Entropy-Constrained Gaussian Channel Capacity

Adway Girish
Information theory lab, EPFL

EPFL Information Processing Group

October 30, 2025
SiA group meeting

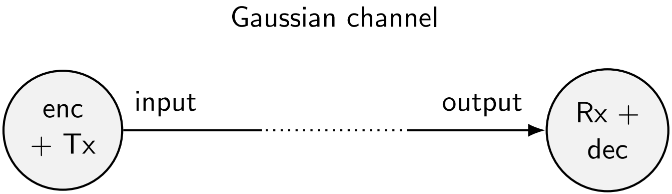
Outline

1. Entropy-constrained Gaussian channel
2. High-SNR asymptotics
3. Low-SNR asymptotics via moment problem

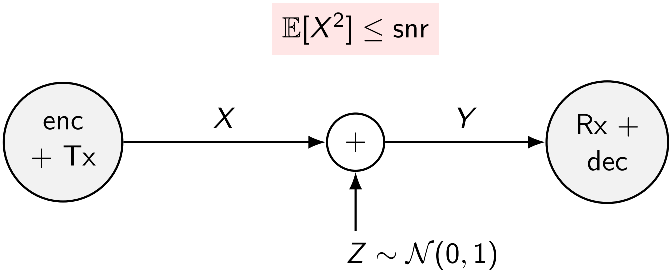
Outline

1. Entropy-constrained Gaussian channel
2. High-SNR asymptotics
3. Low-SNR asymptotics via moment problem

Gaussian channel



Gaussian channel

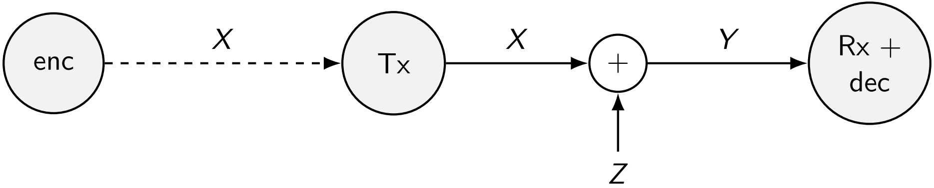


$$C(\mathsf{snr}) = \max_{X:\mathbb{E}[X^2]\leq \mathsf{snr}} I(X; X+Z) = \frac{1}{2}\log(1+\mathsf{snr})$$

Gaussian channel

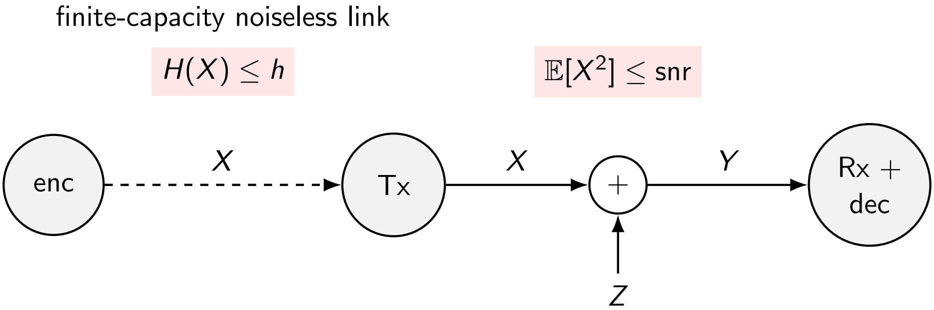
finite-capacity noiseless link

$\mathbb{E}[X^2] \leq \mathsf{snr}$



$$C(\mathsf{snr}) = \max_{X : \mathbb{E}[X^2] \leq \mathsf{snr}} I(X; X + Z) = \frac{1}{2} \log(1 + \mathsf{snr})$$

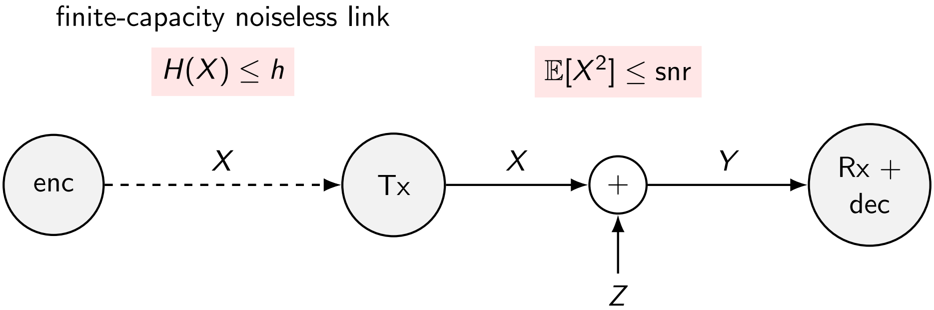
Entropy-constrained Gaussian channel



$$C(\mathsf{snr}) = \max_{X:\mathbb{E}[X^2]\leq\mathsf{snr}} I(X; X+Z) = \frac{1}{2}\log(1+\mathsf{snr})$$

$$C_H(h,\mathsf{snr}) = \max_{X:\substack{\mathbb{E}[X^2]\leq\mathsf{snr} \\ H(X)\leq h}} I(X; X+Z)$$

Entropy-constrained Gaussian channel



$$C(\mathsf{snr}) = \max_{X:\mathbb{E}[X^2]\leq \mathsf{snr}} I(X; X+Z) = \frac{1}{2}\log(1+\mathsf{snr})$$

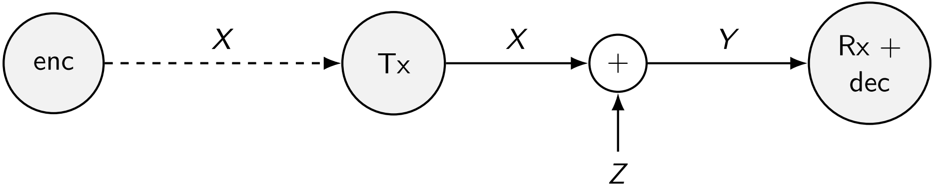
$$C_H(h, \mathsf{snr}) = \max_{\substack{X:\mathbb{E}[X^2]\leq \mathsf{snr} \\ H(X)\leq h}} I(X; X+Z) = \max_{\substack{X:\mathbb{E}[X^2]\leq 1 \\ H(X)\leq h}} I(X; \sqrt{\mathsf{snr}}X+Z)$$

Entropy-constrained Gaussian channel

finite-capacity noiseless link

$H(X) \le h$

$\mathbb{E}[X^2] \le \mathsf{snr}$



$$C(\mathsf{snr}) = \max_{X:\mathbb{E}[X^2]\le \mathsf{snr}} I(X; X+Z) = \frac{1}{2}\log(1+\mathsf{snr})$$

$$C_H(h, \mathsf{snr}) = \max_{X:\substack{\mathbb{E}[X^2]\le \mathsf{snr} \\ H(X)\le h}} I(X; X+Z) = \max_{X:\substack{\mathbb{E}[X^2]\le 1 \\ H(X)\le h}} \underbrace{I(X; \sqrt{\mathsf{snr}}X+Z)}_{I(X,\mathsf{snr})}$$

Approximation perspective

Approximation perspective

- wlog let $\mathbb{E}[X] = 0$, $\mathbb{E}[X^2] = 1$; $G \sim \mathcal{N}(0, 1)$ independent of Z

Approximation perspective

- wlog let $\mathbb{E}[X] = 0$, $\mathbb{E}[X^2] = 1$; $G \sim \mathcal{N}(0, 1)$ independent of Z
- $I(G, \mathsf{snr}) - I(X, \mathsf{snr}) = D(\sqrt{\mathsf{snr}}X + Z \,\|\, \sqrt{\mathsf{snr}}G + Z)$

Approximation perspective

- wlog let $\mathbb{E}[X] = 0$, $\mathbb{E}[X^2] = 1$; $G \sim \mathcal{N}(0, 1)$ independent of Z
- $I(G, \mathsf{snr}) - I(X, \mathsf{snr}) = D(\sqrt{\mathsf{snr}}X + Z \,\|\, \sqrt{\mathsf{snr}}G + Z)$
- $C(\mathsf{snr}) - C_H(h, \mathsf{snr}) = \min\limits_{X:\substack{\mathbb{E}[X^2]\le 1,\\ H(X)\le h}} D(\sqrt{\mathsf{snr}}X + Z \,\|\, \sqrt{\mathsf{snr}}G + Z)$

Approximation perspective

- wlog let $\mathbb{E}[X] = 0$, $\mathbb{E}[X^2] = 1$; $G \sim \mathcal{N}(0,1)$ independent of Z
- $I(G, \mathsf{snr}) - I(X, \mathsf{snr}) = D(\sqrt{\mathsf{snr}}X + Z \,\|\, \sqrt{\mathsf{snr}}G + Z)$
- $C(\mathsf{snr}) - C_H(h, \mathsf{snr}) = \min_{\substack{X:\, \mathbb{E}[X^2] \le 1, \\ H(X) \le h}} D(\sqrt{\mathsf{snr}}X + Z \,\|\, \sqrt{\mathsf{snr}}G + Z)$
- optimal X at h: discrete X with $H(X) \le h$ that is closest to $\mathcal{N}(0, 1 + \mathsf{snr})$ after "Gaussian smoothing"

Estimation perspective

Estimation perspective

- MMSE of estimating X from $Y = \sqrt{\mathsf{snr}}X + Z$:

$$\mathsf{mmse}(X, \mathsf{snr}) = \mathbb{E}\left[(X - \mathbb{E}[X \mid Y])^2\right]$$

Estimation perspective

- MMSE of estimating X from $Y = \sqrt{\mathsf{snr}}X + Z$:

$$\mathsf{mmse}(X, \mathsf{snr}) = \mathbb{E}\left[(X - \mathbb{E}[X \mid Y])^2\right]$$

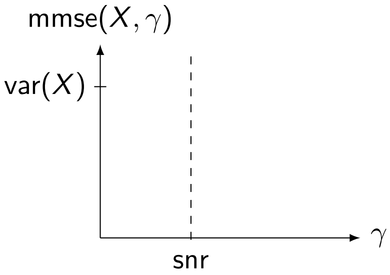
- I-MMSE relationship: $I(X, \mathsf{snr}) = \frac{1}{2}\int_0^{\mathsf{snr}} \mathsf{mmse}(X, \gamma)\,\mathrm{d}\gamma, \qquad H(X) = I(X, \text{“}\infty\text{”})$

Estimation perspective

- MMSE of estimating X from $Y = \sqrt{\mathsf{snr}}X + Z$:

$$\mathsf{mmse}(X, \mathsf{snr}) = \mathbb{E}\left[(X - \mathbb{E}[X \mid Y])^2\right]$$

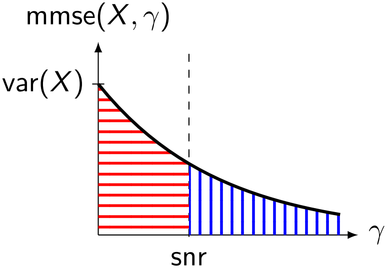
- I-MMSE relationship: $I(X, \mathsf{snr}) = \dfrac{1}{2}\displaystyle\int_0^{\mathsf{snr}} \mathsf{mmse}(X, \gamma)\, \mathrm{d}\gamma, \qquad H(X) = I(X, \text{“}\infty\text{”})$



Estimation perspective

- MMSE of estimating X from $Y = \sqrt{\mathsf{snr}}X + Z$:
$$\mathsf{mmse}(X, \mathsf{snr}) = \mathbb{E}\left[(X - \mathbb{E}[X \mid Y])^2\right]$$

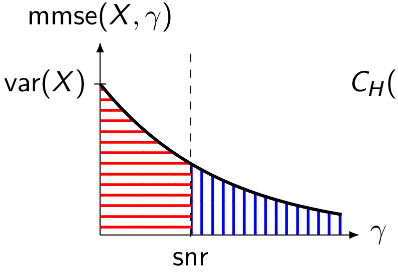
- I-MMSE relationship: $I(X, \mathsf{snr}) = \frac{1}{2}\int_0^{\mathsf{snr}} \mathsf{mmse}(X, \gamma)\, \mathrm{d}\gamma, \qquad H(X) = I(X, \text{“}\infty\text{”})$

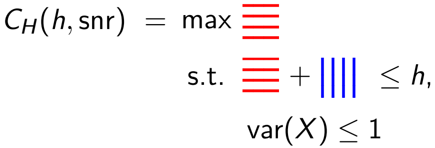


Estimation perspective

- MMSE of estimating X from $Y = \sqrt{\mathsf{snr}}X + Z$:
$$\mathsf{mmse}(X, \mathsf{snr}) = \mathbb{E}\left[(X - \mathbb{E}[X \mid Y])^2\right]$$

- I-MMSE relationship: $I(X, \mathsf{snr}) = \frac{1}{2}\int_0^{\mathsf{snr}} \mathsf{mmse}(X, \gamma)\, d\gamma, \quad H(X) = I(X, \text{``}\infty\text{''})$



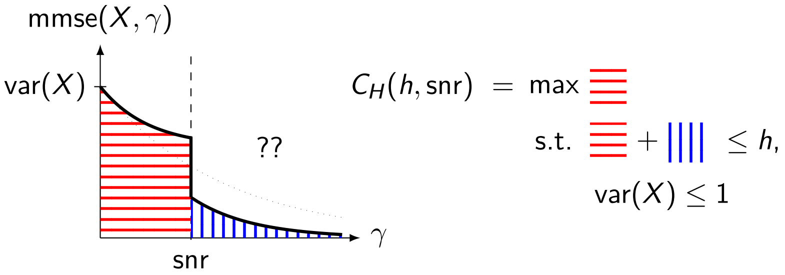


Estimation perspective

- MMSE of estimating X from $Y = \sqrt{\mathsf{snr}}X + Z$:

$$\mathsf{mmse}(X, \mathsf{snr}) = \mathbb{E}\left[(X - \mathbb{E}[X \mid Y])^2\right]$$

- I-MMSE relationship: $I(X, \mathsf{snr}) = \dfrac{1}{2}\displaystyle\int_0^{\mathsf{snr}} \mathsf{mmse}(X, \gamma)\, \mathrm{d}\gamma, \qquad H(X) = I(X, \text{“}\infty\text{”})$

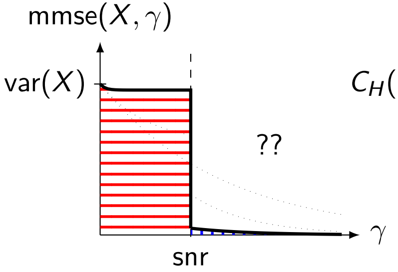


Estimation perspective

- MMSE of estimating X from $Y = \sqrt{\mathsf{snr}}X + Z$:

$$\mathsf{mmse}(X, \mathsf{snr}) = \mathbb{E}\left[(X - \mathbb{E}[X \mid Y])^2\right]$$

- I-MMSE relationship: $I(X, \mathsf{snr}) = \dfrac{1}{2}\displaystyle\int_0^{\mathsf{snr}} \mathsf{mmse}(X, \gamma)\,\mathrm{d}\gamma, \qquad H(X) = I(X, \text{“}\infty\text{”})$



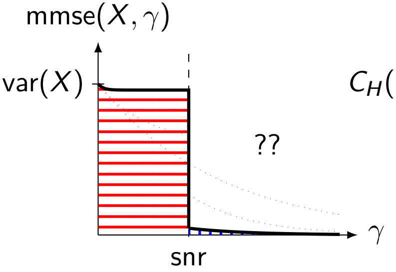
$$\begin{aligned} C_H(h, \mathsf{snr}) \;&=\; \max \\ &\quad \text{s.t.} \quad + \quad \le h, \\ &\quad \mathsf{var}(X) \le 1 \end{aligned}$$

Estimation perspective

- MMSE of estimating X from $Y = \sqrt{\mathsf{snr}}X + Z$:

$$\mathsf{mmse}(X, \mathsf{snr}) = \mathbb{E}\left[(X - \mathbb{E}[X \mid Y])^2\right]$$

- I-MMSE relationship: $I(X, \mathsf{snr}) = \frac{1}{2}\int_0^{\mathsf{snr}} \mathsf{mmse}(X, \gamma)\,\mathrm{d}\gamma, \qquad H(X) = I(X, \text{“}\infty\text{”})$

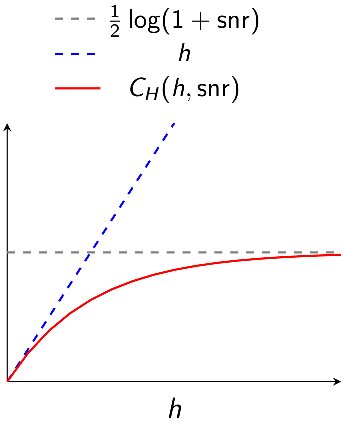


$$\begin{aligned} C_H(h, \mathsf{snr}) \;&=\; \max \\ &\quad \text{s.t.} \quad + \quad \leq h, \\ &\quad \mathsf{var}(X) \leq 1 \end{aligned}$$

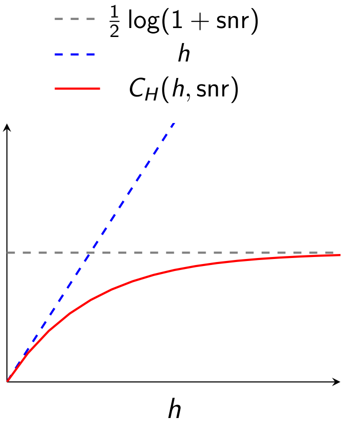
- optimal X at snr: indistinguishable at SNR $<$ snr, distinguishable at SNR $>$ snr

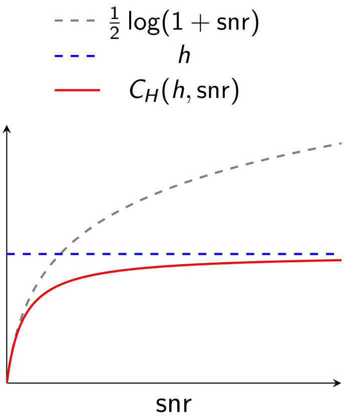
C_H as a function of h and snr

C_H as a function of h and snr

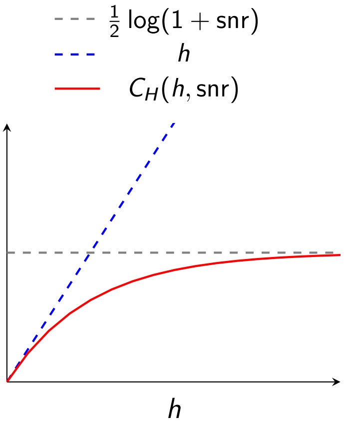


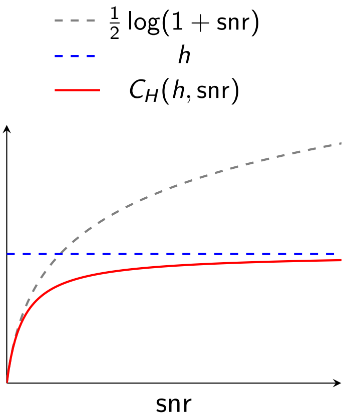
C_H as a function of h and snr





C_H as a function of h and snr





we look at asymptotic behaviour of C_H as snr $\to 0, \infty$

Outline

1. Entropy-constrained Gaussian channel
2. High-SNR asymptotics
3. Low-SNR asymptotics via moment problem

High-SNR asymptotics

- $\mathsf{snr} \to \infty$: $C_H(h, \mathsf{snr}) \to h$, so gap of interest $= h - C_H(h, \mathsf{snr}) = H(X^*) - I(X^*, \mathsf{snr})$

High-SNR asymptotics

- $\mathsf{snr} \to \infty$: $C_H(h, \mathsf{snr}) \to h$, so gap of interest $= h - C_H(h, \mathsf{snr}) = H(X^*) - I(X^*, \mathsf{snr})$

Lemma

For any X with $\mathbb{E}[X^2] = 1$ and finite $H(X)$, and $Z \sim \mathcal{N}(0, 1)$ independent of X, we have

$$H(X \mid \sqrt{\mathsf{snr}}X + Z) = \exp\left(-\mathsf{snr}\frac{\mathrm{d}_{\min}(X)^2}{8} + \mathrm{o}(\mathsf{snr})\right).$$

High-SNR asymptotics

- $\mathsf{snr} \to \infty$: $C_H(h, \mathsf{snr}) \to h$, so gap of interest $= h - C_H(h, \mathsf{snr}) = H(X^*) - I(X^*, \mathsf{snr})$

Lemma

For any X with $\mathbb{E}[X^2] = 1$ and finite $H(X)$, and $Z \sim \mathcal{N}(0,1)$ independent of X, we have

$$H(X \mid \sqrt{\mathsf{snr}}X + Z) = \exp\left(-\mathsf{snr}\frac{\mathrm{d}_{\min}(X)^2}{8} + \mathrm{o}(\mathsf{snr})\right).$$

- $X_h \;:=\; X$ with largest $\mathrm{d}_{\min}$ among all X with $H(X) = h$ and $\mathbb{E}[X^2] = 1$

High-SNR asymptotics

- $\mathsf{snr} \to \infty$: $C_H(h, \mathsf{snr}) \to h$, so gap of interest $= h - C_H(h, \mathsf{snr}) = H(X^*) - I(X^*, \mathsf{snr})$

Lemma

For any X with $\mathbb{E}[X^2] = 1$ and finite $H(X)$, and $Z \sim \mathcal{N}(0,1)$ independent of X, we have

$$H(X \mid \sqrt{\mathsf{snr}}X + Z) = \exp\left(-\mathsf{snr}\frac{\mathsf{d_{min}}(X)^2}{8} + \mathrm{o}(\mathsf{snr})\right).$$

- X_h $:=$ X with largest $\mathsf{d_{min}}$ among all X with $H(X) = h$ and $\mathbb{E}[X^2] = 1$
  $=$ supported on $\beta\mathbb{Z}$ with pmf $\mathsf{P}_{X_h}(x) \propto \exp(-\lambda x^2)$

High-SNR asymptotics

- $\mathsf{snr} \to \infty : C_H(h, \mathsf{snr}) \to h$, so gap of interest $= h - C_H(h, \mathsf{snr}) = H(X^*) - I(X^*, \mathsf{snr})$

Lemma

For any X with $\mathbb{E}[X^2] = 1$ and finite $H(X)$, and $Z \sim \mathcal{N}(0,1)$ independent of X, we have

$$H(X \mid \sqrt{\mathsf{snr}}X + Z) = \exp\left(-\mathsf{snr}\frac{\mathrm{d_{min}}(X)^2}{8} + \mathrm{o}(\mathsf{snr})\right).$$

- $X_h \;:=\; X$ with largest $\mathrm{d_{min}}$ among all X with $H(X) = h$ and $\mathbb{E}[X^2] = 1$
  $=$ supported on $\beta\mathbb{Z}$ with pmf $\mathrm{P}_{X_h}(x) \propto \exp(-\lambda x^2)$
- then "X_h is optimal as $\mathsf{snr} \to \infty$":

High-SNR asymptotics

- $\mathsf{snr} \to \infty$: $C_H(h, \mathsf{snr}) \to h$, so gap of interest $= h - C_H(h, \mathsf{snr}) = H(X^*) - I(X^*, \mathsf{snr})$

Lemma

For any X with $\mathbb{E}[X^2] = 1$ and finite $H(X)$, and $Z \sim \mathcal{N}(0,1)$ independent of X, we have

$$H(X \mid \sqrt{\mathsf{snr}}X + Z) = \exp\left(-\mathsf{snr}\frac{\mathrm{d}_{\min}(X)^2}{8} + \mathrm{o}(\mathsf{snr})\right).$$

- $X_h := X$ with largest $\mathrm{d}_{\min}$ among all X with $H(X) = h$ and $\mathbb{E}[X^2] = 1$
  $=$ supported on $\beta\mathbb{Z}$ with pmf $\mathrm{P}_{X_h}(x) \propto \exp(-\lambda x^2)$

- then "X_h is optimal as $\mathsf{snr} \to \infty$":
  for any other X, there exists snr_0 such that $I(X_h, \mathsf{snr}) > I(X, \mathsf{snr})$ for all $\mathsf{snr} > \mathsf{snr}_0$

High-SNR asymptotics

- $\mathrm{snr} \to \infty$: $C_H(h, \mathrm{snr}) \to h$, so **gap** of interest $= h - C_H(h, \mathrm{snr}) = H(X^*) - I(X^*, \mathrm{snr})$

Lemma

For any X with $\mathbb{E}[X^2] = 1$ and finite $H(X)$, and $Z \sim \mathcal{N}(0,1)$ independent of X, we have

$$H(X \mid \sqrt{\mathrm{snr}}X + Z) = \exp\left(-\mathrm{snr}\frac{\mathrm{d}_{\min}(X)^2}{8} + \mathrm{o}(\mathrm{snr})\right).$$

- $X_h \;:=\; X$ with largest $\mathrm{d}_{\min}$ among all X with $H(X) = h$ and $\mathbb{E}[X^2] = 1$
  $=$ supported on $\beta\mathbb{Z}$ with pmf $\mathrm{P}_{X_h}(x) \propto \exp(-\lambda x^2)$

- then "X_h is optimal as $\mathrm{snr} \to \infty$":
  for any other X, there exists snr_0 such that $I(X_h, \mathrm{snr}) > I(X, \mathrm{snr})$ for all $\mathrm{snr} > \mathrm{snr}_0$

- need stronger $X^* \xrightarrow{d} X_h$ as $\mathrm{snr} \to \infty$ to characterize **gap** in terms of $\mathrm{d}_{\min}(X_h) = \beta$

Outline

1. Entropy-constrained Gaussian channel
2. High-SNR asymptotics
3. Low-SNR asymptotics via moment problem

Low SNR asymptotics via Taylor expansion

Low SNR asymptotics via Taylor expansion

- $\mathbb{E}[X^{2n}] < \infty$: $I(X, \mathsf{snr}) = \sum_{i=1}^{n-1} a_{i,X}\, \mathsf{snr}^i + r_{n,X}\, \mathsf{snr}^n$

Low SNR asymptotics via Taylor expansion

- $\mathbb{E}[X^{2n}] < \infty$: $I(X, \mathsf{snr}) = \sum_{i=1}^{n-1} a_{i,X}\, \mathsf{snr}^i + r_{n,X}\, \mathsf{snr}^n$
- $a_{k,X}$ is a polynomial of $\{\mathbb{E}[X^n] : n = 1, \ldots, k\}$ for $k \geq 2$

Low SNR asymptotics via Taylor expansion

- $\mathbb{E}[X^{2n}] < \infty$: $I(X, \mathsf{snr}) = \sum_{i=1}^{n-1} a_{i,X}\,\mathsf{snr}^i + r_{n,X}\,\mathsf{snr}^n$
- $a_{k,X}$ is a polynomial of $\{\mathbb{E}[X^n] : n = 1, \ldots, k\}$ for $k \geq 2$
- for X with $\mathbb{E}[X^{2(k+2)}] < \infty$:

Low SNR asymptotics via Taylor expansion

- $\mathbb{E}[X^{2n}] < \infty$: $I(X, \mathsf{snr}) = \sum_{i=1}^{n-1} a_{i,X}\,\mathsf{snr}^i + r_{n,X}\,\mathsf{snr}^n$
- $a_{k,X}$ is a polynomial of $\{\mathbb{E}[X^n] : n = 1, \ldots, k\}$ for $k \geq 2$
- for X with $\mathbb{E}[X^{2(k+2)}] < \infty$: if $\mathbb{E}[X^n] = \mathbb{E}[G^n]$, $n = 1, \ldots, k$ and $\mathbb{E}[X^{k+1}] \neq \mathbb{E}[G^{k+1}]$,

Low SNR asymptotics via Taylor expansion

- $\mathbb{E}[X^{2n}] < \infty$: $I(X, \mathsf{snr}) = \sum_{i=1}^{n-1} a_{i,X}\, \mathsf{snr}^i + r_{n,X}\, \mathsf{snr}^n$

- $a_{k,X}$ is a polynomial of $\{\mathbb{E}[X^n] : n = 1, \dots, k\}$ for $k \geq 2$

- for X with $\mathbb{E}[X^{2(k+2)}] < \infty$: if $\mathbb{E}[X^n] = \mathbb{E}[G^n]$, $n = 1, \dots, k$ and $\mathbb{E}[X^{k+1}] \neq \mathbb{E}[G^{k+1}]$, then $I(G, \mathsf{snr}) - I(X, \mathsf{snr}) = \Theta(\mathsf{snr}^{k+1})$ as $\mathsf{snr} \to 0$

Low SNR asymptotics via Taylor expansion

- $\mathbb{E}[X^{2n}] < \infty$: $I(X, \mathsf{snr}) = \sum_{i=1}^{n-1} a_{i,X}\, \mathsf{snr}^i + r_{n,X}\, \mathsf{snr}^n$
- $a_{k,X}$ is a polynomial of $\{\mathbb{E}[X^n] : n = 1, \dots, k\}$ for $k \geq 2$
- for X with $\mathbb{E}[X^{2(k+2)}] < \infty$: if $\mathbb{E}[X^n] = \mathbb{E}[G^n]$, $n = 1, \dots, k$ and $\mathbb{E}[X^{k+1}] \neq \mathbb{E}[G^{k+1}]$, then $I(G, \mathsf{snr}) - I(X, \mathsf{snr}) = \Theta(\mathsf{snr}^{k+1})$ as $\mathsf{snr} \to 0$
- $k_h :=$ maximum number of moments of G matched by X with $H(X) \leq h$

Low SNR asymptotics via Taylor expansion

- $\mathbb{E}[X^{2n}] < \infty$: $I(X, \mathsf{snr}) = \sum_{i=1}^{n-1} a_{i,X}\, \mathsf{snr}^i + r_{n,X}\, \mathsf{snr}^n$
- $a_{k,X}$ is a polynomial of $\{\mathbb{E}[X^n] : n = 1, \ldots, k\}$ for $k \geq 2$
- for X with $\mathbb{E}[X^{2(k+2)}] < \infty$: if $\mathbb{E}[X^n] = \mathbb{E}[G^n]$, $n = 1, \ldots, k$ and $\mathbb{E}[X^{k+1}] \neq \mathbb{E}[G^{k+1}]$, then $I(G, \mathsf{snr}) - I(X, \mathsf{snr}) = \Theta(\mathsf{snr}^{k+1})$ as $\mathsf{snr} \to 0$
- $k_h :=$ maximum number of moments of G matched by X with $H(X) \leq h$
- $\boxed{C(\mathsf{snr}) - C_H(h, \mathsf{snr}) = \mathrm{O}(\mathsf{snr}^{k_h+1})}$

Aside: computing integrals numerically

- $\mathbb{E}[g(W)] = \int_{\mathbb{R}} g(w) f_W(w) \, \mathrm{d}w$

Aside: computing integrals numerically

- $\mathbb{E}[g(W)] = \int_{\mathbb{R}} g(w) f_W(w)\,\mathrm{d}w \approx \sum_{i=1}^{m} p_i g(x_i)$

Aside: computing integrals numerically

- $\mathbb{E}[g(W)] = \int_{\mathbb{R}} g(w) f_W(w)\, \mathrm{d}w \approx \sum_{i=1}^{m} p_i g(x_i) = \mathbb{E}[g(X)]$

Aside: computing integrals numerically

- $\mathbb{E}[g(W)] = \int_{\mathbb{R}} g(w) f_W(w)\, \mathrm{d}w \approx \sum_{i=1}^{m} p_i g(x_i) = \mathbb{E}[g(X)]$

- if X has $\mathbb{E}[X^n] = \mathbb{E}[W^n]$ for $n = 1, \ldots, k$

Aside: computing integrals numerically

- $\mathbb{E}[g(W)] = \int_{\mathbb{R}} g(w) f_W(w)\,\mathrm{d}w \approx \sum_{i=1}^{m} p_i g(x_i) = \mathbb{E}[g(X)]$

- if X has $\mathbb{E}[X^n] = \mathbb{E}[W^n]$ for $n = 1, \ldots, k$, then

$$\int_{\mathbb{R}} g(w) f_W(w)\,\mathrm{d}w = \sum_{i=1}^{m} p_i g(x_i)$$

  for any polynomial g of degree $\leq k$

Aside: computing integrals numerically

- $\mathbb{E}[g(W)] = \int_{\mathbb{R}} g(w) f_W(w)\, \mathrm{d}w \approx \sum_{i=1}^{m} p_i g(x_i) = \mathbb{E}[g(X)]$

- if X has $\mathbb{E}[X^n] = \mathbb{E}[W^n]$ for $n = 1, \ldots, k$, then

$$\int_{\mathbb{R}} g(w) f_W(w)\, \mathrm{d}w = \sum_{i=1}^{m} p_i g(x_i)$$

  for any polynomial g of degree $\leq k$

- given continuous $W \sim f_W$ and [C] constraints on discrete X, how large can $k_{[\mathtt{C}]}$ be?

Classical moment problem

Classical moment problem

- Q: Given $s_1, s_2, s_3, \ldots$, does there exist X on $\mathbb{R}$ such that $\mathbb{E}[X^n] = s_n$ for $n = 1, 2, 3, \ldots$?

Classical moment problem

- Q: Given $s_1, s_2, s_3, \dots$, does there exist X on $\mathbb{R}$ such that $\mathbb{E}[X^n] = s_n$ for $n = 1, 2, 3, \dots$?

  A: iff $\mathrm{H}_n(s_1, \dots, s_{2n}) = \begin{pmatrix} 1 & s_1 & \dots & s_n \\ s_1 & s_2 & \dots & s_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ s_n & s_{n+1} & \dots & s_{2n} \end{pmatrix} \succeq 0$ for $n = 1, 2, 3, \dots$

Classical moment problem

- Q: Given $s_1, s_2, s_3, \dots$, does there exist X on $\mathbb{R}$ such that $\mathbb{E}[X^n] = s_n$ for $n = 1, 2, 3, \dots$?

  A: iff $\mathrm{H}_n(s_1, \dots, s_{2n}) = \begin{pmatrix} 1 & s_1 & \dots & s_n \\ s_1 & s_2 & \dots & s_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ s_n & s_{n+1} & \dots & s_{2n} \end{pmatrix} \succeq 0$ for $n = 1, 2, 3, \dots$;

  infinite support iff $\mathrm{H}_n \succ 0$ for all n

Truncated moment problem

- Q: Given $s_1, s_2, s_3, \ldots$, does there exist X on $\mathbb{R}$ such that $\mathbb{E}[X^n] = s_n$ for $n = 1, 2, 3, \ldots$?

  A: iff $\mathrm{H}_n(s_1, \ldots, s_{2n}) = \begin{pmatrix} 1 & s_1 & \ldots & s_n \\ s_1 & s_2 & \ldots & s_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ s_n & s_{n+1} & \ldots & s_{2n} \end{pmatrix} \succeq 0$ for $n = 1, 2, 3, \ldots$;

  infinite support iff $\mathrm{H}_n \succ 0$ for all n

- Q: Given $s_1, s_2, s_3, \ldots, s_k$, does there exist X on $\mathbb{R}$ such that $\mathbb{E}[X^n] = s_n$ for $n = 1, 2, 3, \ldots, k$?

Truncated moment problem

- Q: Given $s_1, s_2, s_3, \ldots$, does there exist X on $\mathbb{R}$ such that $\mathbb{E}[X^n] = s_n$ for $n = 1, 2, 3, \ldots$?

  A: iff $\mathrm{H}_n(s_1, \ldots, s_{2n}) = \begin{pmatrix} 1 & s_1 & \ldots & s_n \\ s_1 & s_2 & \ldots & s_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ s_n & s_{n+1} & \ldots & s_{2n} \end{pmatrix} \succeq 0$ for $n = 1, 2, 3, \ldots$;

  infinite support iff $\mathrm{H}_n \succ 0$ for all n

- Q: Given $s_1, s_2, s_3, \ldots, s_k$, does there exist X on $\mathbb{R}$ such that $\mathbb{E}[X^n] = s_n$ for $n = 1, 2, 3, \ldots, k$?

  A: iff (k odd)

  (k even)

Truncated moment problem

- Q: Given $s_1, s_2, s_3, \ldots,$ does there exist X on $\mathbb{R}$ such that $\mathbb{E}[X^n] = s_n$ for $n = 1, 2, 3, \ldots$?

  A: iff $\mathrm{H}_n(s_1, \ldots, s_{2n}) = \begin{pmatrix} 1 & s_1 & \ldots & s_n \\ s_1 & s_2 & \ldots & s_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ s_n & s_{n+1} & \ldots & s_{2n} \end{pmatrix} \succeq 0$ for $n = 1, 2, 3, \ldots;$

  infinite support iff $\mathrm{H}_n \succ 0$ for all n

- Q: Given $s_1, s_2, s_3, \ldots, s_k,$ does there exist X on $\mathbb{R}$ such that $\mathbb{E}[X^n] = s_n$ for $n = 1, 2, 3, \ldots, k$?

  A: iff $(k$ odd$)$ there exists $\tilde{s}_{k+1}$ such that $\mathrm{H}_{\frac{k+1}{2}}(s_1, \ldots, s_k, \tilde{s}_{k+1}) \succeq 0$

  $(k$ even$)$

Truncated moment problem

- Q: Given $s_1, s_2, s_3, \ldots$, does there exist X on $\mathbb{R}$ such that $\mathbb{E}[X^n] = s_n$ for $n = 1, 2, 3, \ldots$?

  A: iff $\mathrm{H}_n(s_1, \ldots, s_{2n}) = \begin{pmatrix} 1 & s_1 & \ldots & s_n \\ s_1 & s_2 & \ldots & s_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ s_n & s_{n+1} & \ldots & s_{2n} \end{pmatrix} \succeq 0$ for $n = 1, 2, 3, \ldots$;

  infinite support iff $\mathrm{H}_n \succ 0$ for all n

- Q: Given $s_1, s_2, s_3, \ldots, s_k$, does there exist X on $\mathbb{R}$ such that $\mathbb{E}[X^n] = s_n$ for $n = 1, 2, 3, \ldots, k$?

  A: iff (k odd) there exists $\tilde{s}_{k+1}$ such that $\mathrm{H}_{\frac{k+1}{2}}(s_1, \ldots, s_k, \tilde{s}_{k+1}) \succeq 0$

  (k even) there exist $\tilde{s}_{k+1}, \tilde{s}_{k+2}$ such that $\mathrm{H}_{\frac{k}{2}+1}(s_1, \ldots, s_k, \tilde{s}_{k+1}, \tilde{s}_{k+2}) \succeq 0$

Truncated moment problem

- Q: Given $s_1, s_2, s_3, \ldots$, does there exist X on $\mathbb{R}$ such that $\mathbb{E}[X^n] = s_n$ for $n = 1, 2, 3, \ldots$?

  A: iff $\mathrm{H}_n(s_1, \ldots, s_{2n}) = \begin{pmatrix} 1 & s_1 & \ldots & s_n \\ s_1 & s_2 & \ldots & s_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ s_n & s_{n+1} & \ldots & s_{2n} \end{pmatrix} \succeq 0$ for $n = 1, 2, 3, \ldots$;

  infinite support iff $\mathrm{H}_n \succ 0$ for all n

- Q: Given $s_1, s_2, s_3, \ldots, s_k$, does there exist X on $\mathbb{R}$ such that $\mathbb{E}[X^n] = s_n$ for $n = 1, 2, 3, \ldots, k$?

  A: iff (k odd) there exists $\tilde{s}_{k+1}$ such that $\mathrm{H}_{\frac{k+1}{2}}(s_1, \ldots, s_k, \tilde{s}_{k+1}) \succeq 0$

  (k even) there exist $\tilde{s}_{k+1}, \tilde{s}_{k+2}$ such that $\mathrm{H}_{\frac{k}{2}+1}(s_1, \ldots, s_k, \tilde{s}_{k+1}, \tilde{s}_{k+2}) \succeq 0$;

  finite support: at most $\lfloor k/2 \rfloor + 1$ atoms (if it exists)

Gauss–Hermite quadrature

Gauss–Hermite quadrature

- $W = G \sim \mathcal{N}(0,1)$ with density $\varphi(w) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{w^2}{2})$

Gauss–Hermite quadrature

- $W = G \sim \mathcal{N}(0,1)$ with density $\varphi(w) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{w^2}{2})$
- "Gauss–Hermite quadrature" X_m^{Q} with m atoms matches $2m-1$ moments:

Gauss–Hermite quadrature

- $W = G \sim \mathcal{N}(0,1)$ with density $\varphi(w) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{w^2}{2})$
- "Gauss–Hermite quadrature" X_m^{Q} with m atoms matches $2m-1$ moments:

$$\text{atom } x_i = \text{root of } P_m(x) = \frac{(-1)^m}{\varphi(x)} \frac{\mathrm{d}^m \varphi(x)}{\mathrm{d}x^m},$$

Gauss–Hermite quadrature

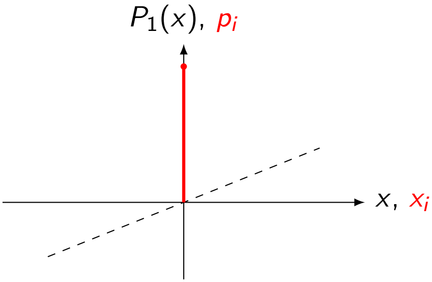
- $W = G \sim \mathcal{N}(0,1)$ with density $\varphi(w) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{w^2}{2})$
- "Gauss–Hermite quadrature" X_m^{Q} with m atoms matches $2m-1$ moments:

$$\text{atom } x_i = \text{root of } P_m(x) = \frac{(-1)^m}{\varphi(x)} \frac{\mathrm{d}^m \varphi(x)}{\mathrm{d}x^m}, \quad \text{with probability } p_i = \frac{(m-1)!}{m P_{m-1}^2(x_i)}$$

Gauss–Hermite quadrature

- $W = G \sim \mathcal{N}(0,1)$ with density $\varphi(w) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{w^2}{2})$
- "Gauss–Hermite quadrature" X_m^{Q} with m atoms matches $2m-1$ moments:

$$\text{atom } x_i = \text{root of } P_m(x) = \frac{(-1)^m}{\varphi(x)} \frac{\mathrm{d}^m \varphi(x)}{\mathrm{d}x^m}, \quad \text{with probability } p_i = \frac{(m-1)!}{m P_{m-1}^2(x_i)}$$

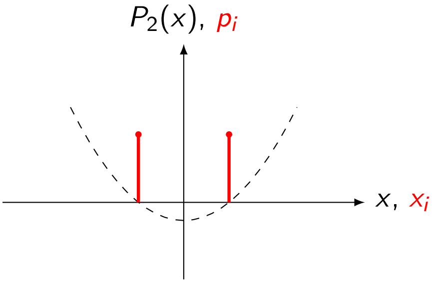


$\mathbb{E}[X] = 0 \quad \mathbb{E}[X^2] \neq 1$

Gauss–Hermite quadrature

- $W = G \sim \mathcal{N}(0,1)$ with density $\varphi(w) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{w^2}{2})$
- "Gauss–Hermite quadrature" X_m^{Q} with m atoms matches $2m-1$ moments:

$$\text{atom } x_i = \text{root of } P_m(x) = \frac{(-1)^m}{\varphi(x)} \frac{\mathrm{d}^m \varphi(x)}{\mathrm{d}x^m}, \quad \text{with probability } p_i = \frac{(m-1)!}{m P_{m-1}^2(x_i)}$$

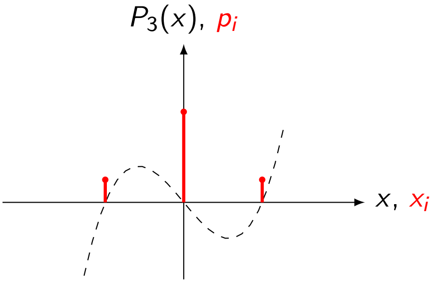


$$\mathbb{E}[X] = 0 \;\; \mathbb{E}[X^2] = 1$$
$$\mathbb{E}[X^3] = 0 \;\; \mathbb{E}[X^4] \neq 3$$

Gauss–Hermite quadrature

- $W = G \sim \mathcal{N}(0,1)$ with density $\varphi(w) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{w^2}{2})$
- "Gauss–Hermite quadrature" X_m^{Q} with m atoms matches $2m-1$ moments:

$$\text{atom } x_i = \text{root of } P_m(x) = \frac{(-1)^m}{\varphi(x)} \frac{\mathrm{d}^m \varphi(x)}{\mathrm{d}x^m}, \quad \text{with probability } p_i = \frac{(m-1)!}{m P_{m-1}^2(x_i)}$$

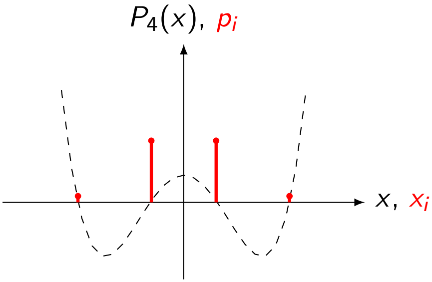


$$\mathbb{E}[X] = 0 \quad \mathbb{E}[X^2] = 1$$
$$\mathbb{E}[X^3] = 0 \quad \mathbb{E}[X^4] = 3$$
$$\mathbb{E}[X^5] = 0 \quad \mathbb{E}[X^6] \neq 15$$

Gauss–Hermite quadrature

- $W = G \sim \mathcal{N}(0,1)$ with density $\varphi(w) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{w^2}{2})$
- "Gauss–Hermite quadrature" X_m^{Q} with m atoms matches $2m-1$ moments:

$$\text{atom } x_i = \text{root of } P_m(x) = \frac{(-1)^m}{\varphi(x)} \frac{\mathrm{d}^m \varphi(x)}{\mathrm{d}x^m}, \quad \text{with probability } p_i = \frac{(m-1)!}{m P_{m-1}^2(x_i)}$$

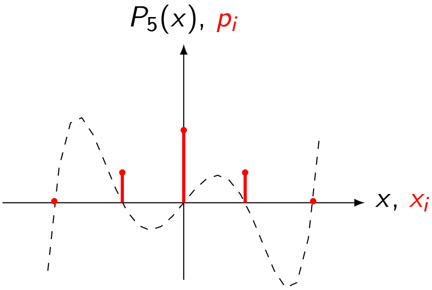


$$\mathbb{E}[X] = 0 \;\; \mathbb{E}[X^2] = 1$$
$$\mathbb{E}[X^3] = 0 \;\; \mathbb{E}[X^4] = 3$$
$$\mathbb{E}[X^5] = 0 \;\; \mathbb{E}[X^6] = 15$$
$$\mathbb{E}[X^7] = 0 \;\; \mathbb{E}[X^8] \neq 105$$

Gauss–Hermite quadrature

- $W = G \sim \mathcal{N}(0,1)$ with density $\varphi(w) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{w^2}{2})$
- "Gauss–Hermite quadrature" X_m^{Q} with m atoms matches $2m-1$ moments:

$$\text{atom } x_i = \text{root of } P_m(x) = \frac{(-1)^m}{\varphi(x)} \frac{\mathrm{d}^m \varphi(x)}{\mathrm{d}x^m}, \quad \text{with probability } p_i = \frac{(m-1)!}{m P_{m-1}^2(x_i)}$$

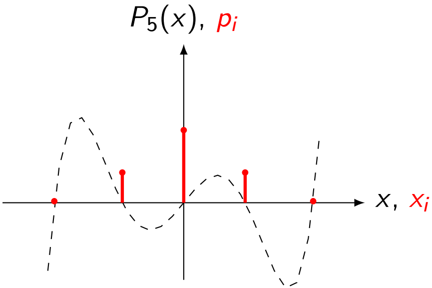


$$\mathbb{E}[X] = 0 \quad \mathbb{E}[X^2] = 1$$
$$\mathbb{E}[X^3] = 0 \quad \mathbb{E}[X^4] = 3$$
$$\mathbb{E}[X^5] = 0 \quad \mathbb{E}[X^6] = 15$$
$$\mathbb{E}[X^7] = 0 \quad \mathbb{E}[X^8] = 105$$
$$\mathbb{E}[X^9] = 0 \quad \mathbb{E}[X^{10}] \neq 1155$$

Gauss–Hermite quadrature

- $W = G \sim \mathcal{N}(0,1)$ with density $\varphi(w) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{w^2}{2})$
- "Gauss–Hermite quadrature" X_m^{Q} with m atoms matches $2m-1$ moments:

$$\text{atom } x_i = \text{root of } P_m(x) = \frac{(-1)^m}{\varphi(x)} \frac{\mathrm{d}^m \varphi(x)}{\mathrm{d}x^m}, \quad \text{with probability } p_i = \frac{(m-1)!}{m P_{m-1}^2(x_i)}$$



$$\mathbb{E}[X] = 0 \quad \mathbb{E}[X^2] = 1$$
$$\mathbb{E}[X^3] = 0 \quad \mathbb{E}[X^4] = 3$$
$$\mathbb{E}[X^5] = 0 \quad \mathbb{E}[X^6] = 15$$
$$\mathbb{E}[X^7] = 0 \quad \mathbb{E}[X^8] = 105$$
$$\mathbb{E}[X^9] = 0 \quad \mathbb{E}[X^{10}] \neq 1155$$

- optimal trade-off of atoms to moments matched, but $H(X_m^{\mathrm{Q}}) \approx \frac{1}{2} \log m \to \infty$

Only three moments can be matched with small entropy

Theorem

For any continuous W*, there exists* $\eta_W \in (0, \frac{1}{2})$ *such that*

Only three moments can be matched with small entropy

Theorem

For any continuous W, there exists $\eta_W \in (0, \frac{1}{2})$ such that

(i) *if X has $\mathbb{E}[X^n] = \mathbb{E}[W^n]$ for $n = 1, 2, 3, 4$, then $H(X) \geq h_2(\eta_W)$.*

Only three moments can be matched with small entropy

Theorem

For any continuous W, there exists $\eta_W \in (0, \frac{1}{2})$ such that

(i) *if X has $\mathbb{E}[X^n] = \mathbb{E}[W^n]$ for $n = 1, 2, 3, 4$, then $H(X) \geq h_2(\eta_W)$.*

(ii) *for any $h > 0$, there is X with $H(X) \leq h$ and $\mathbb{E}[X^n] = \mathbb{E}[W^n]$ for $n = 1, 2, 3$.*

Only three moments can be matched with small entropy

Theorem

For any continuous W, there exists $\eta_W \in (0, \frac{1}{2})$ such that

(i) *if X has $\mathbb{E}[X^n] = \mathbb{E}[W^n]$ for $n = 1, 2, 3, 4$, then $H(X) \geq h_2(\eta_W)$.*

(ii) *for any $h > 0$, there is X with $H(X) \leq h$ and $\mathbb{E}[X^n] = \mathbb{E}[W^n]$ for $n = 1, 2, 3$.*

Proof idea

□

Only three moments can be matched with small entropy

Theorem

For any continuous W, there exists $\eta_W \in (0, \frac{1}{2})$ such that

(i) *if X has $\mathbb{E}[X^n] = \mathbb{E}[W^n]$ for $n = 1, 2, 3, 4$, then $H(X) \geq h_2(\eta_W)$.*

(ii) *for any $h > 0$, there is X with $H(X) \leq h$ and $\mathbb{E}[X^n] = \mathbb{E}[W^n]$ for $n = 1, 2, 3$.*

Proof idea

- $H(X) \leq h < \log 2 \iff X = \begin{cases} \tilde{X} & \text{w.p. } \epsilon < 1/2 \\ x_0 & \text{w.p. } 1 - \epsilon > 1/2 \end{cases}$

$\square$

Only three moments can be matched with small entropy

Theorem

For any continuous W, there exists $\eta_W \in (0, \frac{1}{2})$ such that

(i) *if X has $\mathbb{E}[X^n] = \mathbb{E}[W^n]$ for $n = 1, 2, 3, 4$, then $H(X) \geq h_2(\eta_W)$.*

(ii) *for any $h > 0$, there is X with $H(X) \leq h$ and $\mathbb{E}[X^n] = \mathbb{E}[W^n]$ for $n = 1, 2, 3$.*

Proof idea

- $H(X) \leq h < \log 2 \iff X = \begin{cases} \tilde{X} & \text{w.p. } \epsilon < 1/2 \\ x_0 & \text{w.p. } 1 - \epsilon > 1/2 \end{cases}$
- need $\mathbb{E}[\tilde{X}^n] = s_n := \frac{1}{\epsilon}(\mathbb{E}[W^n] - (1 - \epsilon)x_0^n)$

$\square$

Only three moments can be matched with small entropy

Theorem

For any continuous W, there exists $\eta_W \in (0, \frac{1}{2})$ such that

(i) *if X has $\mathbb{E}[X^n] = \mathbb{E}[W^n]$ for $n = 1, 2, 3, 4$, then $H(X) \geq h_2(\eta_W)$.*

(ii) *for any $h > 0$, there is X with $H(X) \leq h$ and $\mathbb{E}[X^n] = \mathbb{E}[W^n]$ for $n = 1, 2, 3$.*

Proof idea

- $H(X) \leq h < \log 2 \iff X = \begin{cases} \tilde{X} & \text{w.p. } \epsilon < 1/2 \\ x_0 & \text{w.p. } 1 - \epsilon > 1/2 \end{cases}$
- need $\mathbb{E}[\tilde{X}^n] = s_n := \frac{1}{\epsilon}(\mathbb{E}[W^n] - (1 - \epsilon)x_0^n)$
- check: $s_1, \ldots, s_4$ "valid" iff $\epsilon > \eta_W$, but s_1, s_2, s_3 always "valid" $\square$

Only three moments can be matched with small entropy

Theorem

For any continuous W, there exists $\eta_W \in (0, \frac{1}{2})$ such that

(i) *if X has $\mathbb{E}[X^n] = \mathbb{E}[W^n]$ for $n = 1, 2, 3, 4$, then $H(X) \geq h_2(\eta_W)$.*

(ii) *for any $h > 0$, there is X with $H(X) \leq h$ and $\mathbb{E}[X^n] = \mathbb{E}[W^n]$ for $n = 1, 2, 3$.*

Proof idea

- $H(X) \leq h < \log 2 \iff X = \begin{cases} \tilde{X} & \text{w.p. } \epsilon < 1/2 \\ x_0 & \text{w.p. } 1 - \epsilon > 1/2 \end{cases}$
- need $\mathbb{E}[\tilde{X}^n] = s_n := \frac{1}{\epsilon}(\mathbb{E}[W^n] - (1-\epsilon)x_0^n)$
- check: $s_1, \ldots, s_4$ "valid" iff $\epsilon > \eta_W$, but s_1, s_2, s_3 always "valid" $\square$

Corollary

$\eta_G = \frac{1}{3}$

Only three moments can be matched with small entropy

Theorem

For any continuous W, there exists $\eta_W \in (0, \frac{1}{2})$ such that

(i) *if X has $\mathbb{E}[X^n] = \mathbb{E}[W^n]$ for $n = 1, 2, 3, 4$, then $H(X) \geq h_2(\eta_W)$.*

(ii) *for any $h > 0$, there is X with $H(X) \leq h$ and $\mathbb{E}[X^n] = \mathbb{E}[W^n]$ for $n = 1, 2, 3$.*

Proof idea

- $H(X) \leq h < \log 2 \iff X = \begin{cases} \tilde{X} & \text{w.p. } \epsilon < 1/2 \\ x_0 & \text{w.p. } 1 - \epsilon > 1/2 \end{cases}$
- need $\mathbb{E}[\tilde{X}^n] = s_n := \frac{1}{\epsilon}(\mathbb{E}[W^n] - (1-\epsilon)x_0^n)$
- check: $s_1, \ldots, s_4$ "valid" iff $\epsilon > \eta_W$, but s_1, s_2, s_3 always "valid" $\square$

Corollary

$\eta_G = \frac{1}{3}$, *so for $h < h_2(\frac{1}{3})$, as* $\mathrm{snr} \to 0$, $C(\mathrm{snr}) - C_H(h, \mathrm{snr}) = \mathrm{O}(\mathrm{snr}^4)$.

Only three moments can be matched with small entropy

Theorem

For any continuous W, there exists $\eta_W \in (0, \frac{1}{2})$ such that

(i) *if X has $\mathbb{E}[X^n] = \mathbb{E}[W^n]$ for $n = 1, 2, 3, 4$, then $H(X) \geq h_2(\eta_W)$.*

(ii) *for any $h > 0$, there is X with $H(X) \leq h$ and $\mathbb{E}[X^n] = \mathbb{E}[W^n]$ for $n = 1, 2, 3$.*

Only three moments can be matched with small entropy

Theorem

For any continuous W, there exists $\eta_W \in (0, \frac{1}{2})$ such that

(i) *if X has $\mathbb{E}[X^n] = \mathbb{E}[W^n]$ for $n = 1, 2, 3, 4$, then $H(X) \geq h_2(\eta_W)$.*

(ii) *for any $h > 0$, there is X with $H(X) \leq h$ and $\mathbb{E}[X^n] = \mathbb{E}[W^n]$ for $n = 1, 2, 3$.*

For symmetric W,

$$\eta_W = \begin{cases} \dfrac{\mathbb{E}[W^2]^2}{\mathbb{E}[X^4]} & \text{if } \frac{\mathbb{E}[W^2]^2}{\mathbb{E}[X^4]} \leq \frac{1}{3}, \\[2ex] \dfrac{5\mathbb{E}[W^2]^2 - \mathbb{E}[W^4]}{9\mathbb{E}[W^2]^2 - \mathbb{E}[W^4]} & \text{if } \frac{\mathbb{E}[W^2]^2}{\mathbb{E}[X^4]} > \frac{1}{3}. \end{cases}$$

Only three moments can be matched with small entropy

Theorem

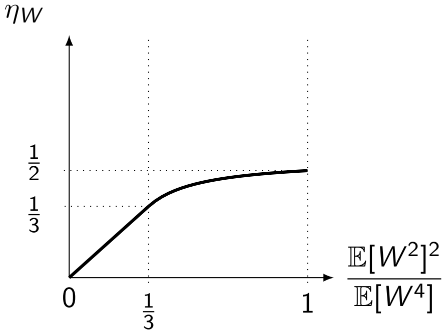
For any continuous W, there exists $\eta_W \in (0, \frac{1}{2})$ such that

(i) *if X has $\mathbb{E}[X^n] = \mathbb{E}[W^n]$ for $n = 1, 2, 3, 4$, then $H(X) \geq h_2(\eta_W)$.*

(ii) *for any $h > 0$, there is X with $H(X) \leq h$ and $\mathbb{E}[X^n] = \mathbb{E}[W^n]$ for $n = 1, 2, 3$.*

For symmetric W,

$$\eta_W = \begin{cases} \dfrac{\mathbb{E}[W^2]^2}{\mathbb{E}[X^4]} & \text{if } \frac{\mathbb{E}[W^2]^2}{\mathbb{E}[X^4]} \leq \frac{1}{3}, \\[2ex] \dfrac{5\mathbb{E}[W^2]^2 - \mathbb{E}[W^4]}{9\mathbb{E}[W^2]^2 - \mathbb{E}[W^4]} & \text{if } \frac{\mathbb{E}[W^2]^2}{\mathbb{E}[X^4]} > \frac{1}{3}. \end{cases}$$



Only three moments can be matched with small entropy

Theorem

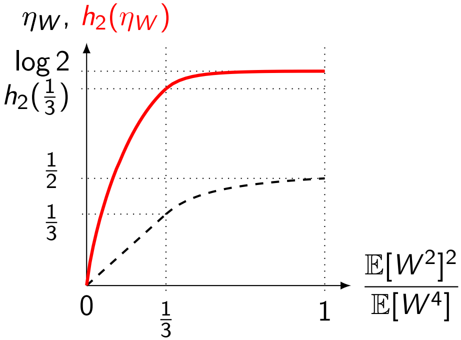
For any continuous W, there exists $\eta_W \in (0, \frac{1}{2})$ such that

(i) *if X has $\mathbb{E}[X^n] = \mathbb{E}[W^n]$ for $n = 1, 2, 3, 4$, then $H(X) \geq h_2(\eta_W)$.*

(ii) *for any $h > 0$, there is X with $H(X) \leq h$ and $\mathbb{E}[X^n] = \mathbb{E}[W^n]$ for $n = 1, 2, 3$.*

For symmetric W,

$$\eta_W = \begin{cases} \dfrac{\mathbb{E}[W^2]^2}{\mathbb{E}[X^4]} & \text{if } \frac{\mathbb{E}[W^2]^2}{\mathbb{E}[X^4]} \leq \frac{1}{3}, \\[2ex] \dfrac{5\mathbb{E}[W^2]^2 - \mathbb{E}[W^4]}{9\mathbb{E}[W^2]^2 - \mathbb{E}[W^4]} & \text{if } \frac{\mathbb{E}[W^2]^2}{\mathbb{E}[X^4]} > \frac{1}{3}. \end{cases}$$



"Application": Testing discrete against Gaussian in high noise

- observe n i.i.d. samples $Y_i = X_i + Z_i$

“Application”: Testing discrete against Gaussian in high noise

- observe n i.i.d. samples $Y_i = X_i + Z_i$ with $Z_i \sim Z \sim \mathcal{N}(0, \sigma^2)$ independent of X_i, $\sigma \to \infty$

“Application”: Testing discrete against Gaussian in high noise

- observe n i.i.d. samples $Y_i = X_i + Z_i$ with $Z_i \sim Z \sim \mathcal{N}(0, \sigma^2)$ independent of X_i, $\sigma \to \infty$

- $$\left.\begin{array}{ll} \mathcal{H} = 0: & X_i \sim X \text{ discrete with } H(X) \le h \\ \mathcal{H} = 1: & X_i \sim G \sim \mathcal{N}(0,1) \end{array}\right\}$$

“Application”: Testing discrete against Gaussian in high noise

- observe n i.i.d. samples $Y_i = X_i + Z_i$ with $Z_i \sim Z \sim \mathcal{N}(0, \sigma^2)$ independent of X_i, $\sigma \to \infty$

- $$\left.\begin{array}{ll} \mathcal{H} = 0: & X_i \sim X \text{ discrete with } H(X) \le h \\ \mathcal{H} = 1: & X_i \sim G \sim \mathcal{N}(0,1) \end{array}\right\} \text{decide } \hat{\mathcal{H}} \in \{0,1\} \text{ from } Y^n$$

"Application": Testing discrete against Gaussian in high noise

- observe n i.i.d. samples $Y_i = X_i + Z_i$ with $Z_i \sim Z \sim \mathcal{N}(0, \sigma^2)$ independent of X_i, $\sigma \to \infty$
- $$\left.\begin{array}{ll} \mathcal{H} = 0: & X_i \sim X \text{ discrete with } H(X) \leq h \\ \mathcal{H} = 1: & X_i \sim G \sim \mathcal{N}(0,1) \end{array}\right\} \text{decide } \hat{\mathcal{H}} \in \{0,1\} \text{ from } Y^n$$
- type-I error: $\Pr\{\hat{\mathcal{H}} = 1 \mid \mathcal{H} = 0\}$ and type-II error: $\Pr\{\hat{\mathcal{H}} = 0 \mid \mathcal{H} = 1\}$

"Application": Testing discrete against Gaussian in high noise

- observe n i.i.d. samples $Y_i = X_i + Z_i$ with $Z_i \sim Z \sim \mathcal{N}(0, \sigma^2)$ independent of X_i, $\sigma \to \infty$

- $$\left.\begin{array}{ll} \mathcal{H} = 0: & X_i \sim X \text{ discrete with } H(X) \leq h \\ \mathcal{H} = 1: & X_i \sim G \sim \mathcal{N}(0,1) \end{array}\right\} \text{decide } \hat{\mathcal{H}} \in \{0,1\} \text{ from } Y^n$$

- type-I error: $\Pr\{\hat{\mathcal{H}} = 1 \mid \mathcal{H} = 0\}$ and type-II error: $\Pr\{\hat{\mathcal{H}} = 0 \mid \mathcal{H} = 1\}$

- Stein regime:
  type-I error $\leq \epsilon < 1$

"Application": Testing discrete against Gaussian in high noise

- observe n i.i.d. samples $Y_i = X_i + Z_i$ with $Z_i \sim Z \sim \mathcal{N}(0, \sigma^2)$ independent of X_i, $\sigma \to \infty$

- $$\left.\begin{array}{ll} \mathcal{H} = 0: & X_i \sim X \text{ discrete with } H(X) \le h \\ \mathcal{H} = 1: & X_i \sim G \sim \mathcal{N}(0,1) \end{array}\right\} \text{decide } \hat{\mathcal{H}} \in \{0,1\} \text{ from } Y^n$$

- type-I error: $\Pr\{\hat{\mathcal{H}} = 1 \mid \mathcal{H} = 0\}$ and type-II error: $\Pr\{\hat{\mathcal{H}} = 0 \mid \mathcal{H} = 1\}$

- Stein regime:
  type-I error $\le \epsilon < 1$ and type-II error $\le \exp\left[-n\, \mathrm{D}_{\mathrm{KL}}(X + Z \,\|\, G + Z) + \mathrm{o}(n)\right]$

“Application”: Testing discrete against Gaussian in high noise

- observe n i.i.d. samples $Y_i = X_i + Z_i$ with $Z_i \sim Z \sim \mathcal{N}(0, \sigma^2)$ independent of X_i, $\sigma \to \infty$

- $$\left.\begin{array}{ll} \mathcal{H} = 0: & X_i \sim X \text{ discrete with } H(X) \le h \\ \mathcal{H} = 1: & X_i \sim G \sim \mathcal{N}(0,1) \end{array}\right\} \text{decide } \hat{\mathcal{H}} \in \{0,1\} \text{ from } Y^n$$

- type-I error: $\Pr\{\hat{\mathcal{H}} = 1 \mid \mathcal{H} = 0\}$ and type-II error: $\Pr\{\hat{\mathcal{H}} = 0 \mid \mathcal{H} = 1\}$

- Stein regime:
  type-I error $\le \epsilon < 1$ and type-II error $\le \exp\left[-n\,\mathrm{D_{KL}}(X + Z \,\|\, G + Z) + o(n)\right]$

- worst-case exponent $= \min\limits_{X:\,H(X) \le h} \mathrm{D_{KL}}(X + Z \,\|\, G + Z)$

"Application": Testing discrete against Gaussian in high noise

- observe n i.i.d. samples $Y_i = X_i + Z_i$ with $Z_i \sim Z \sim \mathcal{N}(0, \sigma^2)$ independent of X_i, $\sigma \to \infty$
- $$\left.\begin{array}{ll} \mathcal{H} = 0: & X_i \sim X \text{ discrete with } H(X) \le h \\ \mathcal{H} = 1: & X_i \sim G \sim \mathcal{N}(0,1) \end{array}\right\} \text{decide } \hat{\mathcal{H}} \in \{0,1\} \text{ from } Y^n$$
- type-I error: $\Pr\{\hat{\mathcal{H}} = 1 \mid \mathcal{H} = 0\}$ and type-II error: $\Pr\{\hat{\mathcal{H}} = 0 \mid \mathcal{H} = 1\}$
- Stein regime:
  type-I error $\le \epsilon < 1$ and type-II error $\le \exp\left[-n\,\mathrm{D}_{\mathrm{KL}}(X + Z \,\|\, G + Z) + o(n)\right]$
- worst-case exponent $= \min_{X:H(X)\le h} \mathrm{D}_{\mathrm{KL}}(X + Z \,\|\, G + Z) \quad = \Theta\left(\frac{1}{\sigma^8}\right)$ for $h < h_2\left(\frac{1}{3}\right)$

"Application": Testing discrete against Gaussian in high noise

- observe n i.i.d. samples $Y_i = X_i + Z_i$ with $Z_i \sim Z \sim \mathcal{N}(0, \sigma^2)$ independent of X_i, $\sigma \to \infty$

- $$\left.\begin{array}{ll} \mathcal{H} = 0: & X_i \sim X \text{ discrete with } H(X) \leq h \\ \mathcal{H} = 1: & X_i \sim G \sim \mathcal{N}(0,1) \end{array}\right\} \text{decide } \hat{\mathcal{H}} \in \{0,1\} \text{ from } Y^n$$

- type-I error: $\Pr\{\hat{\mathcal{H}} = 1 \mid \mathcal{H} = 0\}$ and type-II error: $\Pr\{\hat{\mathcal{H}} = 0 \mid \mathcal{H} = 1\}$

- Stein regime:
  type-I error $\leq \epsilon < 1$ and type-II error $\leq \exp\left[-n \, \mathrm{D_{KL}}(X + Z \,\|\, G + Z) + o(n)\right]$

- worst-case exponent $= \min\limits_{X : H(X) \leq h} \mathrm{D_{KL}}(X + Z \,\|\, G + Z) \qquad = \Theta\left(\frac{1}{\sigma^8}\right)$ for $h < h_2\left(\frac{1}{3}\right)$

- proof: Taylor expansion, k^{th} term depends on first k moments, can match only 3 moments

Summary

- entropy-constrained Gaussian channel

Summary

- entropy-constrained Gaussian channel capacity $C_H(h, \mathsf{snr})$

Summary

- entropy-constrained Gaussian channel capacity $C_H(h, \mathrm{snr})$
- snr $\to 0$, $0 < h < h_2(\frac{1}{3})$: $C_H(h, \mathrm{snr}) = C(\mathrm{snr}) - \mathrm{O}(\mathrm{snr}^4)$

Summary

- entropy-constrained Gaussian channel capacity $C_H(h, \mathsf{snr})$
- $\mathsf{snr} \to 0$, $0 < h < h_2(\frac{1}{3})$: $C_H(h, \mathsf{snr}) = C(\mathsf{snr}) - \mathrm{O}(\mathsf{snr}^4)$, via entropy-constrained version of truncated moment problem

Summary

- entropy-constrained Gaussian channel capacity $C_H(h, \mathsf{snr})$
- $\mathsf{snr} \to 0$, $0 < h < h_2(\frac{1}{3})$: $C_H(h, \mathsf{snr}) = C(\mathsf{snr}) - \mathrm{O}(\mathsf{snr}^4)$, via entropy-constrained version of truncated moment problem
- for any continuous distribution, only three moments can be matched by a discrete distribution of sufficiently small entropy

Summary

- entropy-constrained Gaussian channel capacity $C_H(h, \mathrm{snr})$
- $\mathrm{snr} \to 0$, $0 < h < h_2(\frac{1}{3})$: $C_H(h, \mathrm{snr}) = C(\mathrm{snr}) - \mathrm{O}(\mathrm{snr}^4)$, via entropy-constrained version of truncated moment problem
- for any continuous distribution, only three moments can be matched by a discrete distribution of sufficiently small entropy

A.G., S.Shamai, E.Telatar, "On Entropy-Constrained Gaussian Channel Capacity via the Moment Problem."
— ISIT 2025
— INFORMS APS 2025
— arXiv:2501.13814

Summary

- entropy-constrained Gaussian channel capacity $C_H(h, \mathsf{snr})$
- $\mathsf{snr} \to 0$, $0 < h < h_2(\frac{1}{3})$: $C_H(h, \mathsf{snr}) = C(\mathsf{snr}) - \mathrm{O}(\mathsf{snr}^4)$, via entropy-constrained version of truncated moment problem
- for any continuous distribution, only three moments can be matched by a discrete distribution of sufficiently small entropy

  A.G., S.Shamai, E.Telatar, "On Entropy-Constrained Gaussian Channel Capacity via the Moment Problem."
  — ISIT 2025
  — INFORMS APS 2025
  — arXiv:2501.13814

- $\mathsf{snr} \to \infty$: "discrete Gaussian" seems to be "optimal"

Summary

- entropy-constrained Gaussian channel capacity $C_H(h, \mathsf{snr})$
- $\mathsf{snr} \to 0$, $0 < h < h_2(\frac{1}{3})$: $C_H(h, \mathsf{snr}) = C(\mathsf{snr}) - \mathrm{O}(\mathsf{snr}^4)$, via entropy-constrained version of truncated moment problem
- for any continuous distribution, only three moments can be matched by a discrete distribution of sufficiently small entropy

  A.G., S.Shamai, E.Telatar, "On Entropy-Constrained Gaussian Channel Capacity via the Moment Problem."
  — ISIT 2025
  — INFORMS APS 2025
  — arXiv:2501.13814

- $\mathsf{snr} \to \infty$: "discrete Gaussian" seems to be "optimal"

Thank you!