On the Minimum Mean p-th Error in Gaussian Noise Channels and its Applications

Alex Dytso, Ronit Bustin, Daniela Tuninetti,
Natasha Devroye, H. Vincent Poor,
and Shlomo Shamai (Shitz)

PRINCETON UNIVERSITY

TECHNION
Israel Institute of Technology

Notation

$$\mathbf{X} \in \mathbb{R}^n, \ \frac{1}{n}\mathrm{Tr}\left(\mathbb{E}[\mathbf{X}\mathbf{X}^T]\right) \leq 1,$$
$$\mathbf{Y}_{\mathrm{snr}} = \sqrt{\mathrm{snr}}\mathbf{X} + \mathbf{Z},$$
$$I_n(\mathbf{X}, \mathrm{snr}) = \frac{1}{n} I(\mathbf{X}; \mathbf{Y}_{\mathrm{snr}}),$$
$$\mathrm{mmse}(\mathbf{X}|\mathbf{Y}_{\mathrm{snr}}) = \mathrm{mmse}(\mathbf{X}, \mathrm{snr}) := \frac{1}{n}\mathrm{Tr}\left(\mathbb{E}\left[\mathrm{cov}(\mathbf{X}|\mathbf{Y}_{\mathrm{snr}})\right]\right)$$

I-MMSE relationship

$$\frac{d}{d\mathrm{snr}} I_n(\mathbf{X}, \mathrm{snr}) = \frac{1}{2}\mathrm{mmse}(\mathbf{X}, \mathrm{snr})$$

$$I_n(\mathbf{X}, \mathrm{snr}) = \frac{1}{2}\int_0^{\mathrm{snr}} \mathrm{mmse}(\mathbf{X}, t)\,dt$$

PRINCETON UNIVERSITY

D. Guo, S. Shamai, and S. Verdu´, “Mutual information and minimum mean-square error in Gaussian channels,” *IEEE Trans. Inf. Theory*, vol. 51, no. 4, pp. 1261–1282, April 2005.

Outline

- Notation and Past Work
- Minimum Mean p^{th} Error (MMPE)
- Properties of the MMPE
- Conditional MMPE
- Applications to Information Theory Problems

Part 1

Pastwork

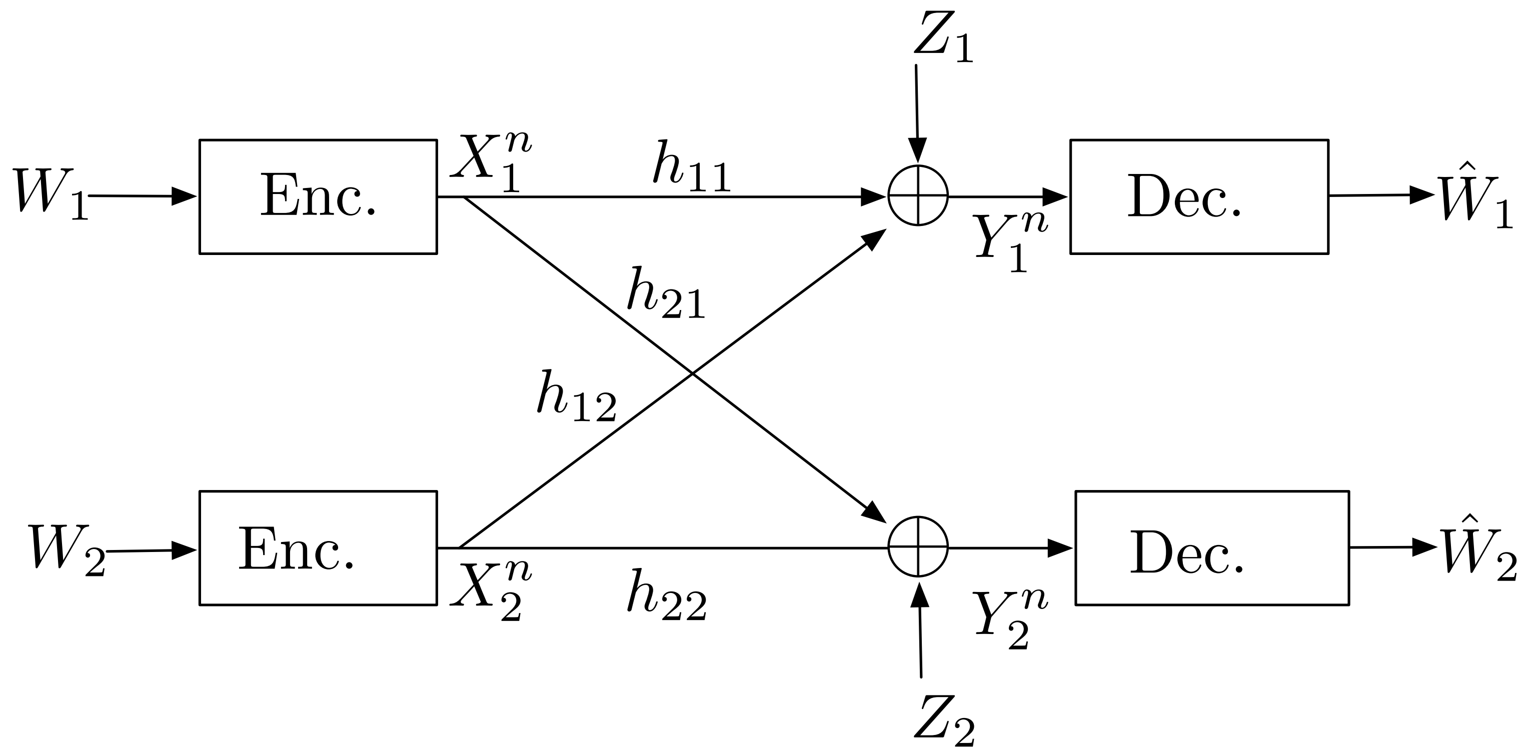
PRINCETON UNIVERSITY

Part 1a

My Thesis

PRINCETON
UNIVERSITY

Interference Channel



with the usual assumptions and capacity region definition.

Symmetric: $$\begin{aligned} |h_{11}|^2 &= |h_{22}|^2 = \mathrm{S} \\ |h_{12}|^2 &= |h_{21}|^2 = \mathrm{I} \end{aligned}$$

TIN Inner Bound

Capacity: $\mathcal{C} = \lim_{n\to\infty} \mathrm{co}\left(\bigcup_{P_{X_1^n X_2^n}=P_{X_1^n}P_{X_2^n}} \left\{ \begin{array}{l} 0 \le R_1 \le \frac{1}{n} I(X_1^n;Y_1^n) \\ 0 \le R_2 \le \frac{1}{n} I(X_2^n;Y_2^n) \end{array} \right\}\right)$

R. Ahlswede, “Multi-way communication channels,” in Proc. IEEE Int. Symp. Inf. Theory, March 1973, pp. 23–52.

$\big\downarrow$ i.i.d. inputs

Treat Interference as Noise (TIN) Inner Bound:

$$\mathcal{R}_{\mathrm{in}}^{\mathrm{TIN+TS}} = \mathrm{co}\left(\bigcup_{P_{X_1X_2}=P_{X_1}P_{X_2}} \left\{ \begin{array}{l} 0 \le R_1 \le I(X_1;Y_1) \\ 0 \le R_2 \le I(X_2;Y_2) \end{array} \right\}\right)$$

With Time Sharing

$$\mathcal{R}_{\mathrm{in}}^{\mathrm{TINnoTS}} = \bigcup_{P_{X_1X_2}=P_{X_1}P_{X_2}} \left\{ \begin{array}{l} 0 \le R_1 \le I(X_1;Y_1) \\ 0 \le R_2 \le I(X_2;Y_2) \end{array} \right\}$$

No Time Sharing

How far away is TINnoTS from the Capacity?

Known Results

Sum-Capacity in Very Weak Interference

$$\sqrt{\frac{\mathrm{S}}{\mathrm{I}}}(1+\mathrm{I}) \leq \frac{1}{2}$$

X. Shang, G. Kramer, and B. Chen, “A new outer bound and the noisy-interference sum rate capacity for Gaussian interference channels,” IEEE Trans. Inf. Theory, vol. 55, no. 2, pp. 689–699, 2009.

Motahari, A.S.; Khandani, A.K., "Capacity Bounds for the Gaussian Interference Channel," IEEE Trans. Inf. Theory, vol.55, no.2, pp.620,643, Feb. 2009

Annapureddy, V.S.; Veeravalli, V.V., "Gaussian Interference Networks: Sum Capacity in the Low-Interference Regime and New Outer Bounds on the Capacity Region," IEEE Trans. Inf. Theory, vol.55, no. 7, pp.3032,3050, July 2009

Known Results

Gaussian is NOT optimal

R. Cheng and S. Verdu, “On limiting characterizations of memoryless multiuser capacity regions,” *IEEE Trans. Inf. Theory*, vol. 39, no. 2, pp. 609-612, 1993.

E. Abbe and L. Zheng, “A coordinate system for Gaussian networks,” *IEEE Trans. Inf. Theory*, vol. 58, no. 2, pp. 721-733, 2012.

M. Vaezi and H.V. Poor. "On limiting expressions for the capacity region of Gaussian interference channels." *Proc. 49th Asilomar Conference on Signals, Systems and Computers.* 2015.

This stimulate our investigation of discrete inputs

PRINCETON
UNIVERSITY

My Thesis

Mixed inputs

$$X_{\text{mix}} = \sqrt{1-\delta} X_D + \sqrt{\delta} X_G,$$
$$\delta \in [0,1],$$
$$X_G \sim \mathcal{N}(0,1)$$
$$\mathbb{E}[X_D^2] \leq 1$$

Key Result:

TIN is $O(1)$ or $O\left(\log\log(\text{snr})\right)$ away from capacity.

A. Dytso, D. Tuninetti, and N. Devroye. "Interference as noise: Friend or foe?." *IEEE Trans. Inf. Theory*, 62.6 (2016): 3561-3596.

PRINCETON UNIVERSITY

Ozarow-Wyner Bound

(Ozarow-Wyner Bound.) Let X_D be a discrete random variable then

$$H(X_D) - \text{gap} \le I(X_D;Y) \le H(X_D),$$

$$\text{gap} = \frac{1}{2}\log\left(\frac{\pi e}{6}\right) + \frac{1}{2}\log\left(1 + \frac{12 \cdot \text{lmmse}(X_D|Y)}{d_{\min}^2(X_D)}\right),$$

$$\text{lmmse}(X_D|Y) = \frac{\mathbb{E}[X_D^2]}{1 + \text{snr} \cdot \mathbb{E}[X_D^2]}$$

$$d_{\min}(X_D) = \inf_{x_i, x_j \in \text{supp}(X_D): i \neq j} |x_i - x_j|.$$

L. Ozarow and A. Wyner, "On the capacity of the Gaussian channel with a finite number of input levels," *IEEE Trans. Inf. Theory*, vol. 36, no. 6, pp. 1426–1428, Nov 1990.

Take X_D to be PAM with $N = \sqrt{1 + \text{snr}}$ then

$$\frac{1}{2}\log(1 + \text{snr}) - I(X_D;Y) = \text{gap} \approx 0.75 \text{ bit}$$

Ozarow-Wyner Bound

(Ozarow-Wyner Bound.) Let X_D be a discrete random variable then

$$H(X_D) - \text{gap} \leq I(X_D;Y) \leq H(X_D),$$

$$\text{gap} = \frac{1}{2}\log\left(\frac{\pi e}{6}\right) + \frac{1}{2}\log\left(1 + \frac{12 \cdot \text{lmmse}(X_D|Y)}{d_{\min}^2(X_D)}\right),$$

$$\text{lmmse}(X_D|Y) = \frac{\mathbb{E}[X_D^2]}{1 + \text{snr} \cdot \mathbb{E}[X_D^2]}$$

$$d_{\min}(X_D) = \inf_{x_i, x_j \in \text{supp}(X_D): i \neq j} |x_i - x_j|.$$

L. Ozarow and A. Wyner, "On the capacity of the Gaussian channel with a finite number of input levels," *IEEE Trans. Inf. Theory*, vol. 36, no. 6, pp. 1426–1428, Nov 1990.

Can this be improved?

Take X_D to be PAM with $N = \sqrt{1+\text{snr}}$ then

$$\frac{1}{2}\log(1+\text{snr}) - I(X_D;Y) = \text{gap} \approx 0.75 \text{ bit}$$

PRINCETON UNIVERSITY

Ozarow-Wyner Bound

(Ozarow-Wyner Bound.) Let X_D be a discrete random variable then

$$\begin{aligned} H(X_D) - \text{gap} \leq I(X_D;Y) &\leq H(X_D), \\ \text{gap} &= \frac{1}{2}\log\left(\frac{\pi e}{6}\right) + \frac{1}{2}\log\left(1 + \frac{12 \cdot \text{lmmse}(X_D|Y)}{d_{\min}^2(X_D)}\right), \\ \text{lmmse}(X_D|Y) &= \frac{\mathbb{E}[X_D^2]}{1 + \text{snr} \cdot \mathbb{E}[X_D^2]}, \\ d_{\min}(X_D) &= \inf_{x_i, x_j \in \text{supp}(X_D): i \neq j} |x_i - x_j|. \end{aligned}$$

L. Ozarow and A. Wyner, "On the capacity of the Gaussian channel with a finite number of input levels," *IEEE Trans. Inf. Theory*, vol. 36, no. 6, pp. 1426–1428, Nov 1990.

Improvement

$$\begin{aligned} \text{gap} &= \frac{1}{2}\log\left(\frac{\pi e}{6}\right) + \frac{1}{2}\log\left(1 + \frac{12 \cdot \text{mmse}(X_D|Y)}{d_{\min}^2(X_D)}\right), \\ \text{mmse}(X_D|Y) &= \mathbb{E}\left[(X_D - \mathbb{E}[X_D|Y])^2\right]. \end{aligned}$$

A. Dytso, D. Tuninetti, and N. Devroye, "Interference as noise: Friend or foe?" *IEEE Trans. Inf. Theory*, vol. 62, no. 6, pp. 3561–3596, 2016.

PRINCETON UNIVERSITY

Ozarow-Wyner Bound

(Ozarow-Wyner Bound.) Let X_D be a discrete random variable then

$$H(X_D) - \text{gap} \le I(X_D;Y) \le H(X_D),$$

$$\text{gap} = \frac{1}{2}\log\left(\frac{\pi e}{6}\right) + \frac{1}{2}\log\left(1 + \frac{12\cdot \text{lmmse}(X_D|Y)}{d_{\min}^2(X_D)}\right),$$

$$\text{lmmse}(X_D|Y) = \frac{\mathbb{E}[X_D^2]}{1+\text{snr}\cdot\mathbb{E}[X_D^2]}$$

$$d_{\min}(X_D) = \inf_{x_i,x_j\in\text{supp}(X_D): i\neq j} |x_i - x_j|.$$

L. Ozarow and A. Wyner, "On the capacity of the Gaussian channel with a finite number of input levels," *IEEE Trans. Inf. Theory*, vol. 36, no. 6, pp. 1426–1428, Nov 1990.

$$\text{mmse}(X_D|Y) \le \text{lmmse}(X_D|Y),$$

$$\text{lmmse}(X_D|Y) = O\left(\frac{1}{\text{snr}}\right),$$

$$\text{mmse}(X_D|Y) = O\left(\mathrm{e}^{-\frac{\text{snr} d_{\min}^2}{4}}\right).$$

Improvement

$$\text{gap} = \frac{1}{2}\log\left(\frac{\pi e}{6}\right) + \frac{1}{2}\log\left(1 + \frac{12\cdot \text{mmse}(X_D|Y)}{d_{\min}^2(X_D)}\right),$$

$$\text{mmse}(X_D|Y) = \mathbb{E}\left[(X_D - \mathbb{E}[X_D|Y])^2\right].$$

A. Dytso, D. Tuninetti, and N. Devroye, "Interference as noise: Friend or foe?" *IEEE Trans. Inf. Theory*, vol. 62, no. 6, pp. 3561–3596, 2016.

PRINCETON UNIVERSITY

Ozarow-Wyner Bound

(Ozarow-Wyner Bound.) Let X_D be a discrete random variable then

$$H(X_D) - \text{gap} \le I(X_D;Y) \le H(X_D),$$

$$\text{gap} = \frac{1}{2}\log\left(\frac{\pi e}{6}\right) + \frac{1}{2}\log\left(1 + \frac{12 \cdot \text{lmmse}(X_D|Y)}{d_{\min}^2(X_D)}\right),$$

$$\text{lmmse}(X_D|Y) = \frac{\mathbb{E}[X_D^2]}{1 + \text{snr} \cdot \mathbb{E}[X_D^2]}$$

$$d_{\min}(X_D) = \inf_{x_i, x_j \in \text{supp}(X_D): i \neq j} |x_i - x_j|.$$

L. Ozarow and A. Wyner, "On the capacity of the Gaussian channel with a finite number of input levels," *IEEE Trans. Inf. Theory*, vol. 36, no. 6, pp. 1426–1428, Nov 1990.

$$\frac{1}{2}\log(1+\text{snr}) - I(X_D;Y) \approx 0.75 \text{ bit.}$$

Improvement

$$\frac{1}{2}\log(1+\text{snr}) - I(X_D;Y) \approx 0.65 \text{ bit.}$$

$$\text{gap} = \frac{1}{2}\log\left(\frac{\pi e}{6}\right) + \frac{1}{2}\log\left(1 + \frac{12 \cdot \text{mmse}(X_D|Y)}{d_{\min}^2(X_D)}\right),$$

$$\text{mmse}(X_D|Y) = \mathbb{E}\left[(X_D - \mathbb{E}[X_D|Y])^2\right].$$

A. Dytso, D. Tuninetti, and N. Devroye, "Interference as noise: Friend or foe?" *IEEE Trans. Inf. Theory*, vol. 62, no. 6, pp. 3561–3596, 2016.

Can we do better?

PRINCETON UNIVERSITY

Part 1b

Bounds on the MMSE

PRINCETON UNIVERSITY

MMSE Bounds: Single Crossing Point Property

(**SCPP.**) For any fixed $\mathbf{X}$, suppose that $\mathrm{mmse}(\mathbf{X}, \mathrm{snr}_0) = \frac{\beta}{1+\beta \mathrm{snr}_0}$, for some fixed $\beta \geq 0$. Then

$$\mathrm{mmse}(\mathbf{X}, \mathrm{snr}) \leq \frac{\beta}{1+\beta \mathrm{snr}}, \quad \mathrm{snr} \in [\mathrm{snr}_0, \infty)$$
$$\mathrm{mmse}(\mathbf{X}, \mathrm{snr}) \geq \frac{\beta}{1+\beta \mathrm{snr}}, \quad \mathrm{snr} \in [0, \mathrm{snr}_0).$$

SCPP+I-MMSE important tool for derive converses:

- Version of EPI
- BC
- Wiretap
- MIMO channels

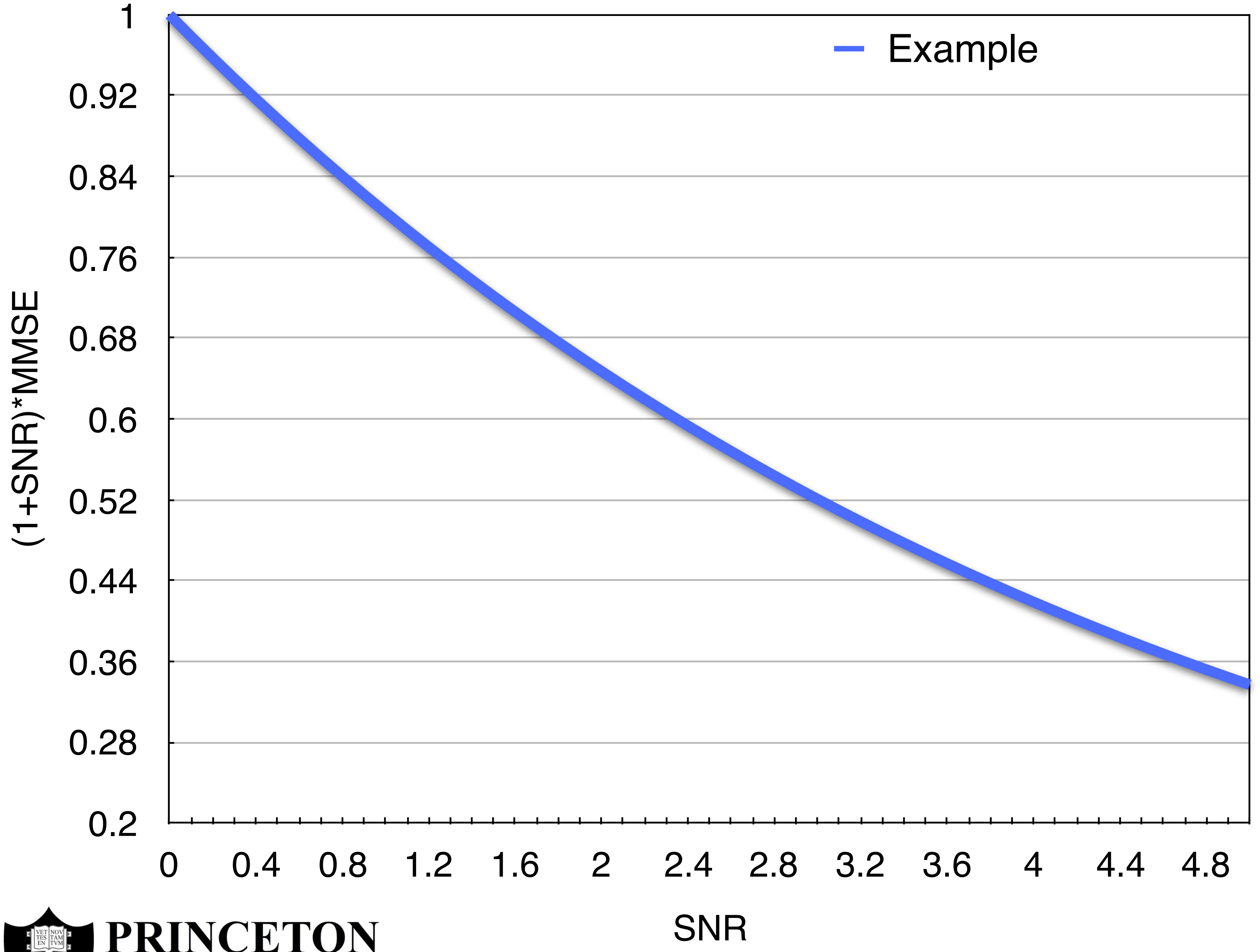
D. Guo, Y. Wu, S. Shamai, and S. Verdu, "Estimation in Gaussian noise: Properties of the minimum mean-square error," IEEE Trans. Inf. Theory, vol. 57, no. 4, pp. 2371–2385, April 2011.

R. Bustin, R. Schaefer, H. Poor, and S. Shamai, "On MMSE properties of optimal codes for the Gaussian wiretap channel," in Proc. IEEE Inf. Theory Workshop, April 2015, pp. 1–5.

R. Bustin, M. Payaro, D. Palomar, and S. Shamai, "On MMSE crossing properties and implications in parallel vector Gaussian channels," IEEE Trans. Inf. Theory, vol. 59, no. 2, pp. 818–844, Feb 2013.

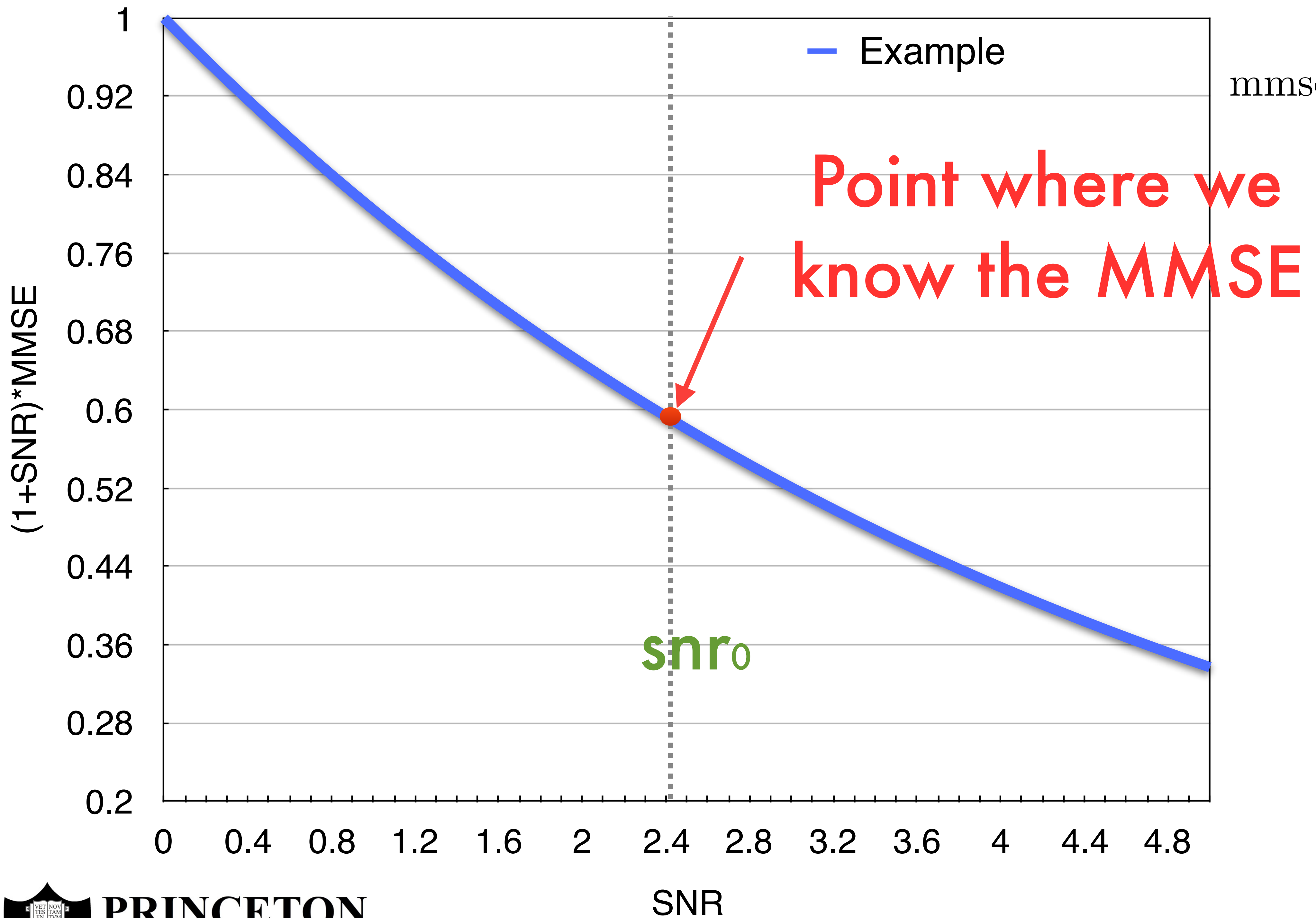
PRINCETON UNIVERSITY

Review: MMSE bounds



PRINCETON UNIVERSITY

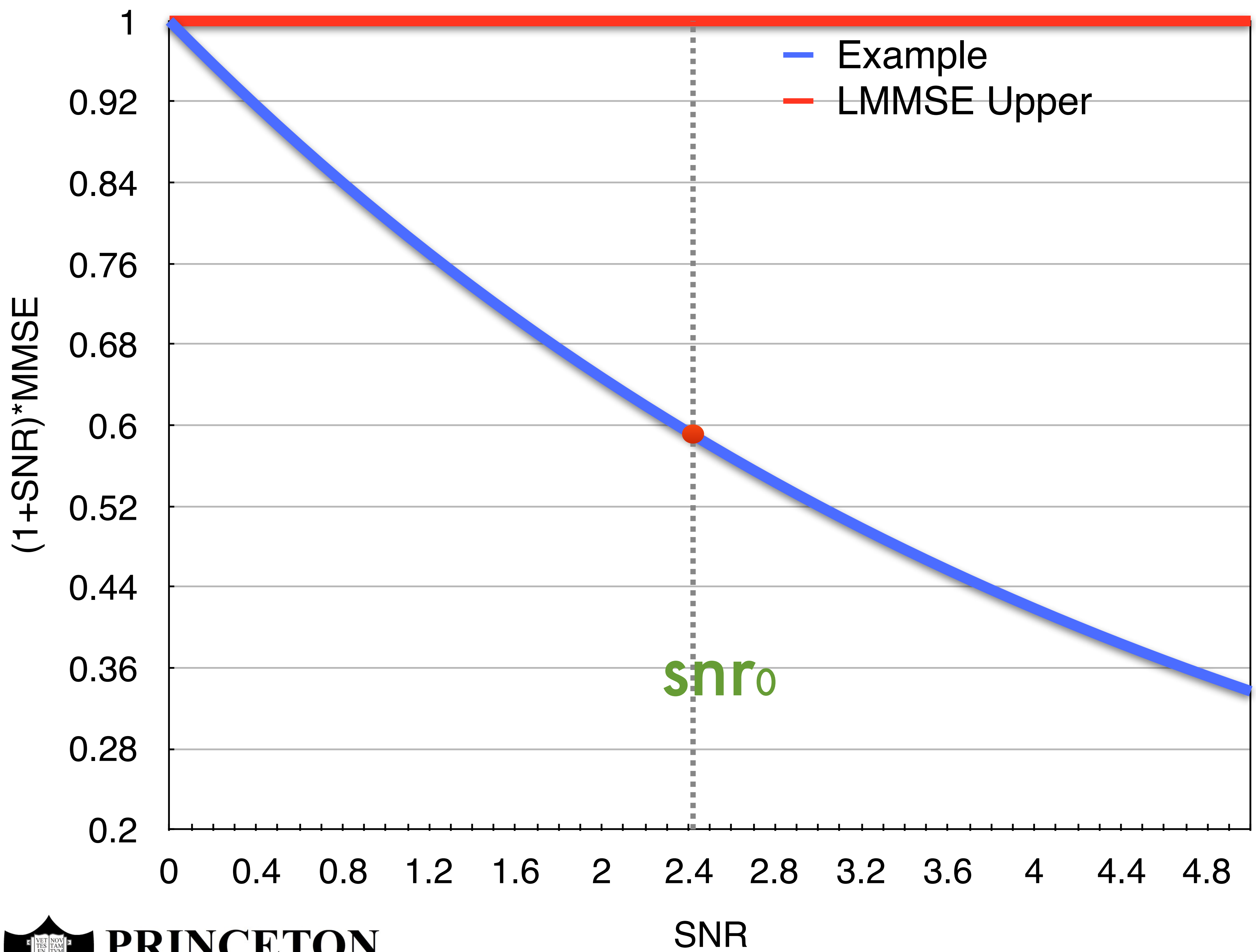
Review: MMSE bounds



$$\mathrm{mmse}(\mathbf{X}, \mathrm{snr}_0) = \frac{\beta}{1 + \beta \mathrm{snr}_0}$$

PRINCETON UNIVERSITY

Review: MMSE bounds



$$\mathrm{mmse}(\mathbf{X}, \mathrm{snr}_0) = \frac{\beta}{1 + \beta \mathrm{snr}_0}$$

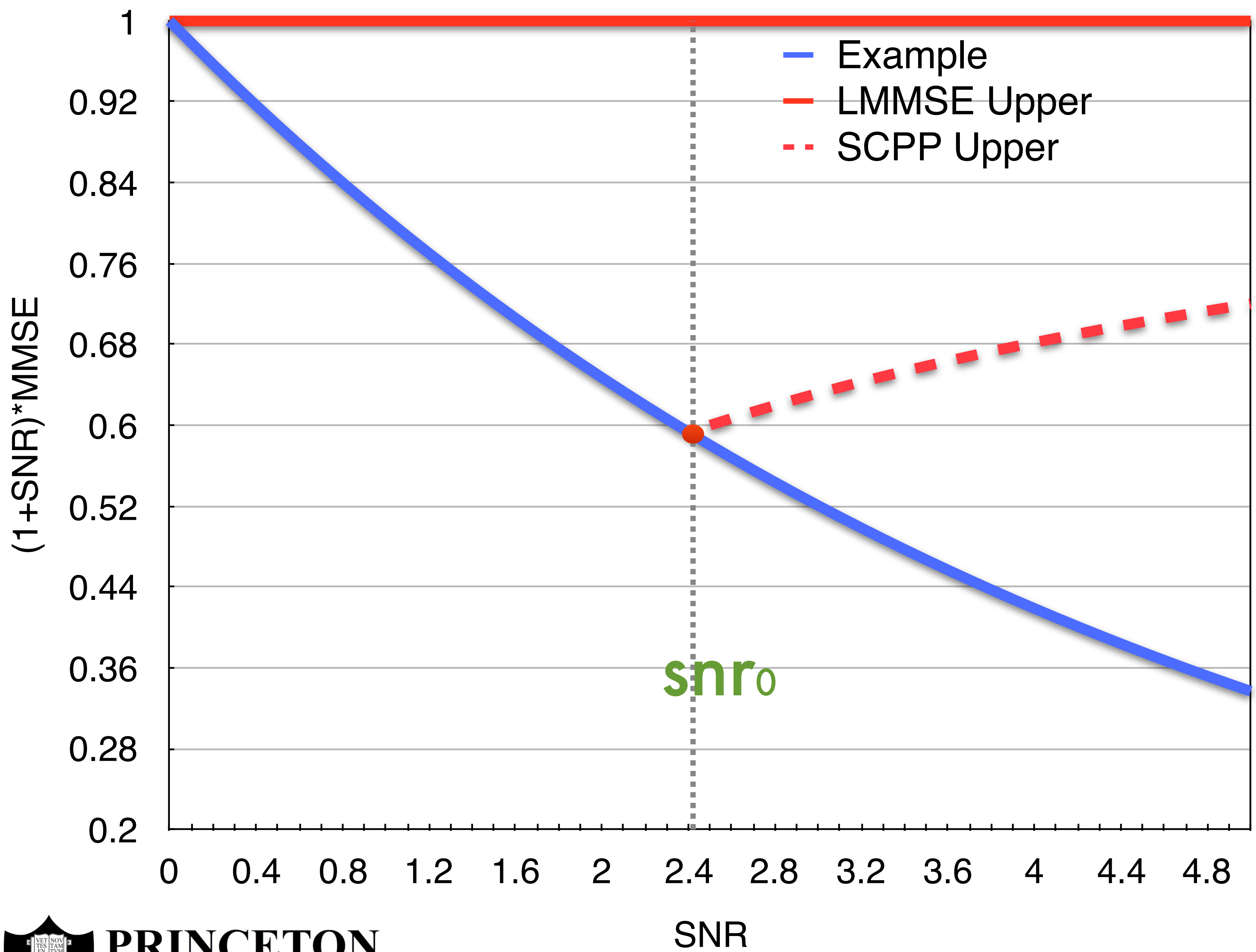
LMMSE u.b.

$$\mathrm{mmse}(\mathbf{X}, \mathrm{snr}) \leq \frac{1}{1 + \mathrm{snr}}$$

only power constraint

PRINCETON UNIVERSITY

Review: MMSE bounds



$$\mathrm{mmse}(\mathbf{X}, \mathrm{snr}_0) = \frac{\beta}{1+\beta\mathrm{snr}_0}$$

LMMSE u.b.

$$\mathrm{mmse}(\mathbf{X}, \mathrm{snr}) \leq \frac{1}{1+\mathrm{snr}}$$

SCPP u.b.

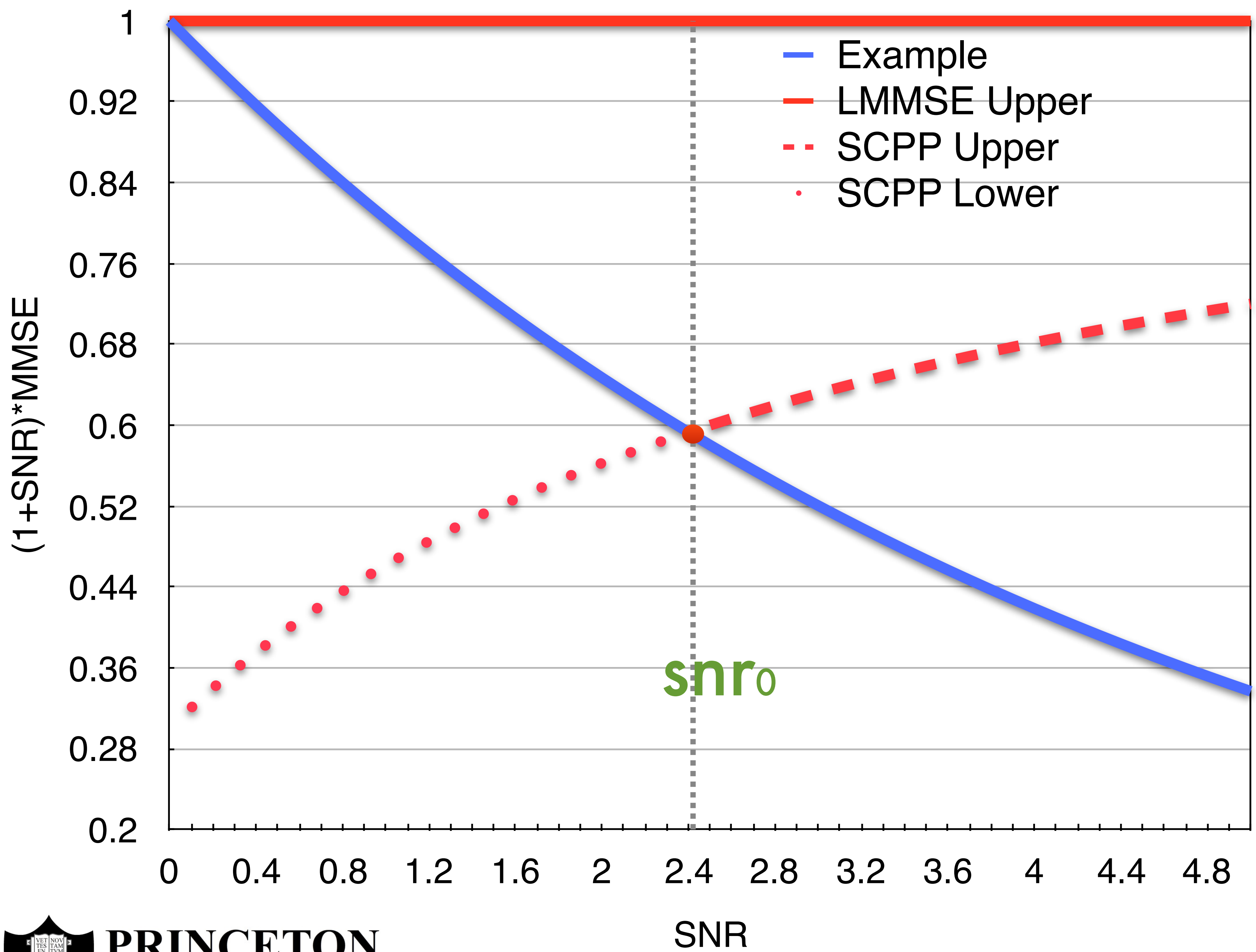
$$\mathrm{mmse}(\mathbf{X}, \mathrm{snr}) \leq \frac{\beta}{1+\beta\mathrm{snr}}$$

for $\mathrm{snr} \geq \mathrm{snr}_0$

MMSE constraint at snr_0

PRINCETON UNIVERSITY

MMSE bounds



$$\mathrm{mmse}(\mathbf{X}, \mathrm{snr}_0) = \frac{\beta}{1 + \beta \mathrm{snr}_0}$$

LMMSE u.b.

$$\mathrm{mmse}(\mathbf{X}, \mathrm{snr}) \leq \frac{1}{1 + \mathrm{snr}}$$

SCPP u.b.

$$\mathrm{mmse}(\mathbf{X}, \mathrm{snr}) \leq \frac{\beta}{1 + \beta \mathrm{snr}}$$

$$\text{for } \mathrm{snr} \geq \mathrm{snr}_0$$

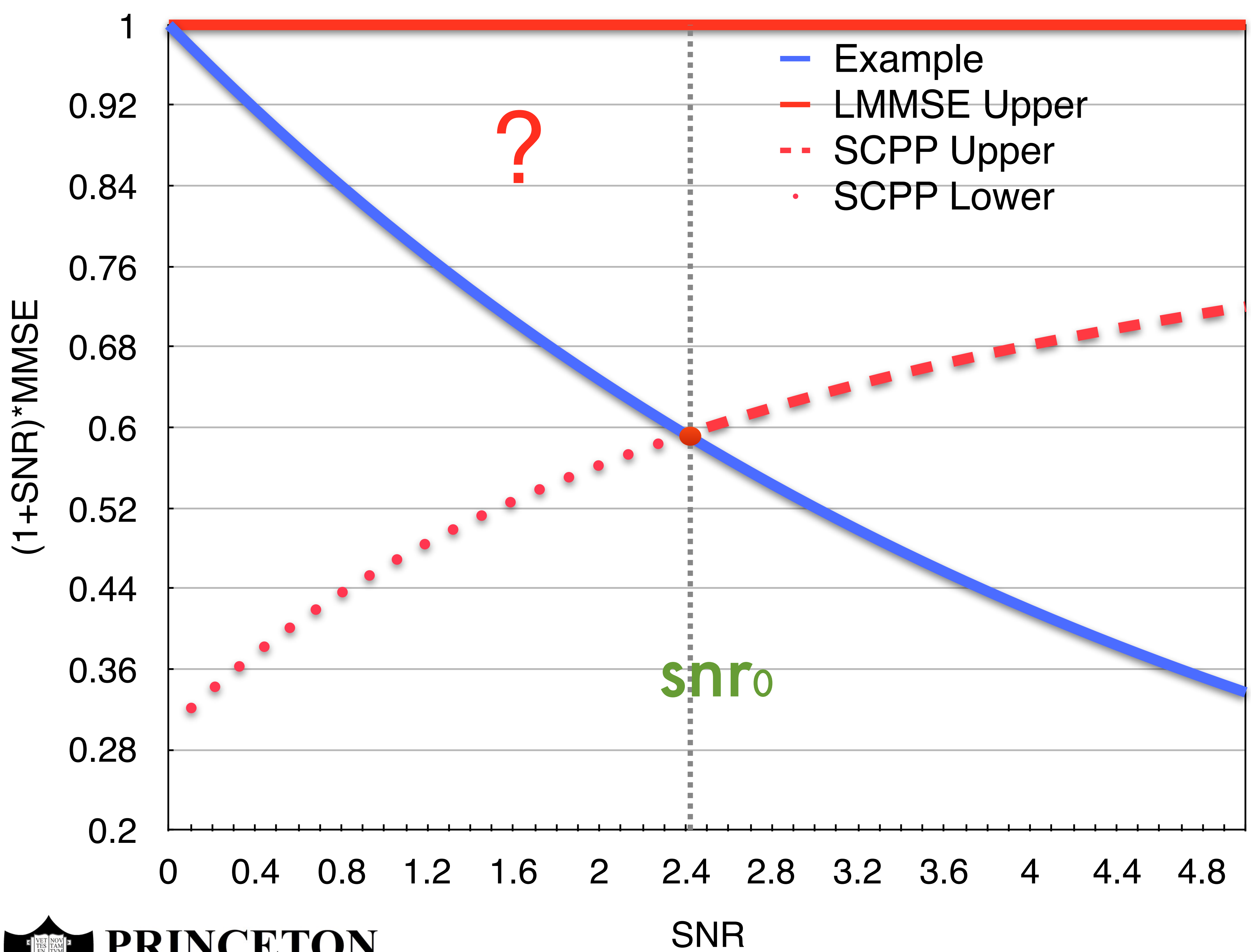
SCPP l.b.

$$\mathrm{mmse}(\mathbf{X}, \mathrm{snr}) \geq \frac{\beta}{1 + \beta \mathrm{snr}}$$

$$\text{for } \mathrm{snr} \leq \mathrm{snr}_0$$

PRINCETON
UNIVERSITY

Review: MMSE bounds



$$\mathrm{mmse}(\mathbf{X}, \mathrm{snr}_0) = \frac{\beta}{1 + \beta \mathrm{snr}_0}$$

LMMSE u.b.

$$\mathrm{mmse}(\mathbf{X}, \mathrm{snr}) \leq \frac{1}{1 + \mathrm{snr}}$$

SCPP u.b.

$$\mathrm{mmse}(\mathbf{X}, \mathrm{snr}) \leq \frac{\beta}{1 + \beta \mathrm{snr}}$$

$$\text{for } \mathrm{snr} \geq \mathrm{snr}_0$$

SCPP l.b.

$$\mathrm{mmse}(\mathbf{X}, \mathrm{snr}) \geq \frac{\beta}{1 + \beta \mathrm{snr}}$$

$$\text{for } \mathrm{snr} \leq \mathrm{snr}_0$$

PRINCETON UNIVERSITY

max-MMSE problem

Motivated by the need of a complementary (to the SCPP) upper bound for MMSE

max-MMSE problem

$$\mathrm{M}_n(\mathrm{snr}, \mathrm{snr}_0, \beta) := \sup_{\mathbf{X}} \mathrm{mmse}(\mathbf{X}, \mathrm{snr}),$$

$$\text{s.t. } \frac{1}{n}\mathrm{Tr}\left(\mathbb{E}[\mathbf{X}\mathbf{X}^{\mathrm{T}}]\right) \le 1,$$

$$\text{and } \mathrm{mmse}(\mathbf{X}, \mathrm{snr}_0) \le \frac{\beta}{1+\beta \mathrm{snr}_0}.$$

From SCPP

$$\mathrm{mmse}(\mathbf{X}, \mathrm{snr}) \le \frac{\beta}{1+\beta \mathrm{snr}}$$

$$\text{for } \mathrm{snr} \ge \mathrm{snr}_0$$

solution for $\mathrm{snr} > \mathrm{snr}_0$: Gaussian input with reduced power

In the rest focus on $\mathrm{snr} < \mathrm{snr}_0$

PRINCETON UNIVERSITY

Bounds on max-MMSE

$$\mathrm{M}_\infty(\mathrm{snr}, \mathrm{snr}_0, \beta) = \begin{cases} \frac{1}{1+\mathrm{snr}}, & \mathrm{snr} < \mathrm{snr}_0, \\ \frac{\beta}{1+\beta\mathrm{snr}}, & \mathrm{snr} \geq \mathrm{snr}_0, \end{cases} .$$

- Converse: LMMSE+SCPP.
- Achievability: superposition coding with Gaussian codebooks.

Discontinuity, or *phase transition,* at $\mathrm{snr}=\mathrm{snr}_0$

PRINCETON UNIVERSITY

N. Merhav, D. Guo, and S. Shamai, "Statistical physics of signal estimation in Gaussian noise: Theory and examples of phase transitions," *IEEE Trans. Inf. Theory*, vol. 56, no. 3, pp. 1400–1416, March 2010.

Bounds on max-MMSE

$$\mathrm{M}_\infty(\mathrm{snr}, \mathrm{snr}_0, \beta) = \begin{cases} \frac{1}{1+\mathrm{snr}}, & \mathrm{snr} < \mathrm{snr}_0, \\ \frac{\beta}{1+\beta \mathrm{snr}}, & \mathrm{snr} \geq \mathrm{snr}_0, \end{cases} .$$

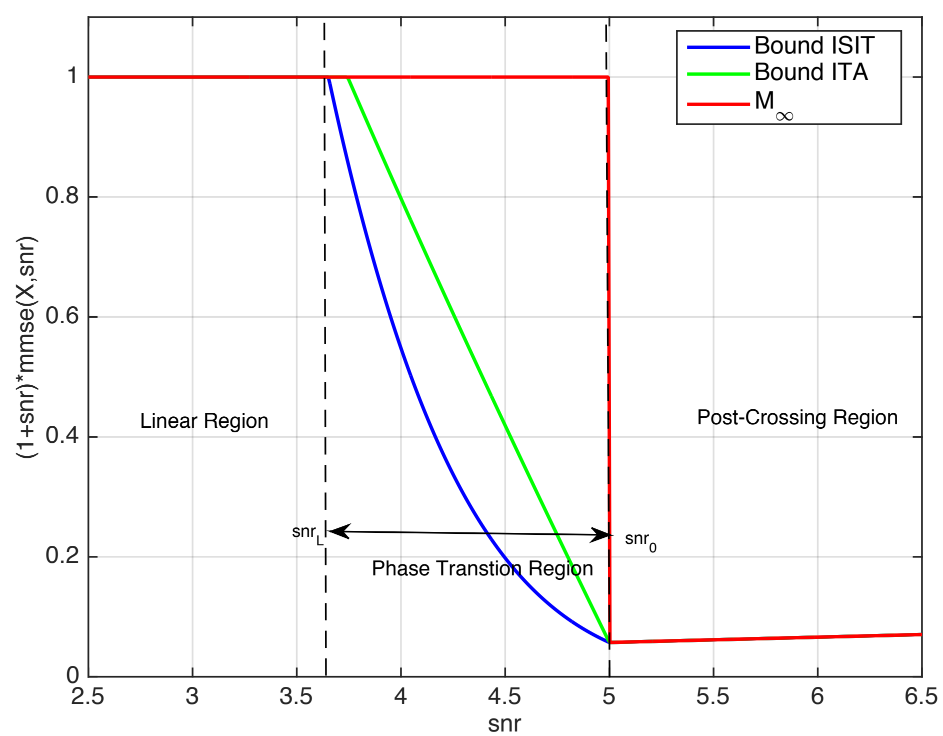
Discontinuity, or *phase transition,* at snr=snr$_0$

For Finite *n*

$$\mathrm{M}_n(\mathrm{snr}, \mathrm{snr}_0, \beta) = \begin{cases} \frac{1}{1+\mathrm{snr}}, & \mathrm{snr} \leq \mathrm{snr}_L, \\ T_n(\mathrm{snr}, \mathrm{snr}_0, \beta), & \mathrm{snr}_L \leq \mathrm{snr} \leq \mathrm{snr}_0, \\ \frac{\beta}{1+\beta \mathrm{snr}}, & \mathrm{snr}_0 \leq \mathrm{snr}, \end{cases}$$

***phase transition* region**

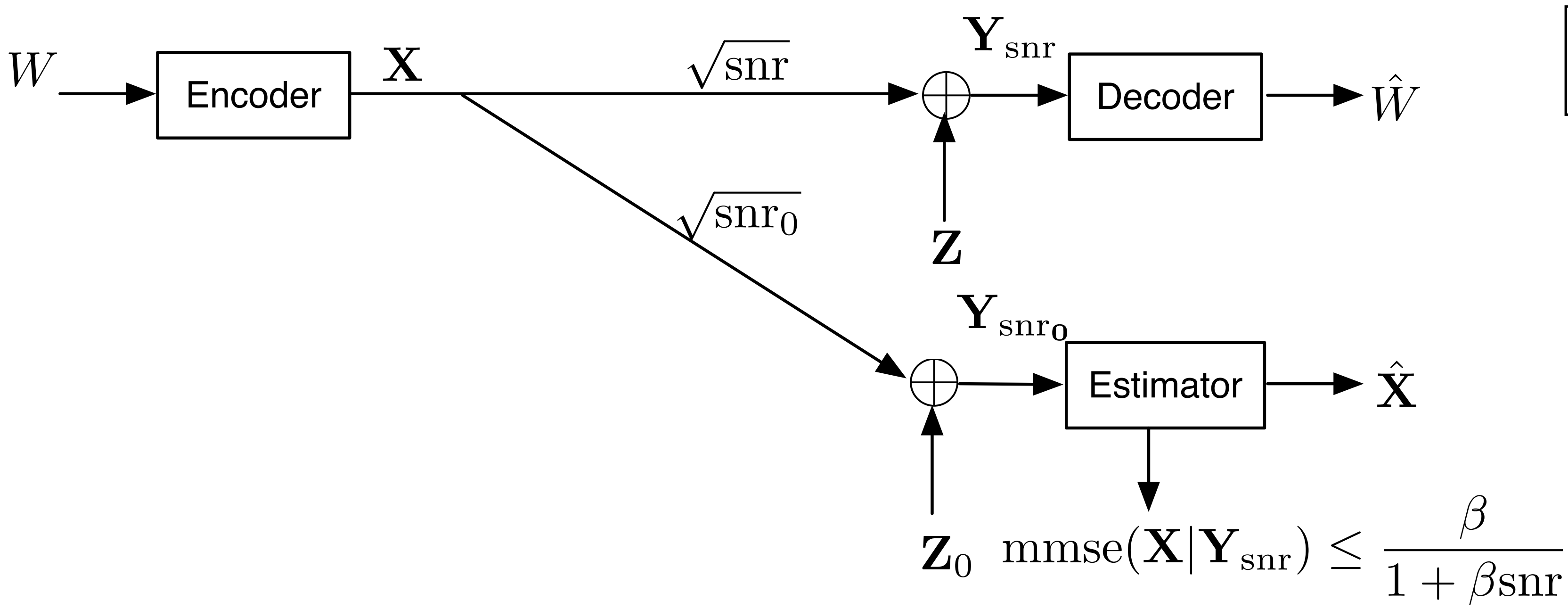
for some snr_L.



PRINCETON UNIVERSITY

Communication with the Disturbance Constraint

R. Bustin and S. Shamai, "MMSE of 'bad' codes," *IEEE Trans. Inf. Theory*, vol. 59, no. 2, pp. 733–743, Feb 2013.



max-I problem

$$\mathcal{C}_n(\mathrm{snr}, \mathrm{snr}_0, \beta) := \sup_{\mathbf{X}} \frac{1}{n} I(\mathbf{X};\mathbf{Y}),$$

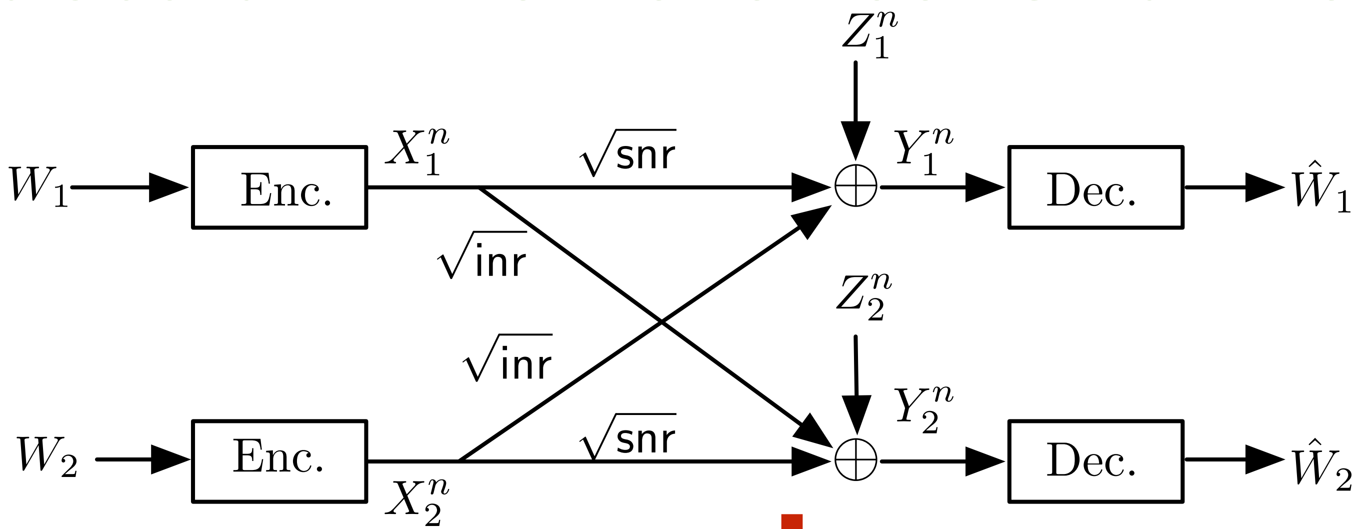
$$\mathrm{s.t.} \mathrm{Tr}\left(\mathbb{E}\left[\mathbf{X}\mathbf{X}^T\right]\right) \le n,$$

$$\text{and } \mathrm{mmse}(\mathbf{X}, \mathrm{snr}_0) \le \frac{\beta}{1+\beta \mathrm{snr}_0}.$$

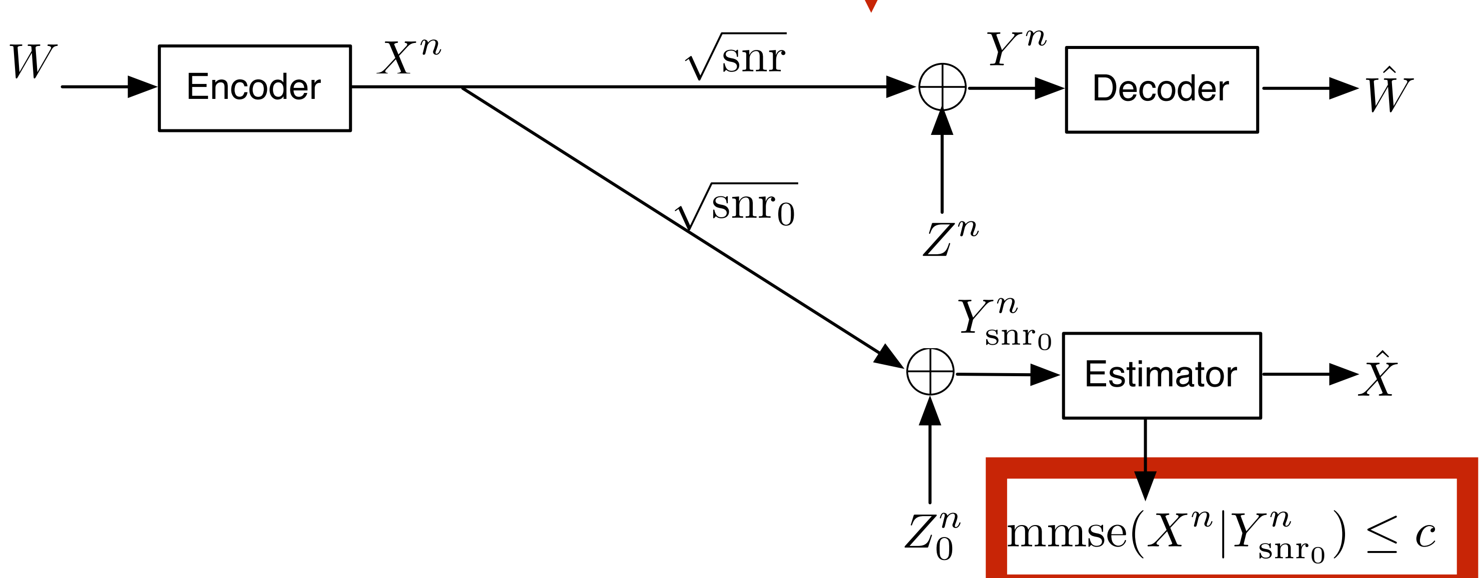
S. Shamai, "From constrained signaling to network interference alignment via an information-estimation perspective," IEEE Information Theory Society Newsletter, vol. 62, no. 7, pp. 6–24, September 2012.

Motivation

Gaussian Interference Channel



Simplified Model



MMSE Constraint captures the undecodable interference

PRINCETON UNIVERSITY

S. Shamai, "From constrained signaling to network interference alignment via an information-estimation perspective," IEEE Information Theory Society Newsletter, vol. 62, no. 7, pp. 6–24, September 2012.

max-MMSE problem

max-MMSE problem

$$\mathrm{M}_n(\mathrm{snr}, \mathrm{snr}_0, \beta) := \sup_{\mathbf{X}} \mathrm{mmse}(\mathbf{X}, \mathrm{snr}),$$
$$\text{s.t. } \frac{1}{n}\mathrm{Tr}\left(\mathbb{E}[\mathbf{X}\mathbf{X}^{\mathrm{T}}]\right) \le 1,$$
$$\text{and } \mathrm{mmse}(\mathbf{X}, \mathrm{snr}_0) \le \frac{\beta}{1+\beta \mathrm{snr}_0}.$$

max-I problem

$$\mathcal{C}_n(\mathrm{snr}, \mathrm{snr}_0, \beta) := \sup_{\mathbf{X}} \frac{1}{n} I(\mathbf{X}; \mathbf{Y}),$$
$$\text{s.t.} \mathrm{Tr}\left(\mathbb{E}\left[\mathbf{X}\mathbf{X}^{T}\right]\right) \le n,$$
$$\text{and } \mathrm{mmse}(\mathbf{X}, \mathrm{snr}_0) \le \frac{\beta}{1+\beta \mathrm{snr}_0}.$$

Relationship to max-I problem

$$\frac{\mathrm{snr}}{2}\mathrm{M}_n(\mathrm{snr}, \mathrm{snr}_0, \beta) \le \mathcal{C}_n(\mathrm{snr}, \mathrm{snr}_0, \beta) \le \frac{1}{2}\int_0^{\mathrm{snr}} \mathrm{M}_n(t, \mathrm{snr}_0, \beta) dt.$$

PRINCETON UNIVERSITY

Summary

- Improve OW Bound
- Bounds on the MMSE
- max-MMSE
- max-I

PRINCETON UNIVERSITY

Part 1b

Minimum Mean p-th Error

PRINCETON UNIVERSITY

MMPE

The p-norm of a random vector $\mathbf{U} \in \mathbb{R}^n$ is given by

$$\|\mathbf{U}\|_p^p = \frac{1}{n}\mathrm{E}\left[\mathrm{Tr}^{\frac{p}{2}}\left(\mathbf{U}\mathbf{U}^T\right)\right].$$

For $n = 1$, $\|U\|_p = \mathrm{E}[|U|^p]$.

Example: $\mathbf{U} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$

$$n\|\mathbf{U}\|_p^p = 2^{\frac{p}{2}} \frac{\Gamma\left(\frac{n+p}{2}\right)}{\Gamma\left(\frac{n}{2}\right)},$$

$$\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt.$$

PRINCETON UNIVERSITY

MMPE

We define the Minimum Mean p-th Error (MMPE) of estimating $\mathbf{X}$ from $\mathbf{Y}$ as

$$\boxed{\begin{gathered}\mathrm{mmpe}(\mathbf{X}|\mathbf{Y};p) := \inf_f \|\mathbf{X} - f(\mathbf{Y})\|_p^p = \inf_f \mathbb{E}\left[\mathsf{Err}^{\frac{p}{2}}(\mathbf{X}, f(\mathbf{Y}))\right], \\ \mathsf{Err} = \sqrt{(\mathbf{X} - f(\mathbf{Y}))^T(\mathbf{X} - f(\mathbf{Y}))}.\end{gathered}}$$

We denote the optimal estimator as $\boxed{f_p(\mathbf{X}|\mathbf{Y}).}$
For $n = 1$:

$$\mathrm{mmpe}(X|Y;p) = \inf_f \mathbb{E}\left[|X - f(Y)|^p\right].$$

Example: for $p = 2$

$$\mathrm{mmpe}(\mathbf{X}|\mathbf{Y};p=2) = \mathrm{mmse}(\mathbf{X}|\mathbf{Y}),$$
$$f_{p=2}(\mathbf{X}|\mathbf{Y}) = \mathbb{E}[\mathbf{X}|\mathbf{Y}].$$

We shall denote

Study Properties

$$\boxed{\mathrm{mmpe}(\mathbf{X}|\mathbf{Y};p) = \mathrm{mmpe}(\mathbf{X}, \mathrm{snr}, p).}$$

Properties of the MMPE

Optimal Estimator

Theorem. For $\text{mmpe}(\mathbf{X}, \text{snr}, p)$ and $p \geq 0$ the optimal estimator is given by the following point-wise relationship:

$$f_p(\mathbf{X}|\mathbf{Y}=\mathbf{y}) = \arg\min_{\mathbf{v}\in\mathbb{R}^n} \mathrm{E}\left[\mathsf{Err}^{\frac{p}{2}}(\mathbf{X},\mathbf{v})|\mathbf{Y}=\mathbf{y}\right].$$

Orthogonality-like Property

Theorem. For any $1 \leq p < \infty$ optimal estimator $f_p(\mathbf{X}|\mathbf{Y})$ satisfies

$$\mathrm{E}\left[\left(\mathbf{W}^T\mathbf{W}\right)^{\frac{p-2}{2}} \cdot \mathbf{W}^T \cdot g(\mathbf{Y})\right] = 0,$$

where $\mathbf{W} = \mathbf{X} - f_p(\mathbf{X}|\mathbf{Y})$ for any deterministic function $g(\cdot)$.

PRINCETON UNIVERSITY

Properties of Optimal Estimators

$\mathrm{E}[X|Y]$

1. if $0 \leq X \in \mathbb{R}^1$ then $0 \leq \mathrm{E}[X|Y]$,
2. (Linearity) $\mathrm{E}[a\mathbf{X} + b|\mathbf{Y}] = a\mathrm{E}[\mathbf{X}|\mathbf{Y}] + b$,
3. (Stability) $\mathrm{E}[g(\mathbf{Y})|\mathbf{Y}] = g(\mathbf{Y})$ for any function $g(\cdot)$,
4. (Idempotent) $\mathrm{E}[\mathrm{E}[\mathbf{X}|\mathbf{Y}]|\mathbf{Y}] = \mathrm{E}[\mathbf{X}|\mathbf{Y}]$,
5. (Degradedness)0 $\mathrm{E}\left[\mathbf{X}|\mathbf{Y}_{\mathrm{snr}}, \mathbf{Y}_{\mathrm{snr}_0}\right] = \mathrm{E}\left[\mathbf{X}|\mathbf{Y}_{\mathrm{snr}_0}\right]$, for $\mathbf{X} \to \mathbf{Y}_{\mathrm{snr}_0} \to \mathbf{Y}_{\mathrm{snr}}$,
6. (Shift Invariance) $\mathrm{mmse}(\mathbf{X} + a, \mathrm{snr}) = \mathrm{mmse}(\mathbf{X}, \mathrm{snr})$,
7. (Scaling) $\mathrm{mmse}(a\mathbf{X}, \mathrm{snr}) = a^2 \mathrm{mmse}(\mathbf{X}, a^2\mathrm{snr})$.

$$\mathrm{E}\left[\mathrm{E}\left[X|Y\right]\right] = \mathrm{E}\left[X\right]$$

$f_p(X|Y)$

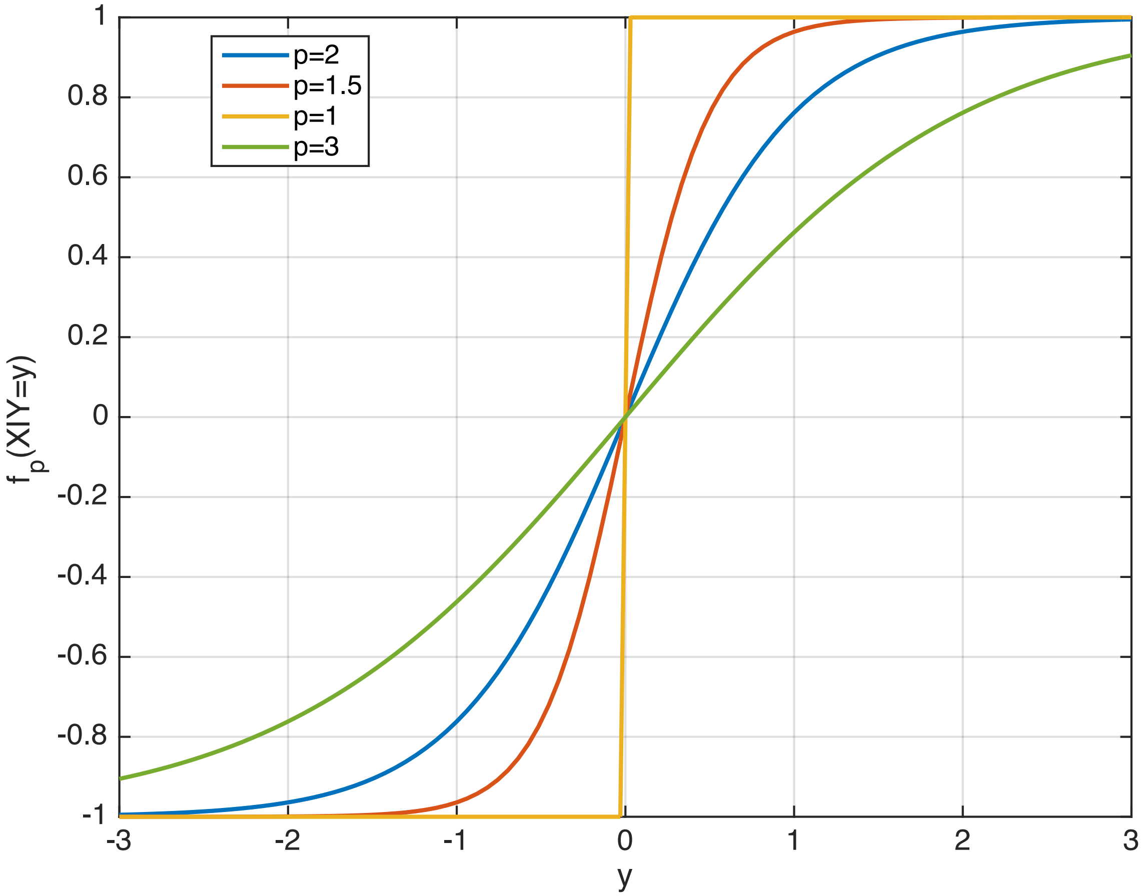
1. if $0 \leq X \in \mathbb{R}^1$ then $0 \leq f_p(X|Y)$,
2. (Linearity) $f_p(a\mathbf{X} + b|\mathbf{Y}) = af_p(\mathbf{X}|\mathbf{Y}) + b$,
3. (Stability) $f_p(g(\mathbf{Y})|\mathbf{Y}) = g(\mathbf{Y})$ for any function $g(\cdot)$,
4. (Idempotent) $f_p(f_p(\mathbf{X}|\mathbf{Y})|\mathbf{Y}) = f_p(\mathbf{X}|\mathbf{Y})$,
5. (Degradedness) $f_p\left(\mathbf{X}|\mathbf{Y}_{\mathrm{snr}_0}, \mathbf{Y}_{\mathrm{snr}}\right) = f_p\left(\mathbf{X}|\mathbf{Y}_{\mathrm{snr}_0}\right)$, for $\mathbf{X} \to \mathbf{Y}_{\mathrm{snr}_0} \to \mathbf{Y}_{\mathrm{snr}}$,
6. (Shift) $\mathrm{mmpe}(\mathbf{X} + a, \mathrm{snr}, p) = \mathrm{mmpe}(\mathbf{X}, \mathrm{snr}, p)$,
7. (Scaling) $\mathrm{mmpe}(a\mathbf{X}, \mathrm{snr}, p) = a^p \mathrm{mmpe}(\mathbf{X}, a^2\mathrm{snr}, p)$.

$$\cancel{\mathrm{E}\left[f_p\left(X|Y\right)\right] = \mathrm{E}\left[X\right]}$$

PRINCETON
UNIVERSITY

Examples of Optimal Estimators

- if $\mathbf{X} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ then for $p \geq 1$ $f_p(\mathbf{X}|\mathbf{Y}=\mathbf{y}) = \frac{\sqrt{\text{snr}}}{1+\text{snr}}\mathbf{y}$ $\boxed{\text{(i.e., linear)}}$;
- if $X = \{\pm 1\}$ equally likely (BPSK) then for $p \geq 1$ $f_p(X|Y) = \tanh\left(\frac{\sqrt{\text{snr}}}{p-1}y\right)$.



Function of p

S. Sherman, "Non-mean-square error criteria," *IRE Transactions on Information Theory*, vol. 3, no. 4, pp. 125–126, 1958.

PRINCETON UNIVERSITY

Bounds on the MMPE

$$\mathrm{mmse}(\mathbf{X}, \mathrm{snr}) \le \min\left(\frac{1}{\mathrm{snr}}, \|\mathbf{X}\|_2^2\right)$$

Theorem. For $\mathrm{snr} \ge 0$, $0 < q \le p$, and input $\mathbf{X}$

$$\mathrm{mmpe}(\mathbf{X}, \mathrm{snr}, p) \le \min\left(\frac{\|\mathbf{Z}\|_p^p}{\mathrm{snr}^{\frac{p}{2}}}, \|\mathbf{X}\|_p^p\right),$$

$$n^{\frac{p}{q}-1}\mathrm{mmpe}^{\frac{p}{q}}(\mathbf{X}, \mathrm{snr}, p) \le \mathrm{mmpe}(\mathbf{X}, \mathrm{snr}, p).$$

PRINCETON UNIVERSITY

MMPE with Gaussian Input

(Estimation of Gaussian Input.) For $\mathbf{X}_{\mathrm{G}} \sim \mathcal{N}(0, \sigma^2 \mathbf{I})$ and $p \geq 1$

$$\mathrm{mmpe}(\mathbf{X}_{\mathrm{G}}, \mathrm{snr}, p) = \frac{\sigma^p \|\mathbf{Z}\|_p^p}{(1 + \mathrm{snr}\sigma^2)^{\frac{p}{2}}},$$

with the optimal estimator given by

$$f_p(\mathbf{X}_{\mathrm{G}} | \mathbf{Y} = \mathbf{y}) = \frac{\sigma^2 \sqrt{\mathrm{snr}}}{1 + \mathrm{snr}\sigma^2} \mathbf{y}.$$

(Asymptotic Optimality of Gaussian Input.) For every $p \geq 1$, and a random variable $\mathbf{X}$ such that $\|\mathbf{X}\|_p^p \leq \sigma^p \|\mathbf{Z}\|_p^p$, we have

$$\mathrm{mmpe}(\mathbf{X}, \mathrm{snr}, p) \leq k_{p,\sigma^2 \mathrm{snr}} \cdot \frac{\sigma^p \|\mathbf{Z}\|_p^p}{(1 + \mathrm{snr}\sigma^2)^{\frac{p}{2}}}.$$

where

$$\begin{cases} k_{p,\sigma^2 \mathrm{snr}} = 1 & p = 2, \\ 1 \leq k_{p,\mathrm{snr}}^{\frac{1}{p}} = \frac{1+\sqrt{\mathrm{snr}}}{\sqrt{1+\mathrm{snr}}} \leq 1 + \frac{1}{\sqrt{1+\mathrm{snr}}} & p \neq 2. \end{cases}$$

PRINCETON UNIVERSITY

Conditional MMPE

We define the conditional MMPE as

$$\mathrm{mmpe}(\mathbf{X}, \mathrm{snr}, p|\mathbf{U}) := \|\mathbf{X} - f_p(\mathbf{X}|\mathbf{Y}_{\mathrm{snr}}, \mathbf{U})\|_p^p.$$

Additional Information

(Conditioning Reduces MMPE.) For every $\mathrm{snr} \geq 0$, $p \geq 0$, and random variable $\mathbf{X}$, we have

$$\mathrm{mmpe}(\mathbf{X}, \mathrm{snr}, p) \geq \mathrm{mmpe}(\mathbf{X}, \mathrm{snr}, p|\mathbf{U}).$$

PRINCETON UNIVERSITY

Bounds on the MMPE

(Extra Independent Noisy Observation.) Let $\mathbf{U} = \sqrt{\Delta} \cdot \mathbf{X} + \mathbf{Z}_\Delta$ where $\mathbf{Z}_\Delta \sim \mathcal{N}(0, \mathbf{I})$ and where $(\mathbf{X}, \mathbf{Z}, \mathbf{Z}_\Delta)$ are mutually independent. Then

$$\mathrm{mmpe}(\mathbf{X}, \mathrm{snr}, p|\mathbf{U}) = \mathrm{mmpe}(\mathbf{X}, \mathrm{snr} + \Delta, p).$$

Consequences:

- MMPE is decreasing function of SNR
- New Proof of the Single-Crossing-Point Property

PRINCETON UNIVERSITY

Part 2

Applications

PRINCETON
UNIVERSITY

Part 2a

Simple proof of SCPP

PRINCETON
UNIVERSITY

MMSE Bounds: Single Crossing Point Property

(**SCPP.**) For any fixed $\mathbf{X}$, suppose that $\mathrm{mmse}(\mathbf{X}, \mathrm{snr}_0) = \frac{\beta}{1+\beta \mathrm{snr}_0}$, for some fixed $\beta \geq 0$. Then

$$\mathrm{mmse}(\mathbf{X}, \mathrm{snr}) \leq \frac{\beta}{1+\beta \mathrm{snr}}, \quad \mathrm{snr} \in [\mathrm{snr}_0, \infty)$$

$$\mathrm{mmse}(\mathbf{X}, \mathrm{snr}) \geq \frac{\beta}{1+\beta \mathrm{snr}}, \quad \mathrm{snr} \in [0, \mathrm{snr}_0).$$

SCPP+I-MMSE important tool for derive converses:

- Version of EPI
- BC
- Wiretap
- MIMO channels

D. Guo, Y. Wu, S. Shamai, and S. Verdu, "Estimation in Gaussian noise: Properties of the minimum mean-square error," IEEE Trans. Inf. Theory, vol. 57, no. 4, pp. 2371–2385, April 2011.

R. Bustin, R. Schaefer, H. Poor, and S. Shamai, "On MMSE properties of optimal codes for the Gaussian wiretap channel," in Proc. IEEE Inf. Theory Workshop, April 2015, pp. 1–5.

R. Bustin, M. Payaro, D. Palomar, and S. Shamai, "On MMSE crossing properties and implications in parallel vector Gaussian channels," IEEE Trans. Inf. Theory, vol. 59, no. 2, pp. 818–844, Feb 2013.

New Proof of the SCPP

$m := \mathrm{mmpe}^{\frac{2}{p}}(\mathbf{X}, \mathrm{snr}_0, p) = \frac{\beta \|\mathbf{Z}\|_p^2}{1+\beta \mathrm{snr}_0}$ for some $\beta \geq 0$, $\Leftrightarrow$, $\beta = \frac{m}{\|\mathbf{Z}\|_p^2 - \mathrm{snr}_0 m}$.

Proof: $\mathrm{mmpe}^{\frac{1}{p}}(\mathbf{X}, \mathrm{snr}, p) = \mathrm{mmpe}^{\frac{1}{p}}(\mathbf{X}, \mathrm{snr}_0 + \Delta, p)$

PRINCETON UNIVERSITY

New Proof of the SCPP

$m := \mathrm{mmpe}^{\frac{2}{p}}(\mathbf{X}, \mathrm{snr}_0, p) = \frac{\beta \|\mathbf{Z}\|_p^2}{1+\beta \mathrm{snr}_0}$ for some $\beta \geq 0$, $\Leftrightarrow$, $\beta = \frac{m}{\|\mathbf{Z}\|_p^2 - \mathrm{snr}_0 m}$.

Proof:

$$\begin{aligned}\mathrm{mmpe}^{\frac{1}{p}}(\mathbf{X}, \mathrm{snr}, p) &= \mathrm{mmpe}^{\frac{1}{p}}(\mathbf{X}, \mathrm{snr}_0 + \Delta, p) \\ &= \mathrm{mmpe}^{\frac{1}{p}}(\mathbf{X}, \mathrm{snr}_0, p | \mathbf{Y}_\Delta), \text{ where } \mathbf{Y}_\Delta = \sqrt{\Delta}\mathbf{X} + \mathbf{Z}_\Delta\end{aligned}$$

PRINCETON UNIVERSITY

New Proof of the SCPP

$m := \mathrm{mmpe}^{\frac{2}{p}}(\mathbf{X}, \mathrm{snr}_0, p) = \frac{\beta \|\mathbf{Z}\|_p^2}{1+\beta \mathrm{snr}_0}$ for some $\beta \geq 0$, $\Leftrightarrow$, $\beta = \frac{m}{\|\mathbf{Z}\|_p^2 - \mathrm{snr}_0 m}$.

Proof:

$$\begin{aligned} \mathrm{mmpe}^{\frac{1}{p}}(\mathbf{X}, \mathrm{snr}, p) &= \mathrm{mmpe}^{\frac{1}{p}}(\mathbf{X}, \mathrm{snr}_0 + \Delta, p) \\ &= \mathrm{mmpe}^{\frac{1}{p}}(\mathbf{X}, \mathrm{snr}_0, p | \mathbf{Y}_\Delta), \text{ where } \mathbf{Y}_\Delta = \sqrt{\Delta}\mathbf{X} + \mathbf{Z}_\Delta \\ &= \|\mathbf{X} - f_p(\mathbf{X} | \mathbf{Y}_\Delta, \mathbf{Y}_{\mathrm{snr}_0})\|_p \end{aligned}$$

PRINCETON UNIVERSITY

New Proof of the SCPP

$m := \mathrm{mmpe}^{\frac{2}{p}}(\mathbf{X}, \mathrm{snr}_0, p) = \frac{\beta \|\mathbf{Z}\|_p^2}{1+\beta \mathrm{snr}_0}$ for some $\beta \geq 0$, $\Leftrightarrow$, $\beta = \frac{m}{\|\mathbf{Z}\|_p^2 - \mathrm{snr}_0 m}$.

Proof:

$$\begin{aligned}
\mathrm{mmpe}^{\frac{1}{p}}(\mathbf{X}, \mathrm{snr}, p) &= \mathrm{mmpe}^{\frac{1}{p}}(\mathbf{X}, \mathrm{snr}_0 + \Delta, p) \\
&= \mathrm{mmpe}^{\frac{1}{p}}(\mathbf{X}, \mathrm{snr}_0, p|\mathbf{Y}_\Delta), \text{ where } \mathbf{Y}_\Delta = \sqrt{\Delta}\mathbf{X} + \mathbf{Z}_\Delta \\
&= \|\mathbf{X} - f_p(\mathbf{X}|\mathbf{Y}_\Delta, \mathbf{Y}_{\mathrm{snr}_0})\|_p \\
&\leq \|\mathbf{X} - \hat{\mathbf{X}}\|_p, \text{ where } \hat{\mathbf{X}} = \frac{(1-\gamma)}{\sqrt{\Delta}}\mathbf{Y}_\Delta + \gamma f_p(\mathbf{X}|\mathbf{Y}_{\mathrm{snr}_0}), \gamma \in [0,1]
\end{aligned}$$

New Proof of the SCPP

$m := \mathrm{mmpe}^{\frac{2}{p}}(\mathbf{X}, \mathrm{snr}_0, p) = \frac{\beta \|\mathbf{Z}\|_p^2}{1+\beta \mathrm{snr}_0}$ for some $\beta \geq 0$, $\Leftrightarrow$, $\beta = \frac{m}{\|\mathbf{Z}\|_p^2 - \mathrm{snr}_0 m}$.

Proof:

$$\begin{aligned}
\mathrm{mmpe}^{\frac{1}{p}}(\mathbf{X}, \mathrm{snr}, p) &= \mathrm{mmpe}^{\frac{1}{p}}(\mathbf{X}, \mathrm{snr}_0 + \Delta, p) \\
&= \mathrm{mmpe}^{\frac{1}{p}}(\mathbf{X}, \mathrm{snr}_0, p|\mathbf{Y}_\Delta), \text{ where } \mathbf{Y}_\Delta = \sqrt{\Delta}\mathbf{X} + \mathbf{Z}_\Delta \\
&= \|\mathbf{X} - f_p(\mathbf{X}|\mathbf{Y}_\Delta, \mathbf{Y}_{\mathrm{snr}_0})\|_p \\
&\leq \|\mathbf{X} - \hat{\mathbf{X}}\|_p, \text{ where } \hat{\mathbf{X}} = \frac{(1-\gamma)}{\sqrt{\Delta}}\mathbf{Y}_\Delta + \gamma f_p(\mathbf{X}|\mathbf{Y}_{\mathrm{snr}_0}), \gamma \in [0,1] \\
&= \left\| \gamma(\mathbf{X} - f_p(\mathbf{X}|\mathbf{Y}_{\mathrm{snr}_0})) - \frac{(1-\gamma)}{\sqrt{\Delta}}\mathbf{Z}_\Delta \right\|_p
\end{aligned}$$

PRINCETON UNIVERSITY

New Proof of the SCPP

$m := \mathrm{mmpe}^{\frac{2}{p}}(\mathbf{X}, \mathrm{snr}_0, p) = \frac{\beta \|\mathbf{Z}\|_p^2}{1+\beta \mathrm{snr}_0}$ for some $\beta \geq 0$, $\Leftrightarrow$, $\beta = \frac{m}{\|\mathbf{Z}\|_p^2 - \mathrm{snr}_0 m}$.

Proof:

$$\begin{aligned}
\mathrm{mmpe}^{\frac{1}{p}}(\mathbf{X}, \mathrm{snr}, p) &= \mathrm{mmpe}^{\frac{1}{p}}(\mathbf{X}, \mathrm{snr}_0 + \Delta, p) \\
&= \mathrm{mmpe}^{\frac{1}{p}}(\mathbf{X}, \mathrm{snr}_0, p | \mathbf{Y}_\Delta), \text{ where } \mathbf{Y}_\Delta = \sqrt{\Delta}\mathbf{X} + \mathbf{Z}_\Delta \\
&= \|\mathbf{X} - f_p(\mathbf{X}|\mathbf{Y}_\Delta, \mathbf{Y}_{\mathrm{snr}_0})\|_p \\
&\leq \|\mathbf{X} - \hat{\mathbf{X}}\|_p, \text{ where } \hat{\mathbf{X}} = \frac{(1-\gamma)}{\sqrt{\Delta}}\mathbf{Y}_\Delta + \gamma f_p(\mathbf{X}|\mathbf{Y}_{\mathrm{snr}_0}), \gamma \in [0,1] \\
&= \left\| \gamma(\mathbf{X} - f_p(\mathbf{X}|\mathbf{Y}_{\mathrm{snr}_0})) - \frac{(1-\gamma)}{\sqrt{\Delta}} \mathbf{Z}_\Delta \right\|_p \\
&= \frac{\left\| \|\mathbf{Z}\|_p^2 (\mathbf{X} - f_p(\mathbf{X}|\mathbf{Y}_{\mathrm{snr}_0})) - \sqrt{\Delta} \cdot m \cdot \mathbf{Z}_\Delta \right\|_p}{\|\mathbf{Z}\|_p^2 + \Delta \cdot m}, \text{ choose } \gamma = \frac{\|\mathbf{Z}\|_p^2}{\|\mathbf{Z}\|_p^2 + \Delta \cdot m}
\end{aligned}$$

PRINCETON UNIVERSITY

New Proof of the SCPP

$$m := \mathrm{mmpe}^{\frac{2}{p}}(\mathbf{X}, \mathrm{snr}_0, p) = \tfrac{\beta\|\mathbf{Z}\|_p^2}{1+\beta\mathrm{snr}_0} \text{ for some } \beta \geq 0, \Leftrightarrow, \beta = \tfrac{m}{\|\mathbf{Z}\|_p^2 - \mathrm{snr}_0 m}.$$

Proof:

$$\begin{aligned}
\mathrm{mmpe}^{\frac{1}{p}}(\mathbf{X}, \mathrm{snr}, p) &= \mathrm{mmpe}^{\frac{1}{p}}(\mathbf{X}, \mathrm{snr}_0 + \Delta, p) \\
&= \mathrm{mmpe}^{\frac{1}{p}}(\mathbf{X}, \mathrm{snr}_0, p|\mathbf{Y}_\Delta), \text{ where } \mathbf{Y}_\Delta = \sqrt{\Delta}\mathbf{X} + \mathbf{Z}_\Delta \\
&= \|\mathbf{X} - f_p(\mathbf{X}|\mathbf{Y}_\Delta, \mathbf{Y}_{\mathrm{snr}_0})\|_p \\
&\leq \|\mathbf{X} - \hat{\mathbf{X}}\|_p, \text{ where } \hat{\mathbf{X}} = \frac{(1-\gamma)}{\sqrt{\Delta}}\mathbf{Y}_\Delta + \gamma f_p(\mathbf{X}|\mathbf{Y}_{\mathrm{snr}_0}), \gamma \in [0,1] \\
&= \left\| \gamma(\mathbf{X} - f_p(\mathbf{X}|\mathbf{Y}_{\mathrm{snr}_0})) - \frac{(1-\gamma)}{\sqrt{\Delta}}\mathbf{Z}_\Delta \right\|_p \\
&= \frac{\left\| \|\mathbf{Z}\|_p^2(\mathbf{X} - f_p(\mathbf{X}|\mathbf{Y}_{\mathrm{snr}_0})) - \sqrt{\Delta} \cdot m \cdot \mathbf{Z}_\Delta \right\|_p}{\|\mathbf{Z}\|_p^2 + \Delta \cdot m}, \text{ choose } \gamma = \frac{\|\mathbf{Z}\|_p^2}{\|\mathbf{Z}\|_p^2 + \Delta \cdot m} \\
&\leq c_p \cdot \frac{\sqrt{m}\|\mathbf{Z}\|_p}{\sqrt{\|\mathbf{Z}\|_p^2 + \Delta \cdot m}}, c_p = \begin{cases} 2 & p \geq 1, \text{ triangle inequality} \\ 1 & p = 2, \text{ expand} \end{cases} \\
&= c_p \cdot \frac{\beta\|\mathbf{Z}\|_p^2}{1 + \beta \text{ snr}}
\end{aligned}$$

PRINCETON UNIVERSITY

New Proof of the SCPP

(**SCPP Bound for MMPE.**) Suppose $\mathrm{mmpe}^{\frac{2}{p}}(\mathbf{X}, \mathrm{snr}_0, p) = \frac{\beta \|\mathbf{Z}\|_p^2}{1+\beta \mathrm{snr}_0}$ for some $\beta \geq 0$. Then

$$\mathrm{mmpe}^{\frac{2}{p}}(\mathbf{X}, \mathrm{snr}, p) \leq c_p \cdot \frac{\beta \|\mathbf{Z}\|_p^2}{1+\beta\ \mathrm{snr}}, \text{ for } \mathrm{snr} \geq \mathrm{snr}_0,$$

where $c_p = \begin{cases} 2 & p \geq 1 \\ 1 & p = 2 \end{cases}.$

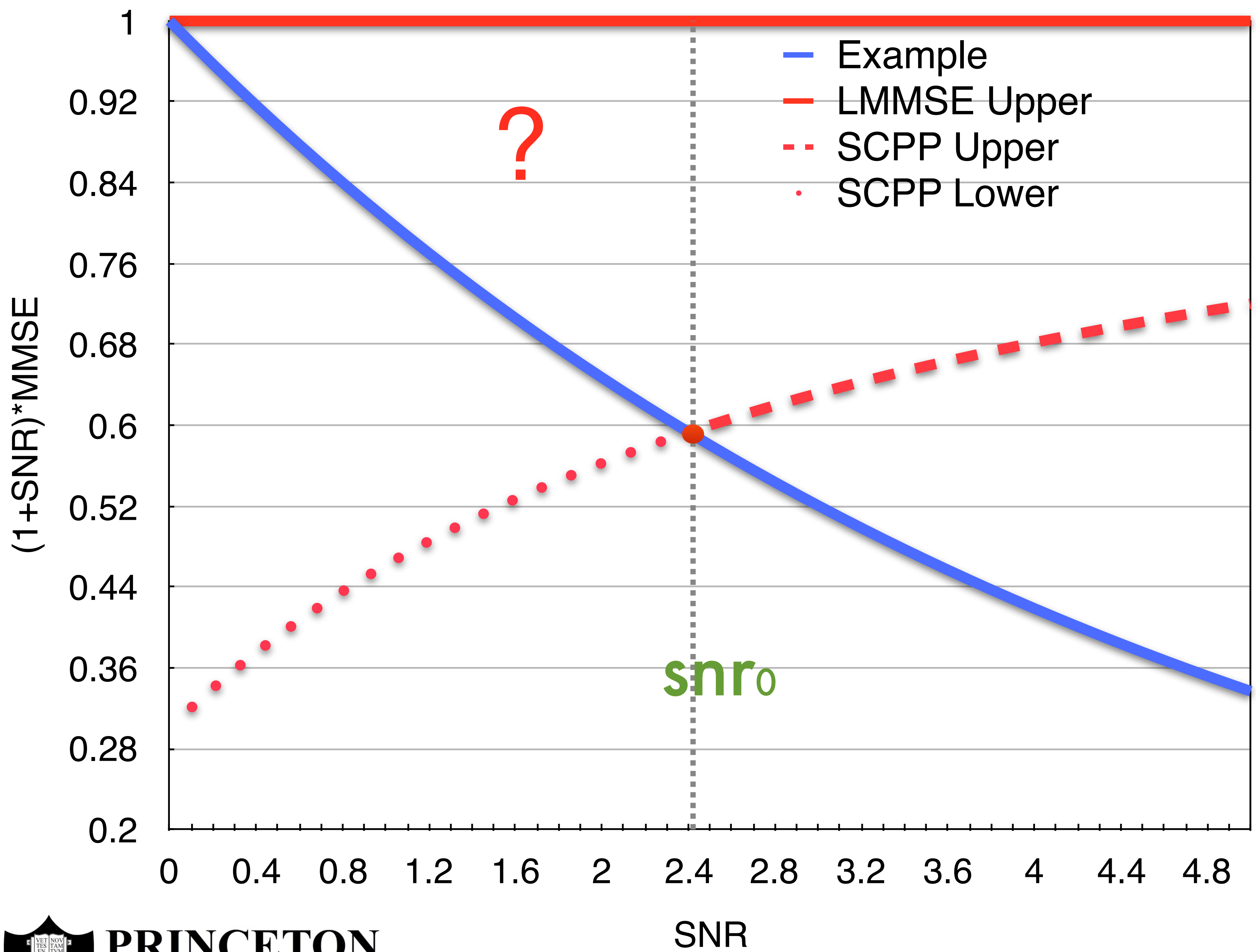
Observations:

- Previous proof relied on the knowing the derivative of the MMSE
- New proof relies on simple estimation observations
- Extends to the MMPE for all p>1

Part 2b

Complementary Bounds to the SCPP

Review: MMSE bounds



$$\mathrm{mmse}(\mathbf{X}, \mathrm{snr}_0) = \frac{\beta}{1 + \beta \mathrm{snr}_0}$$

LMMSE u.b.

$$\mathrm{mmse}(\mathbf{X}, \mathrm{snr}) \leq \frac{1}{1 + \mathrm{snr}}$$

SCPP u.b.

$$\mathrm{mmse}(\mathbf{X}, \mathrm{snr}) \leq \frac{\beta}{1 + \beta \mathrm{snr}}$$

$$\text{for } \mathrm{snr} \geq \mathrm{snr}_0$$

SCPP l.b.

$$\mathrm{mmse}(\mathbf{X}, \mathrm{snr}) \geq \frac{\beta}{1 + \beta \mathrm{snr}}$$

$$\text{for } \mathrm{snr} \leq \mathrm{snr}_0$$

PRINCETON UNIVERSITY

Bounds on the MMPE

Interpolation Bounds

Theorem. For any $0 < p < q < r \leq \infty$ and $\alpha \in (0,1)$ such that

$$\frac{1}{q} = \frac{\alpha}{p} + \frac{\bar{\alpha}}{r} \iff \alpha = \frac{q^{-1} - r^{-1}}{p^{-1} - r^{-1}}, \tag{1a}$$

where $\bar{\alpha} = 1 - \alpha$,

$$\mathrm{mmpe}^{\frac{1}{q}}(\mathbf{X}, \mathrm{snr}, q) \leq \inf_f \|\mathbf{X} - f(\mathbf{Y})\|_p^{\alpha} \|\mathbf{X} - f(\mathbf{Y})\|_r^{\bar{\alpha}}. \tag{1b}$$

In particular,

$$\mathrm{mmpe}^{\frac{1}{q}}(\mathbf{X}, \mathrm{snr}, q) \leq \|\mathbf{X} - f_r(\mathbf{X}|\mathbf{Y})\|_p^{\alpha}\ \mathrm{mmpe}^{\frac{\bar{\alpha}}{r}}(\mathbf{X}, \mathrm{snr}, r), \tag{1c}$$

$$\mathrm{mmpe}^{\frac{1}{q}}(\mathbf{X}, \mathrm{snr}, q) \leq \mathrm{mmpe}^{\frac{\alpha}{p}}(\mathbf{X}, \mathrm{snr}, p) \|\mathbf{X} - f_p(\mathbf{X}|\mathbf{Y})\|_r^{\bar{\alpha}}. \tag{1d}$$

PRINCETON UNIVERSITY

Bounds on The MMPE

Change of Measure Bound

Theorem. For $0 < \text{snr} \leq \text{snr}_0$, $\mathbf{X}$ and $p \geq 0$, we have

$$\text{mmpe}(\mathbf{X}, \text{snr}, p) \leq \kappa_{n,t}\ \text{mmpe}^{\frac{1-t}{1+t}}\left(\mathbf{X}, \text{snr}_0, \frac{1+t}{1-t} \cdot p\right),$$

$$\text{where } \kappa_{n,t} := \left(\frac{2^n}{n^2}\right)^{\frac{t}{t+1}} \left(\frac{1}{1-t}\right)^{\frac{nt}{t+1} - \frac{1}{2}},\ t = \frac{\text{snr}_0 - \text{snr}}{\text{snr}_0}.$$

Fix SNR But Change Order

PRINCETON UNIVERSITY

New Bound on max-MMSE

$$\begin{aligned} \mathrm{mmse}(\mathbf{X}, \mathrm{snr}) &= \mathrm{mmpe}(\mathbf{X}, \mathrm{snr}, 2) \\ &\leq \kappa_{n,t} \ \mathrm{mmpe}^{\frac{1-t}{1+t}} \left(\mathbf{X}, \mathrm{snr}_0, \frac{1+t}{1-t} \cdot 2 \right), \ \text{change of measure bound} \end{aligned}$$

PRINCETON
UNIVERSITY

New Bound on max-MMSE

$$
\begin{aligned}
\mathrm{mmse}(\mathbf{X}, \mathrm{snr}) &= \mathrm{mmpe}(\mathbf{X}, \mathrm{snr}, 2) \\
&\le \kappa_{n,t}\ \mathrm{mmpe}^{\frac{1-t}{1+t}}\left(\mathbf{X}, \mathrm{snr}_0, \frac{1+t}{1-t} \cdot 2\right),\ \text{change of measure bound} \\
&\le \mathrm{mmse}^{\alpha}(\mathbf{X}, \mathrm{snr}_0) \| \mathbf{X} - \mathbb{E}[\mathbf{X} | \mathbf{Y}_{\mathrm{snr}_0}] \|_r^{2(1-\alpha)},\ \text{interpolation bound.}
\end{aligned}
$$

PRINCETON
UNIVERSITY

New Bound on max-MMSE

$$
\begin{aligned}
\mathrm{mmse}(\mathbf{X}, \mathrm{snr}) &= \mathrm{mmpe}(\mathbf{X}, \mathrm{snr}, 2) \\
&\le \kappa_{n,t} \ \mathrm{mmpe}^{\frac{1-t}{1+t}} \left(\mathbf{X}, \mathrm{snr}_0, \frac{1+t}{1-t} \cdot 2 \right), \ \text{change of measure bound} \\
&\le \mathrm{mmse}^{\alpha}(\mathbf{X}, \mathrm{snr}_0) \| \mathbf{X} - \mathbb{E}[\mathbf{X} | \mathbf{Y}_{\mathrm{snr}_0}] \|_r^{2(1-\alpha)}, \ \text{interpolation bound.}
\end{aligned}
$$

Point where we know MMSE

PRINCETON UNIVERSITY

New Bound on max-MMSE

$$\begin{aligned}\mathrm{mmse}(\mathbf{X},\mathrm{snr}) &= \mathrm{mmpe}(\mathbf{X},\mathrm{snr},2)\\ &\le \kappa_{n,t}\ \mathrm{mmpe}^{\frac{1-t}{1+t}}\left(\mathbf{X},\mathrm{snr}_0,\frac{1+t}{1-t}\cdot 2\right),\ \text{change of measure bound}\\ &\le \mathrm{mmse}^{\alpha}(\mathbf{X},\mathrm{snr}_0)\|\mathbf{X}-\mathbb{E}[\mathbf{X}|\mathbf{Y}_{\mathrm{snr}_0}]\|_r^{2(1-\alpha)},\ \text{interpolation bound.}\end{aligned}$$

Point where we know MMSE

Bound with LMMSE type bound

PRINCETON UNIVERSITY

New Bound on max-MMSE

Theorem.

$$T_n(\mathrm{snr}, \mathrm{snr}_0, \beta) \leq \min_{r > \frac{2}{\gamma}} k_{n,\gamma,r} \cdot \mathrm{mmse}^{\frac{\gamma r - 2}{r-2}}(\mathbf{X}, \mathrm{snr}_0),$$

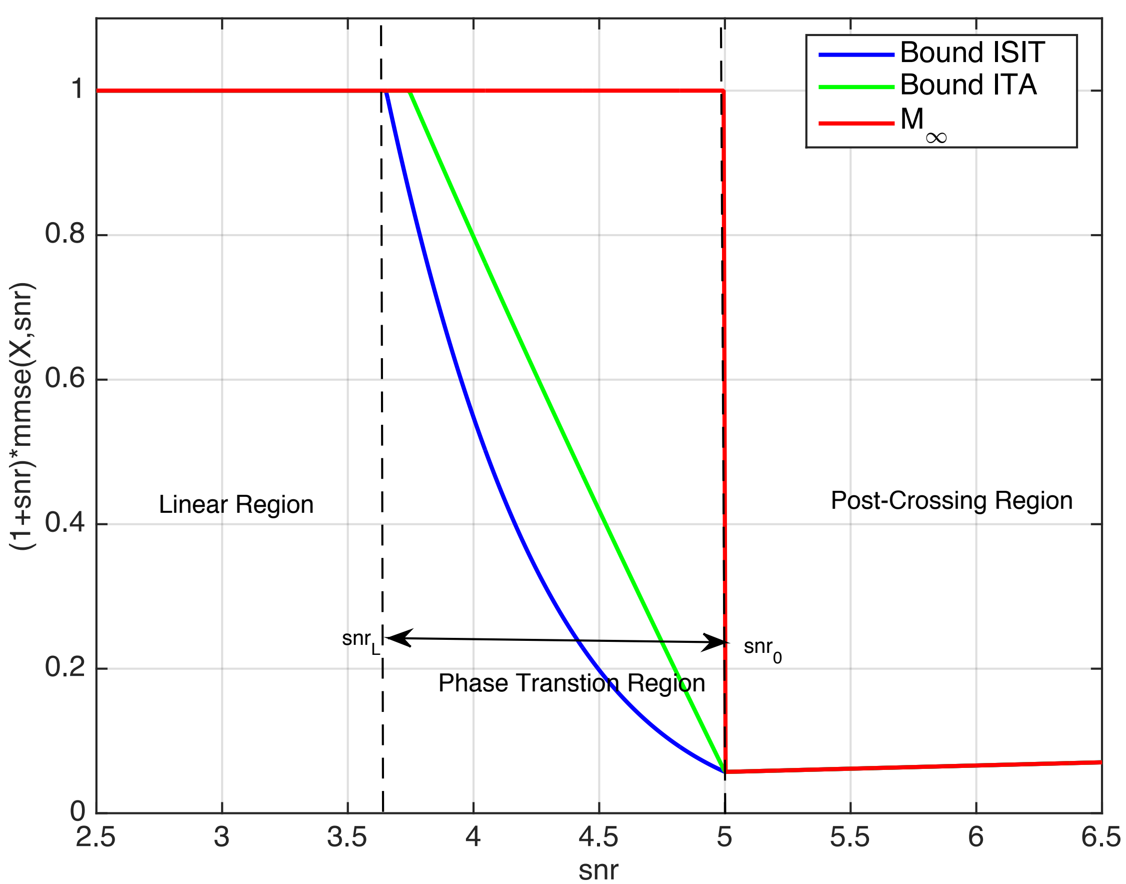
$$\gamma := \frac{\mathrm{snr}}{2\mathrm{snr}_0 - \mathrm{snr}} \in (0,1],$$

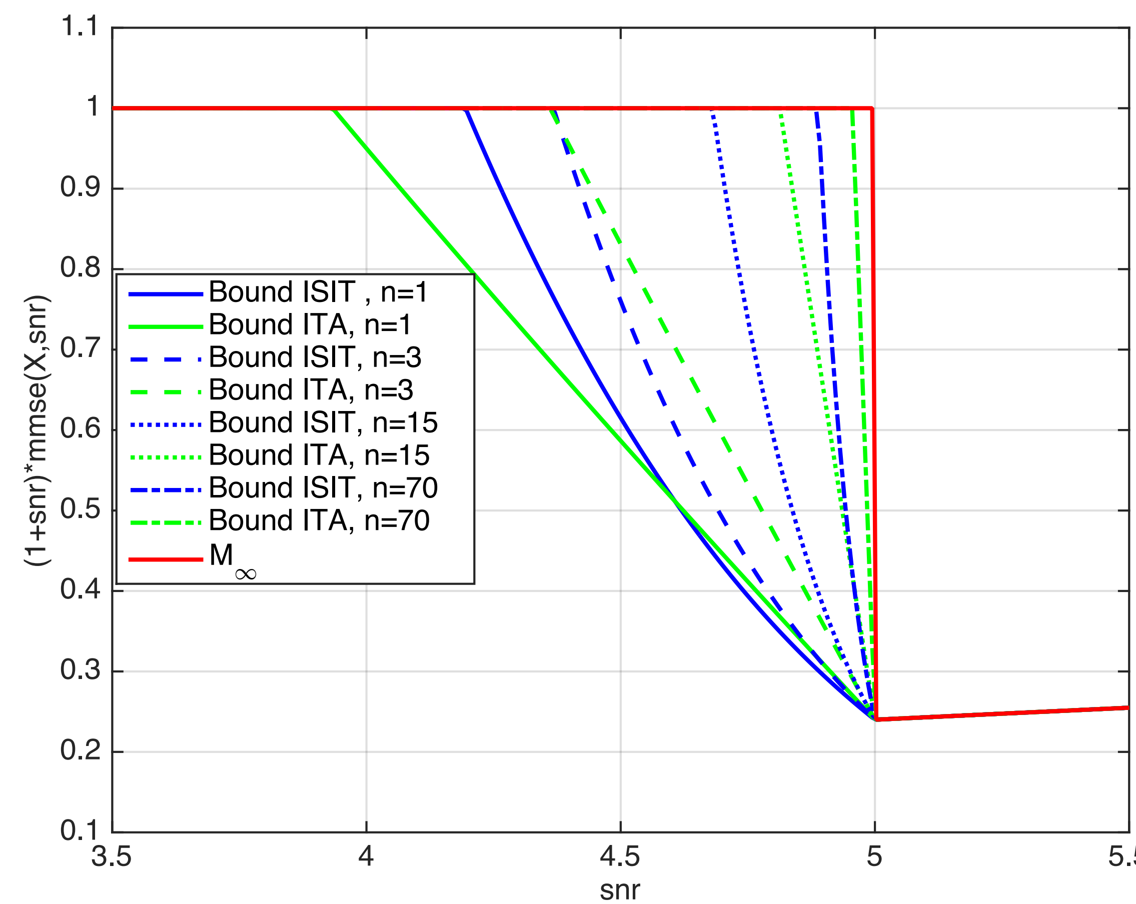
$$k_{n,\gamma,r} := \frac{\sqrt{2}}{n^{1-\gamma}} \left(\frac{1+\gamma}{\gamma}\right)^{\frac{n(1-\gamma)-1}{2}} \left(\frac{2^r \|\mathbf{Z}\|_r^r}{\mathrm{snr}_0^{\frac{r}{2}}}\right)^{\frac{2(1-\gamma)}{r-2}}.$$

Moreover, the width of the phase transition region is given by $W(n) = O\left(n^{-\frac{1}{2}}\right)$.

PRINCETON UNIVERSITY

New Bound on max-MMSE

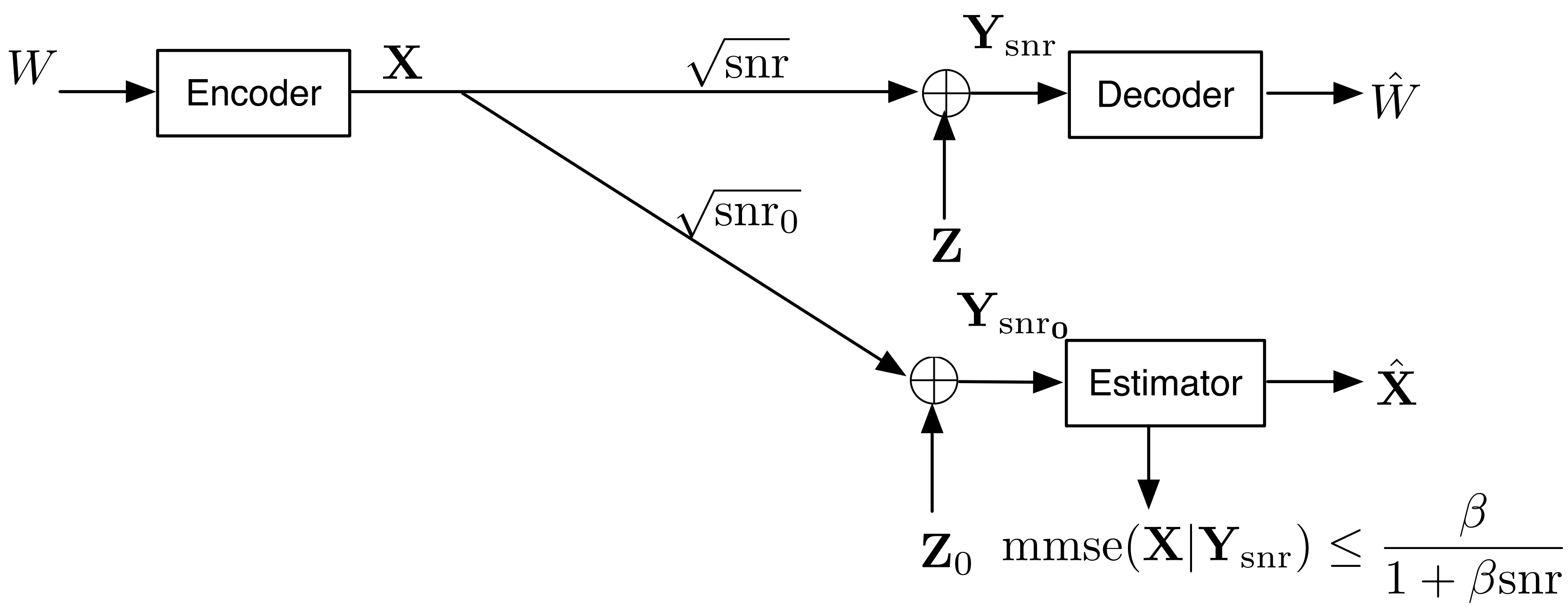




$snr_0=5$, beta=0.01, n=1

$snr_0=5$, beta=0.05

Communication with the Disturbance Constraint



max-I problem

$$\mathcal{C}_n(\mathrm{snr}, \mathrm{snr}_0, \beta) := \sup_{\mathbf{X}} \frac{1}{n} I(\mathbf{X}; \mathbf{Y}),$$

$$\text{s.t.} \operatorname{Tr}\left(\mathbb{E}\left[\mathbf{X}\mathbf{X}^T\right]\right) \le n,$$

$$\text{and } \mathrm{mmse}(\mathbf{X}, \mathrm{snr}_0) \le \frac{\beta}{1+\beta\mathrm{snr}_0}.$$

Relationship to max-I problem

$$\frac{\mathrm{snr}}{2} \mathrm{M}_n(\mathrm{snr}, \mathrm{snr}_0, \beta) \le \mathcal{C}_n(\mathrm{snr}, \mathrm{snr}_0, \beta) \le \frac{1}{2}\int_0^{\mathrm{snr}} \mathrm{M}_n(t, \mathrm{snr}_0, \beta) dt.$$

PRINCETON UNIVERSITY

Achievability

Mixed inputs

$$X_{\text{mix}} = \sqrt{1-\delta}X_D + \sqrt{\delta}X_G,$$
$$\delta \in [0,1],$$
$$X_G \sim \mathcal{N}(0,1)$$
$$\mathbb{E}[X_D^2] \le 1$$

Nice "Decomposition"

$$I(X_{\text{mix}}, \text{snr}) = I\left(X_D, \frac{\text{snr}(1-\delta)}{1+\delta\text{snr}}\right) + I(X_G, \text{snr}\ \delta),$$

$$\text{mmse}(X_{\text{mix}}, \text{snr}) = \frac{1-\delta}{(1+\text{snr}\delta)^2}\text{mmse}\left(X_D, \frac{\text{snr}(1-\delta)}{1+\delta\text{snr}}\right) + \delta\ \text{mmse}(X_G, \text{snr}\ \delta).$$

Bounds for X_D

$$\text{mmse}(X_D, \text{snr}) \le d_{\max}^2 \sum_{i=1}^{N} p_i \mathrm{e}^{-\frac{\text{snr}}{8} d_i^2},$$

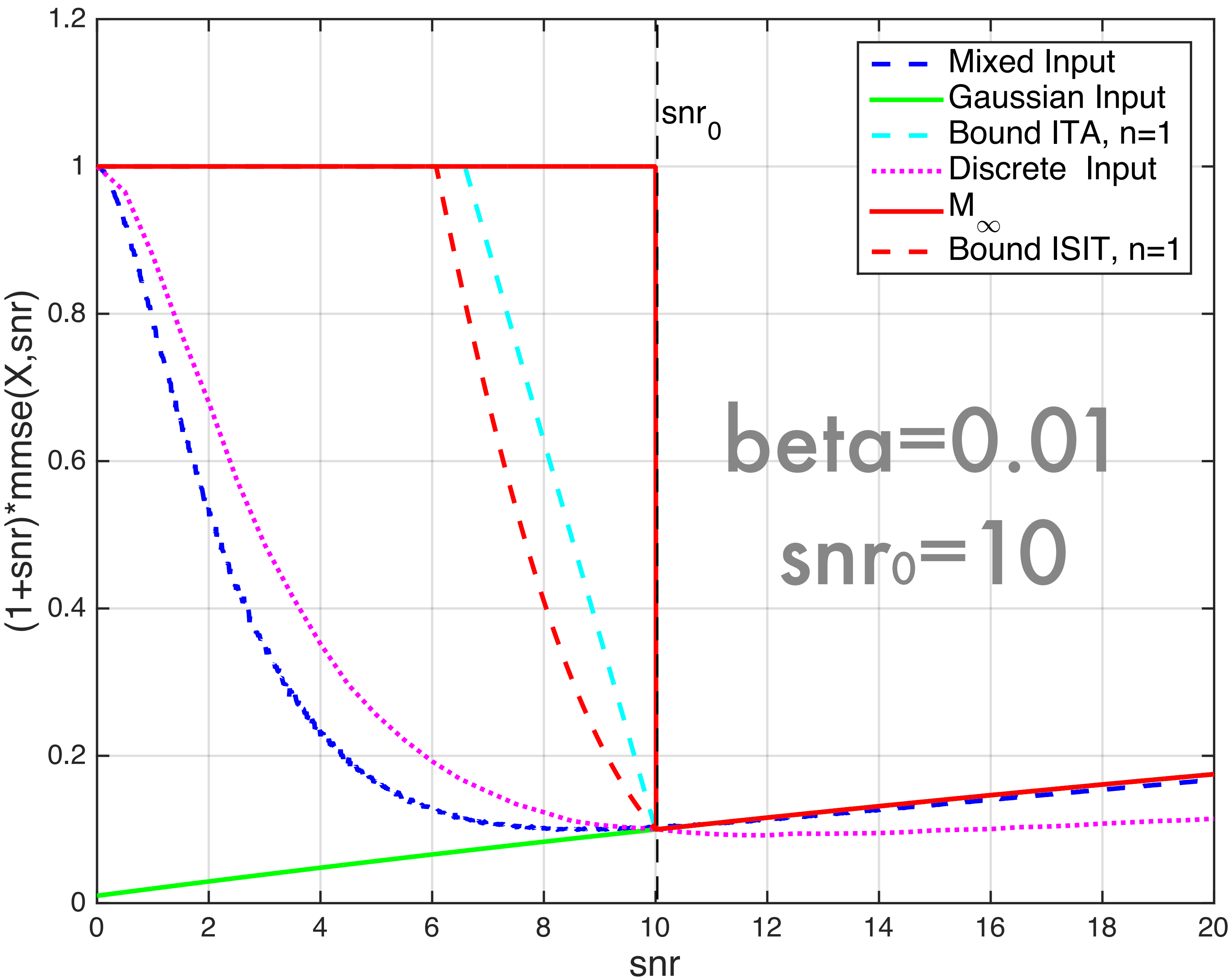
A. Lozano, A. M. Tulino, and S. Verdu´, "Optimum power allocation for parallel Gaussian channels with arbitrary input distributions," *IEEE Trans. Inf. Theory*, vol. 52, no. 7, pp. 3033–3051, July 2006.

$$I(X_D, \text{snr}) \ge H(X_D) - \frac{1}{2}\log\left(\frac{\pi}{6}\right) - \frac{1}{2}\log\left(1 + \frac{12}{d_{\min}^2}\text{mmse}(X_D, \text{snr})\right),$$

A. Dytso, D. Tuninetti, and N. Devroye, "Treating Interference as Noise: Fried or Foe?" arXiv:1506.02597

PRINCETON UNIVERSITY

Bounds for max-MMSE



Mixed Input:

$$\delta = \beta \frac{\text{snr}_0}{1+\text{snr}_0}$$

$$\mathcal{X}_D = [-1.8412, -1.7386, 0.5594]$$

$$\text{with } P_X = [0.1111, 0.1274, 0.7615]$$

Discrete Input:

X_D above only

Mixed Inputs get the best of 'both worlds'

PRINCETON UNIVERSITY

Part 2c

Improved OW Bound

PRINCETON UNIVERSITY

Connections to Conditional Entropy

Bounds on Conditional Entropy

$$h(\mathbf{U}|\mathbf{V}) \le \frac{n}{2}\log(2\pi\mathrm{e}\ \mathrm{mmse}(\mathbf{U}|\mathbf{V})),$$

PRINCETON UNIVERSITY

Connections to Conditional Entropy

Bounds on Conditional Entropy

$$h(\mathbf{U}|\mathbf{V}) \le \frac{n}{2}\log(2\pi\mathrm{e}\ \mathrm{mmse}(\mathbf{U}|\mathbf{V})),$$

Theorem. For any $\mathbf{U} \in \mathbb{R}^n$ such that $h(\mathbf{U}) < \infty$ and $\|\mathbf{U}\|_p < \infty$ for some $p \in (0,\infty)$, and for any $\mathbf{V} \in \mathbb{R}^n$, we have

$$h(\mathbf{U}|\mathbf{V}) \le \frac{n}{2}\log\left(k_{n,\mathrm{p}}^2 \cdot n^{\frac{1}{\mathrm{p}}} \cdot \mathrm{mmpe}^{\frac{2}{\mathrm{p}}}(\mathbf{U}|\mathbf{V};p)\right),$$

where

$$k_{n,\mathrm{p}} := \frac{\sqrt{\pi}\left(\frac{\mathrm{p}}{n}\right)^{\frac{1}{\mathrm{p}}}\mathrm{e}^{\frac{1}{\mathrm{p}}}\Gamma^{\frac{1}{n}}\left(\frac{n}{\mathrm{p}}+1\right)}{\Gamma^{\frac{1}{n}}\left(\frac{n}{2}+1\right)} = \sqrt{2\pi\mathrm{e}}\frac{1}{n^{\frac{1}{2}}\left(\frac{\mathrm{p}}{2}\right)^{\frac{1}{2n}}} + o\left(\frac{n}{\mathrm{p}}\right).$$

- Sharper version of Ozarow-Wyner bound
- Bounds on the derivative of the MMSE

PRINCETON UNIVERSITY

Ozarow-Wyner Bound

(Ozarow-Wyner Bound.) Let X_D be a discrete random variable then

$$
\begin{aligned}
H(X_D) - \text{gap} \leq I(X_D;Y) &\leq H(X_D),\\
\text{gap} &= \frac{1}{2}\log\left(\frac{\pi e}{6}\right) + \frac{1}{2}\log\left(1 + \frac{12\cdot \text{lmmse}(X_D|Y)}{d_{\min}^2(X_D)}\right),\\
\text{lmmse}(X_D|Y) &= \frac{\mathbb{E}[X_D^2]}{1+\text{snr}\cdot\mathbb{E}[X_D^2]}\\
d_{\min}(X_D) &= \inf_{x_i,x_j\in\text{supp}(X_D):i\neq j} |x_i - x_j|.
\end{aligned}
$$

Generalized Ozarow-Wyner Bound

Let $\mathbf{X}_D$ be a discrete random vector, then for any $p \geq 1$

$$[H(\mathbf{X}_D) - \mathsf{gap}_p]^+ \leq I(\mathbf{X}_D; \mathbf{Y}) \leq H(\mathbf{X}_D),$$

where

$$\mathsf{gap}_p = \inf_{\mathbf{U} \in \mathcal{K}} \text{gap}(\mathbf{U}),$$

$$n^{-1} \cdot \text{gap}(\mathbf{U}) = \log\left(\frac{\|\mathbf{U} + \mathbf{X}_D - f_p(\mathbf{X}_D|\mathbf{Y})\|_p}{\|\mathbf{U}\|_p}\right) + \log\left(\frac{k_{n,p} \cdot n^{\frac{1}{p}} \cdot \|\mathbf{U}\|_p}{\mathrm{e}^{\frac{1}{n} h_{\mathrm{e}}(\mathbf{U})}}\right).$$

where $\mathcal{K} = \{\mathbf{U} : \text{diam}(\text{supp}(\mathbf{U})) \leq 2 \cdot d_{\min}(\mathbf{X}_D)\}$.
Moreover,

$$\log\left(\frac{\|\mathbf{U} + \mathbf{X}_D - f_p(\mathbf{X}_D|\mathbf{Y})\|_p}{\|\mathbf{U}\|_p}\right) \overset{\text{for } p \geq 1}{\leq} \log\left(1 + \frac{\text{mmpe}^{\frac{1}{p}}(\mathbf{X}, \text{snr}, p)}{\|\mathbf{U}\|_p}\right).$$

PRINCETON UNIVERSITY

Generalized OW Bound

Proof:

$$I(\mathbf{X}_D;\mathbf{Y}) \geq I(\mathbf{X}_D + \mathbf{U};\mathbf{Y}), \text{ data processing}$$

PRINCETON UNIVERSITY

Generalized OW Bound

Proof:

$$
\begin{aligned}
I(\mathbf{X}_D;\mathbf{Y}) &\geq I(\mathbf{X}_D+\mathbf{U};\mathbf{Y}), \text{ data processing} \\
&= h(\mathbf{X}_D+\mathbf{U}) - h(\mathbf{X}_D+\mathbf{U}|\mathbf{Y})
\end{aligned}
$$

PRINCETON UNIVERSITY

Generalized OW Bound

Proof:

$$\begin{aligned} I(\mathbf{X}_D;\mathbf{Y}) &\geq I(\mathbf{X}_D+\mathbf{U};\mathbf{Y}),\ \text{data processing} \\ &= h(\mathbf{X}_D+\mathbf{U}) - h(\mathbf{X}_D+\mathbf{U}|\mathbf{Y}) \\ &= H(\mathbf{X}_D) + h(\mathbf{U}) - h(\mathbf{X}_D+\mathbf{U}|\mathbf{Y}),\ \text{compact support of } \mathbf{U}, \end{aligned}$$

PRINCETON
UNIVERSITY

Generalized OW Bound

Proof:

$$\begin{aligned} I(\mathbf{X}_D;\mathbf{Y}) &\geq I(\mathbf{X}_D+\mathbf{U};\mathbf{Y}),\ \text{data processing} \\ &= h(\mathbf{X}_D+\mathbf{U}) - h(\mathbf{X}_D+\mathbf{U}|\mathbf{Y}) \\ &= H(\mathbf{X}_D) + h(\mathbf{U}) - h(\mathbf{X}_D+\mathbf{U}|\mathbf{Y}),\ \text{compact support of } \mathbf{U}, \end{aligned}$$

Next, observe

$$n^{-1}h(\mathbf{X}_D+\mathbf{U}|\mathbf{Y}) \leq \log\left(k_{n,p}\cdot n^{\frac{1}{p}}\cdot \|\mathbf{X}_D+\mathbf{U}-g(\mathbf{Y})\|_p\right)$$

PRINCETON UNIVERSITY

Generalized OW Bound

Proof:

$$\begin{aligned} I(\mathbf{X}_D;\mathbf{Y}) &\geq I(\mathbf{X}_D+\mathbf{U};\mathbf{Y}), \text{ data processing} \\ &= h(\mathbf{X}_D+\mathbf{U}) - h(\mathbf{X}_D+\mathbf{U}|\mathbf{Y}) \\ &= H(\mathbf{X}_D) + h(\mathbf{U}) - h(\mathbf{X}_D+\mathbf{U}|\mathbf{Y}), \text{ compact support of } \mathbf{U}, \end{aligned}$$

Next, observe

$$n^{-1}h(\mathbf{X}_D+\mathbf{U}|\mathbf{Y}) \leq \log\left(k_{n,p}\cdot n^{\frac{1}{p}}\cdot \|\mathbf{X}_D+\mathbf{U}-g(\mathbf{Y})\|_p\right)$$

By combining all the bounds

$$I(\mathbf{X}_D;\mathbf{Y}) \geq H(\mathbf{X}_D) - n\cdot\log\left(\frac{\|\mathbf{U}+\mathbf{X}_D-g(\mathbf{Y})\|_p}{\|\mathbf{U}\|_p}\right) - n\cdot\log\left(\frac{k_{n,p}\cdot n^{\frac{1}{p}}\cdot\|\mathbf{U}\|_p}{\mathrm{e}^{\frac{1}{n}h_{\mathrm{e}}(\mathbf{U})}}\right).$$

PRINCETON UNIVERSITY

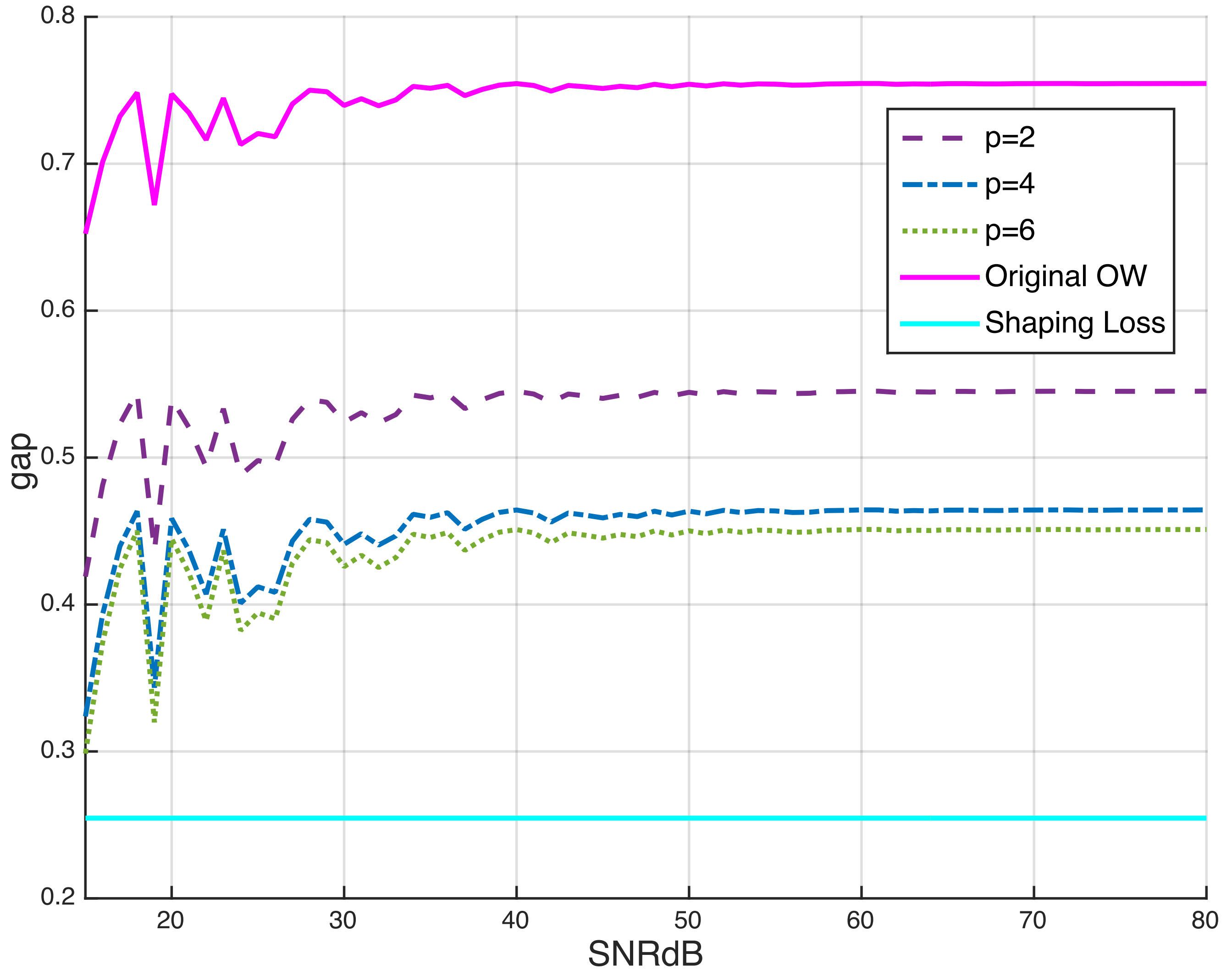
Generalized OW Bound

Take X_D to be PAM with $N \approx \sqrt{1+\text{snr}}$ then

$$H(X_D) \approx \frac{1}{2}\log(1+\text{snr})$$

and we are interested in

$$\frac{1}{2}\log(1+\text{snr}) - I(X_D;Y) = \text{gap}$$



PRINCETON UNIVERSITY

Concluding Remarks

- Properties of the MMPE functional
- Conditional MMPE
- New Proof of the SCPP and its extension
- Complementary SCPP bounds
- Bounds on differential entropy
- Generalized Ozarow-Wyner bound

Outlook

- Non-Gaussian Channels or Channels with Memory [Shamai-Ozarow-Wyner, IT'91]
- New Bounds on the Derivative of the MMSE [Prelov-Verdu, IT'04];
- Bounds on Reyni Entropy
- Relationship with Statistical Physics

PRINCETON UNIVERSITY

A. Dytso, R. Bustin, D. Tuninetti, N. Devroye, S. Shamai, and H. V. Poor, “On the Minimum Mean p-th Error in Gaussian Noise Channels and its Applications ,” Submitted to *IEEE Trans. Inf. Theory,* https://arxiv.org/pdf/1603.07628, 2016.

Thank you

arXiv:1607.01461

adytso@princeton.edu

Back Up Slides

arXiv:1607.01461

adytso@princeton.edu