

Let $\mathbf{A} \in \mathbb{S}_+^n$ be a PSD matrix. Then, the function $q_{\mathbf{A}}(\mathbf{X}, \sigma^2, \gamma)$, has no nonnegative-to-negative zero crossings and, at most, a single negative-to-nonnegative zero crossing in the range $\gamma \in [0, \infty)$. Moreover, let $\mathsf{snr}_0 \in [0, \infty)$ be that crossing point. Then,

1. $q_{\mathbf{A}}(\mathbf{X}, \sigma^2, 0) \leq 0$.
2. $q_{\mathbf{A}}(\mathbf{X}, \sigma^2, \gamma)$ is a strictly increasing in $\gamma \in [0, \mathsf{snr}_0)$.
3. $q_{\mathbf{A}}(\mathbf{X}, \sigma^2, \gamma) \geq 0$ for all $\gamma \in [\mathsf{snr}_0, \infty)$.
4. $\lim_{\gamma \to \infty} q_{\mathbf{A}}(\mathbf{X}, \sigma^2, \gamma) = 0$.

In this work we set $\mathbf{A} = \mathbf{I}$, the identity matrix:

$$\frac{1}{n} q_{\mathbf{I}}(\mathbf{X}, \sigma^2, \gamma) = \frac{\sigma^2}{1 + \sigma^2 \gamma} - \frac{1}{n} \mathrm{Tr}\left(\mathbf{A} \mathbf{E}_{\mathbf{X}}(\gamma)\right) = \mathrm{mmse}_G(\gamma) - \mathrm{MMSE}^{c_n}(\gamma)$$

where $\mathrm{mmse}_G(\gamma)$ assumes an independent Gaussian input $\sim \mathcal{N}(0, \sigma^2)$.

The I-MMSE approach - Cont.

Theorem [Bustin, Payaró, Palomar and Shamai <arXiv >]

Let $\mathbf{A} \in \mathbb{S}_+^n$ be a PSD matrix. Then, the function $q_{\mathbf{A}}(\mathbf{X}, \sigma^2, \gamma)$, has no nonnegative-to-negative zero crossings and, at most, a single negative-to-nonnegative zero crossing in the range $\gamma \in [0, \infty)$. Moreover, let $\mathsf{snr}_0 \in [0, \infty)$ be that crossing point. Then,

1. $q_{\mathbf{A}}(\mathbf{X}, \sigma^2, 0) \leq 0$.
2. $q_{\mathbf{A}}(\mathbf{X}, \sigma^2, \gamma)$ is a strictly increasing in $\gamma \in [0, \mathsf{snr}_0)$.
3. $q_{\mathbf{A}}(\mathbf{X}, \sigma^2, \gamma) \geq 0$ for all $\gamma \in [\mathsf{snr}_0, \infty)$.
4. $\lim_{\gamma \to \infty} q_{\mathbf{A}}(\mathbf{X}, \sigma^2, \gamma) = 0$.

In this work we set $\mathbf{A} = \mathbf{I}$, the identity matrix:

$$\frac{1}{n} q_{\mathbf{I}}(\mathbf{X}, \sigma^2, \gamma) = \frac{\sigma^2}{1 + \sigma^2 \gamma} - \frac{1}{n} \mathrm{Tr}\left(\mathbf{A}\mathbf{E}_{\mathbf{X}}(\gamma)\right) = \mathrm{mmse}_G(\gamma) - \mathrm{MMSE}^{c_n}(\gamma)$$

where $\mathrm{mmse}_G(\gamma)$ assumes an independent Gaussian input $\sim \mathcal{N}(0, \sigma^2)$.

Superposition codes

$I(\gamma)$ and $\mathrm{MMSE}^c(\gamma)$ - known exactly [Merhav, Guo and Shamai, IT 2010]:

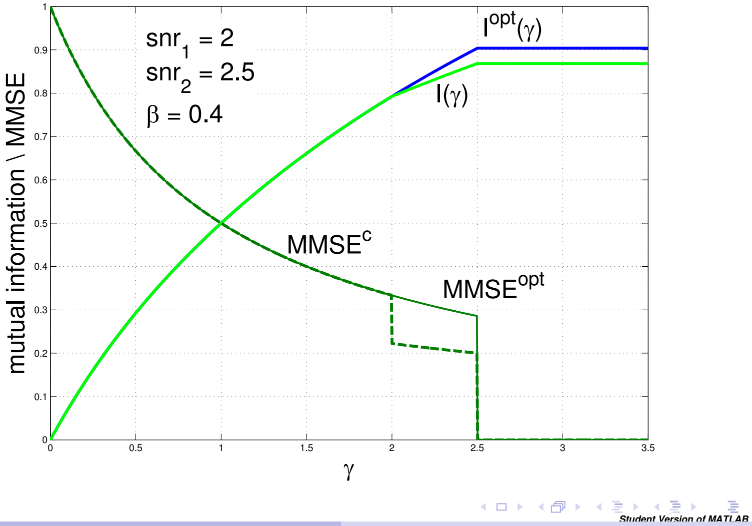
A superposition codebook designed for $(\mathsf{snr}_1, \mathsf{snr}_2)$ with the rate-splitting coefficient $\beta < 1$.

$$I(\gamma) =$$
$$\begin{cases} \frac{1}{2}\log\left(1+\gamma\right), & \text{if} \quad 0 \le \gamma < \mathsf{snr}_1 \\ \frac{1}{2}\log\left(\frac{1+\mathsf{snr}_1}{1+\beta\mathsf{snr}_1}\right) + \frac{1}{2}\log\left(1+\beta\gamma\right), & \text{if} \quad \mathsf{snr}_1 \le \gamma \le \mathsf{snr}_2 \\ \frac{1}{2}\log\left(\frac{1+\mathsf{snr}_1}{1+\beta\mathsf{snr}_1}\right) + \frac{1}{2}\log\left(1+\beta\mathsf{snr}_2\right), & \text{if} \quad \mathsf{snr}_2 < \gamma \end{cases}$$

$$\mathrm{MMSE}^c(\gamma) = \begin{cases} \frac{1}{1+\gamma}, & 0 \le \gamma < \mathsf{snr}_1 \\ \frac{\beta}{1+\beta\gamma}, & \mathsf{snr}_1 \le \gamma \le \mathsf{snr}_2 \\ 0, & \mathsf{snr}_2 < \gamma \end{cases}$$

Superposition codes - Cont.

Example:



Main Result

Theorem 1

Assuming $\mathsf{snr}_1 < \mathsf{snr}_2$ the solution of the following optimization problem,

$$\begin{aligned} \max \quad & I(\mathsf{snr}_2) \\ \text{s.t.} \quad & \mathsf{MMSE}^c(\mathsf{snr}_1) \leq \frac{\beta}{1+\beta\mathsf{snr}_1} \end{aligned}$$

for some $\beta \in [0,1]$, is the following

$$I(\mathsf{snr}_2) = \frac{1}{2}\log\left(1+\beta\mathsf{snr}_2\right) + \frac{1}{2}\log\left(\frac{1+\mathsf{snr}_1}{1+\beta\mathsf{snr}_1}\right)$$

and is attainable when using the optimal Gaussian superposition codebook designed for $(\mathsf{snr}_1, \mathsf{snr}_2)$ with a rate-splitting coefficient β.

Proof Sketch

- Optimal Gaussian superposition codebook comply with the above MMSE constraint and attain the maximum rate.
- We need a tight upper bound on the rate.
- We prove an equivalent claim: assume a code of rate $\mathrm{R}_c = \frac{1}{2}\log(1+\alpha\mathsf{snr}_2)$, designed for reliable transmission at snr_2, lower bound $\mathrm{MMSE}^c(\gamma)$. Then specify for $\gamma = \mathsf{snr}_1$.

$\alpha\mathsf{snr}_2 \le \gamma \le 1$

The lower bound is trivially zero using the optimal Gaussian codebook designed for $\alpha\mathsf{snr}_2$. This is equivalent to setting $\beta = 0$.

$\gamma \le \alpha\mathsf{snr}_2$

$$I(\mathsf{snr}_2) - I(\gamma) \ge I(\mathsf{snr}_2) - \frac{1}{2}\log(1+\gamma) = \mathrm{R}_c - \frac{1}{2}\log(1+\gamma)$$

Proof Sketch

- Optimal Gaussian superposition codebook comply with the above MMSE constraint and attain the maximum rate.
- We need a tight upper bound on the rate.
- We prove an equivalent claim: assume a code of rate $\mathsf{R}_c = \frac{1}{2}\log(1+\alpha\mathsf{snr}_2)$, designed for reliable transmission at snr_2, lower bound $\mathsf{MMSE}^c(\gamma)$. Then specify for $\gamma = \mathsf{snr}_1$.

$\alpha\mathsf{snr}_2 \le \gamma \le 1$

The lower bound is trivially zero using the optimal Gaussian codebook designed for $\alpha\mathsf{snr}_2$. This is equivalent to setting $\beta = 0$.

$\gamma \le \alpha\mathsf{snr}_2$

$$I(\mathsf{snr}_2) - I(\gamma) \ge I(\mathsf{snr}_2) - \frac{1}{2}\log(1+\gamma) = \mathsf{R}_c - \frac{1}{2}\log(1+\gamma)$$

Proof Sketch

- Optimal Gaussian superposition codebook comply with the above MMSE constraint and attain the maximum rate.
- We need a tight upper bound on the rate.
- We prove an equivalent claim: assume a code of rate $\mathrm{R}_c = \frac{1}{2}\log(1 + \alpha \mathrm{snr}_2)$, designed for reliable transmission at snr_2, lower bound $\mathrm{MMSE}^c(\gamma)$. Then specify for $\gamma = \mathrm{snr}_1$.

$\alpha \mathrm{snr}_2 \leq \gamma \leq 1$

The lower bound is trivially zero using the optimal Gaussian codebook designed for $\alpha \mathrm{snr}_2$. This is equivalent to setting $\beta = 0$.

$\gamma \leq \alpha \mathrm{snr}_2$

$$I(\mathrm{snr}_2) - I(\gamma) \geq I(\mathrm{snr}_2) - \frac{1}{2}\log(1+\gamma) = \mathrm{R}_c - \frac{1}{2}\log(1+\gamma)$$

Proof Sketch - Cont.

Using the I-MMSE relationship:

$$\frac{1}{2}\int_{\gamma}^{\mathsf{snr}_2} \mathsf{MMSE}^{\mathsf{c}}(\tau)\, d\tau \geq \frac{1}{2}\log\left(1+\alpha \mathsf{snr}_2\right) - \frac{1}{2}\log\left(1+\gamma\right)$$

Defining d through the following equality:

$$\frac{1}{2}\log\left(1+\alpha \mathsf{snr}_2\right) - \frac{1}{2}\log\left(1+\gamma\right) = \frac{1}{2}\log\left(1+\mathsf{d}\mathsf{snr}_2\right) - \frac{1}{2}\log\left(1+\mathsf{d}\gamma\right)$$

we have:

$$\begin{aligned}\frac{1}{2}\int_{\gamma}^{\mathsf{snr}_2} \mathsf{MMSE}^{\mathsf{c}}(\tau)\, d\tau &\geq \frac{1}{2}\log\left(1+\alpha \mathsf{snr}_2\right) - \frac{1}{2}\log\left(1+\gamma\right) \\ &= \frac{1}{2}\log\left(1+\mathsf{d}\mathsf{snr}_2\right) - \frac{1}{2}\log\left(1+\mathsf{d}\gamma\right) \\ &= \frac{1}{2}\int_{\gamma}^{\mathsf{snr}_2} \mathsf{mmse}_G(\tau)\, d\tau, \quad X_G \sim \mathcal{N}(0,\mathsf{d}), \textit{i.i.d.}\end{aligned}$$

Proof Sketch - Cont.

Using the I-MMSE relationship:

$$\frac{1}{2}\int_{\gamma}^{\mathsf{snr}_2} \mathsf{MMSE}^{\mathsf{c}}(\tau)\,\mathsf{d}\tau \geq \frac{1}{2}\log\left(1+\alpha\mathsf{snr}_2\right) - \frac{1}{2}\log\left(1+\gamma\right)$$

Defining d through the following equality:

$$\frac{1}{2}\log\left(1+\alpha\mathsf{snr}_2\right) - \frac{1}{2}\log\left(1+\gamma\right) = \frac{1}{2}\log\left(1+\mathsf{d}\mathsf{snr}_2\right) - \frac{1}{2}\log\left(1+\mathsf{d}\gamma\right)$$

we have:

$$\begin{aligned}
\frac{1}{2}\int_{\gamma}^{\mathsf{snr}_2} \mathsf{MMSE}^{\mathsf{c}}(\tau)\,\mathsf{d}\tau &\geq \frac{1}{2}\log\left(1+\alpha\mathsf{snr}_2\right) - \frac{1}{2}\log\left(1+\gamma\right) \\
&= \frac{1}{2}\log\left(1+\mathsf{d}\mathsf{snr}_2\right) - \frac{1}{2}\log\left(1+\mathsf{d}\gamma\right) \\
&= \frac{1}{2}\int_{\gamma}^{\mathsf{snr}_2} \mathsf{mmse}_G(\tau)\,\mathsf{d}\tau, \quad X_G \sim \mathcal{N}(0,\mathsf{d}), i.i.d.
\end{aligned}$$

Proof Sketch - Cont.

Using the "single crossing point" property and the above inequality we can conclude:

The single crossing point of $\text{mmse}_G(\tau)$ and $\text{MMSE}^c(\tau)$, if exists, will occur somewhere in the region (γ, ∞).

Thus, we have the following lower bound:

$$\text{MMSE}^c(\gamma) \geq \frac{\text{d}(\gamma)}{1 + \text{d}(\gamma)\gamma} = \frac{\alpha\text{snr}_2 - \gamma}{\text{snr}_2 - \gamma}\frac{1}{1+\gamma}$$

Specifically for $\gamma = \text{snr}_1$ we obtain

$$\text{MMSE}^c(\text{snr}_1) \geq \frac{\alpha\text{snr}_2 - \text{snr}_1}{\text{snr}_2 - \text{snr}_1}\frac{1}{1+\text{snr}_1}.$$

Deriving α as a function of the constraining β, and substituting it in $R_c = \frac{1}{2}\log\left(1 + \alpha\text{snr}_2\right)$ results with the superposition rate.

Proof Sketch - Cont.

Using the "single crossing point" property and the above inequality we can conclude:

The single crossing point of $\mathrm{mmse}_G(\tau)$ and $\mathrm{MMSE}^c(\tau)$, if exists, will occur somewhere in the region (γ, ∞).

Thus, we have the following lower bound:

$$\mathrm{MMSE}^c(\gamma) \geq \frac{\mathrm{d}(\gamma)}{1 + \mathrm{d}(\gamma)\gamma} = \frac{\alpha \mathrm{snr}_2 - \gamma}{\mathrm{snr}_2 - \gamma} \frac{1}{1+\gamma}$$

Specifically for $\gamma = \mathrm{snr}_1$ we obtain

$$\mathrm{MMSE}^c(\mathrm{snr}_1) \geq \frac{\alpha \mathrm{snr}_2 - \mathrm{snr}_1}{\mathrm{snr}_2 - \mathrm{snr}_1} \frac{1}{1+\mathrm{snr}_1}.$$

Deriving α as a function of the constraining β, and substituting it in $R_c = \frac{1}{2}\log(1 + \alpha \mathrm{snr}_2)$ results with the superposition rate.

The effect at other snrs

Theorem 2

From the set of reliable codes of rate

$$\mathsf{R}_c = \frac{1}{2}\log\left(1+\beta\mathsf{snr}_2\right) + \frac{1}{2}\log\left(\frac{1+\mathsf{snr}_1}{1+\beta\mathsf{snr}_1}\right)$$

complying with the MMSE constraint at snr_1:

$$\mathsf{MMSE}^c(\mathsf{snr}_1) \leq \frac{\beta}{1+\beta\mathsf{snr}_1}$$

the superposition codebook provides the minimum MMSE for all snrs.

Extension to two MMSE constraints

Theorem 3 [Bustin and Shamai, IZS 2012]

Assuming $\mathsf{snr}_0 < \mathsf{snr}_1 < \mathsf{snr}_2$ the solution of,

$$\begin{aligned} \max \quad & I(\mathsf{snr}_2) \\ \text{s.t.} \quad & \mathsf{MMSE}^c(\mathsf{snr}_1) \leq \frac{\beta_1}{1+\beta_1 \mathsf{snr}_1}, \quad \mathsf{MMSE}^c(\mathsf{snr}_0) \leq \frac{\beta_0}{1+\beta_0 \mathsf{snr}_0} \end{aligned}$$

for some positive β_1, β_0 such that $\beta_1 + \beta_0 \leq 1$ and $\beta_1 < \beta_0$, is

$$I(\mathsf{snr}_2) = \frac{1}{2}\log\left((1+\beta_1 \mathsf{snr}_2)\frac{1+\beta_0 \mathsf{snr}_1}{1+\beta_1 \mathsf{snr}_1}\frac{1+\mathsf{snr}_0}{1+\beta_0 \mathsf{snr}_0}\right)$$

and is attainable when using the optimal three-layers Gaussian superposition codebook designed for $(\mathsf{snr}_0, \mathsf{snr}_1, \mathsf{snr}_2)$ with rate-splitting coefficients (β_0, β_1).
When $\beta_0 < \beta_1$ the first constraint can be removed and we return to the case of a single constraint given in Theorem 1.

Proof Sketch

- Optimal Gaussian three-layers superposition codebook comply with the above MMSE constraints and attain the maximum rate.
- We need a tight upper bound on the rate.

Using Theorem 1

Considering only the constraint on $\text{MMSE}^c(\text{snr}_0)$ we obtain the following upper bound on the rate at snr_1:

$$I(\text{snr}_1) \leq \frac{1}{2}\log\left(1+\beta_0\text{snr}_1\right) + \frac{1}{2}\log\left(\frac{1+\text{snr}_0}{1+\beta_0\text{snr}_0}\right)$$

Proof Sketch

- Optimal Gaussian three-layers superposition codebook comply with the above MMSE constraints and attain the maximum rate.
- We need a tight upper bound on the rate.

Using Theorem 1

Considering only the constraint on $\mathrm{MMSE}^c(\mathsf{snr}_0)$ we obtain the following upper bound on the rate at snr_1:

$$I(\mathsf{snr}_1) \leq \frac{1}{2}\log\left(1+\beta_0\mathsf{snr}_1\right) + \frac{1}{2}\log\left(\frac{1+\mathsf{snr}_0}{1+\beta_0\mathsf{snr}_0}\right)$$

Proof Sketch - Cont.

The I-MMSE approach

$$I(\mathsf{snr}_2) - I(\mathsf{snr}_1) = \frac{1}{2}\int_{\mathsf{snr}_1}^{\mathsf{snr}_2} \mathrm{MMSE}^{c}(\tau)\mathrm{d}\tau \leq \frac{1}{2}\int_{\mathsf{snr}_1}^{\mathsf{snr}_2} \mathrm{mmse}_G(\tau)\mathrm{d}\tau$$

where $X_G \sim \mathcal{N}(0, \beta_1)$ and i.i.d. This is valid since according to the constraint on $\mathrm{MMSE}^{c}(\mathsf{snr}_1)$ we have

$$\mathrm{MMSE}^{c}(\mathsf{snr}_1) \leq \frac{\beta_1}{1+\beta_1\mathsf{snr}_1} = \mathrm{mmse}_G(\mathsf{snr}_1)$$

and according to the single crossing point property

$$\mathrm{MMSE}^{c}(\tau) \leq \mathrm{mmse}_G(\tau), \quad \forall\tau \geq \mathsf{snr}_1$$

Putting the two upper bounds together, we obtain the desired result.

Proof Sketch - Cont.

The I-MMSE approach

$$I(\mathsf{snr}_2) - I(\mathsf{snr}_1) = \frac{1}{2}\int_{\mathsf{snr}_1}^{\mathsf{snr}_2} \mathrm{MMSE}^c(\tau)\mathrm{d}\tau \leq \frac{1}{2}\int_{\mathsf{snr}_1}^{\mathsf{snr}_2} \mathrm{mmse}_G(\tau)\mathrm{d}\tau$$

where $X_G \sim \mathcal{N}(0, \beta_1)$ and i.i.d. This is valid since according to the constraint on $\mathrm{MMSE}^c(\mathsf{snr}_1)$ we have

$$\mathrm{MMSE}^c(\mathsf{snr}_1) \leq \frac{\beta_1}{1+\beta_1\mathsf{snr}_1} = \mathrm{mmse}_G(\mathsf{snr}_1)$$

and according to the single crossing point property

$$\mathrm{MMSE}^c(\tau) \leq \mathrm{mmse}_G(\tau), \quad \forall \tau \geq \mathsf{snr}_1$$

Putting the two upper bounds together, we obtain the desired result.

Mutual information disturbance: single constraint

Theorem 4 [Bandemer and El Gamal, ISIT 2011]

Assuming $\mathsf{snr}_1 < \mathsf{snr}_2$ the solution of the following optimization problem,

$$\begin{aligned} \max \quad & I_n(\mathsf{snr}_2) \\ \text{s.t.} \quad & I_n(\mathsf{snr}_1) \leq \frac{1}{2}\log\left(1 + \alpha^\star \mathsf{snr}_1\right) \end{aligned}$$

for some $\alpha^\star \in [0, 1]$, is the following

$$I_n(\mathsf{snr}_2) = \frac{1}{2}\log\left(1 + \alpha^\star \mathsf{snr}_2\right).$$

Equality is attained, for any n, by choosing X Gaussian with i.i.d. components of variance $\alpha^\star$. For $n \to \infty$ equality is also attained by a Gaussian codebook designed for snr_2 with limited power of $\alpha^\star$.

Mutual information disturbance: single constraint

Alternative I-MMSE proof

Since, $0 \leq I_n(\mathsf{snr}_1) \leq \frac{1}{2}\log\left(1+\mathsf{snr}_1\right)$ there exists an $\alpha^\star \in [0,1]$ such that

$$I_n(\mathsf{snr}_1) = \frac{1}{n}\log\left(1+\alpha^\star \mathsf{snr}_1\right).$$

$\Longrightarrow$ $\mathrm{MMSE}^{c_n}(\gamma)$ and $\mathrm{mmse}_G(\gamma)$ of $X_G \sim \mathcal{N}(0,\alpha^\star)$ cross in $[0,\mathsf{snr}_1]$.

Using the I-MMSE

$$I_n(\mathsf{snr}_2) = \frac{1}{2}\log\left(1+\alpha^* \mathsf{snr}_1\right) + \int_{\mathsf{snr}_1}^{\mathsf{snr}_2} \mathrm{MMSE}^{c_n}(\gamma)\, d\gamma$$
$$\leq \frac{1}{2}\log\left(1+\alpha^* \mathsf{snr}_2\right)$$

due to the "single crossing point" property which ensures

$$\mathrm{MMSE}^{c_n}(\gamma) \leq \mathrm{mmse}_G(\gamma), \quad \forall \gamma \in [\mathsf{snr}_1, \infty)$$

Mutual information disturbance: single constraint

Alternative I-MMSE proof

Since, $0 \leq I_n(\mathsf{snr}_1) \leq \frac{1}{2}\log(1 + \mathsf{snr}_1)$ there exists an $\alpha^\star \in [0,1]$ such that

$$I_n(\mathsf{snr}_1) = \frac{1}{n}\log(1 + \alpha^\star \mathsf{snr}_1).$$

$\Longrightarrow$ $\mathrm{MMSE}^{c_n}(\gamma)$ and $\mathrm{mmse}_G(\gamma)$ of $X_G \sim \mathcal{N}(0, \alpha^\star)$ cross in $[0, \mathsf{snr}_1]$.

Using the I-MMSE

$$\begin{aligned} I_n(\mathsf{snr}_2) &= \frac{1}{2}\log(1 + \alpha^* \mathsf{snr}_1) + \int_{\mathsf{snr}_1}^{\mathsf{snr}_2} \mathrm{MMSE}^{c_n}(\gamma) d\gamma \\ &\leq \frac{1}{2}\log(1 + \alpha^* \mathsf{snr}_2) \end{aligned}$$

due to the "single crossing point" property which ensures

$$\mathrm{MMSE}^{c_n}(\gamma) \leq \mathrm{mmse}_G(\gamma), \quad \forall \gamma \in [\mathsf{snr}_1, \infty)$$

Mutual information disturbance: multiple constraints

Theorem 5

Assuming $\mathsf{snr}_1 < \mathsf{snr}_2 < \cdots < \mathsf{snr}_K$ the solution of

$$\begin{aligned} \max \quad & I_n(\mathsf{snr}_K) \\ \text{s.t.} \quad & \forall i \in \{1, \cdots, K-1\}, \quad I_n(\mathsf{snr}_i) \leq \frac{1}{2}\log\left(1 + \alpha_i \mathsf{snr}_i\right) \end{aligned}$$

for some $\alpha_i \in [0, 1]$, is the following

$$I_n(\mathsf{snr}_K) = \frac{1}{2}\log\left(1 + \alpha_\ell \mathsf{snr}_K\right)$$

where α_ℓ, $\ell \in \{1, \cdots, K-1\}$, is defined such that

$$\forall i \in \{1, \cdots, K-1\} \quad \frac{1}{2}\log\left(1 + \alpha_\ell \mathsf{snr}_i\right) \leq \frac{1}{2}\log\left(1 + \alpha_i \mathsf{snr}_i\right)$$

The maximum rate is attained, for any n, by choosing X Gaussian with i.i.d. components of variance α_ℓ. For $n \to \infty$ equality is also attained by a Gaussian codebook designed for snr_K with limited power of α_ℓ.

Summary and Outlook

- These results provide the engineering insight to the good performance of the Han and Kobayashi scheme on the interference channel. We show that the Han and Kobayashi scheme is optimal MMSE-wise.
- The results can be easily extended to K MMSE constraint [Bustin and Shamai, submitted ISIT 2012].
- The engineering advantage of the MMSE disturbance measure over the mutual information measure in the Gaussian channel are demonstrated.
- Single code I-MMSE tradeoff, bounded by the optimal superposition coding tradeoff

  [Bennatan, Shamai, Calderbank, <arXiv:1008.1766v1-2010>].
- Interesting challenges: optimization of $I_n(\mathsf{snr}_2)$, under the $\mathrm{MMSE}^{c_n}(\mathsf{snr}_1)$ constraint when block-length n is finite. For $n = 1$, conjecture: the optimal X is discrete [Shamai, ISIT 2011].

Thank You!

“MMSE interference in Gaussian Channels”
Ronit Bustin and Shlomo Shamai

We consider the scalar Gaussian channel, and address the problem of maximizing the average mutual information of a power constraint n component ($n \to \infty$) input random vector at a given signal-to-noise ratio (snr), satisfying a minimum mean square error (MMSE) constraint at another lower snr value. We use the MMSE as an effective interference (disturbance) measure, motivated by interference networks, where codes are expected not only to optimize performance for the intended user but inflict minimum interference on other users. We show via the information-estimation relation, that superposition coding is optimal in this respect, providing further intuition to the effectiveness of the Han-Kobayashi coding strategy on the interference channel, and performance of 'bad' codes.
Moreover, the MMSE function of those codes, attaining the best rate at some snr, subjected to a prescribed MMSE demand at some other snr, is completely defined for all snr, and is the one obtained by the corresponding superposition codebooks.
Extensions to two MMSE constraints, are discussed, and compared to the results for a mutual information disturbance measure. Some challenges for this class of interference problems will also be discussed.