An Introduction to Rate-Distortion-Perception Theory

Jun Chen
McMaster University

North American School of Information Theory
June 22, 2026

Fidelity vs. Realism

Distortion Measure vs. Perception Measure

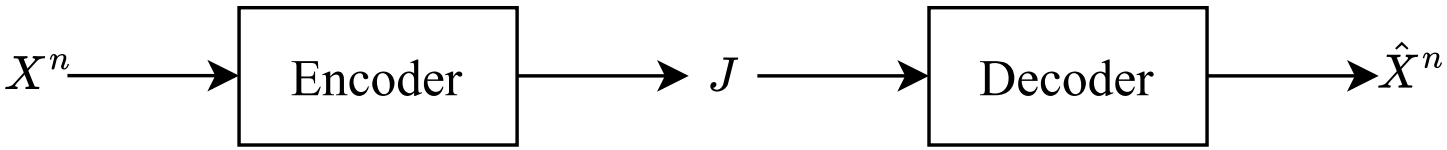
- Distortion measure: $d(x, \hat{x})$

$$d(x, \hat{x}) = \|x - \hat{x}\|^2$$

- Perception measure: $\phi(p_X, p_{\hat{X}})$

$$\phi(p_X, p_{\hat{X}}) = W_2^2(p_X, p_{\hat{X}}) = \inf_{p_{X\hat{X}} \in \Pi(p_X, p_{\hat{X}})} \mathbb{E}[\|X - \hat{X}\|^2]$$

Rate-Distortion Theory: Review



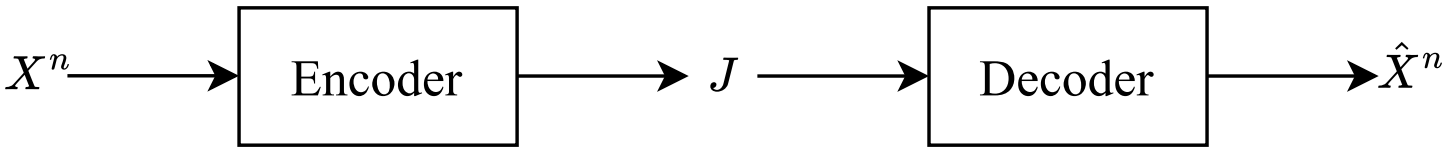
$$J \in \{1, 2, \ldots, M\}$$

Assume $X^n \sim p_X^n$:

$$D^{(\infty)}(R) = \lim_{n\to\infty} \inf_{\mathrm{E,D}} \frac{1}{n} \sum_{t=1}^{n} \mathbb{E}[\|X_t - \hat{X}_t\|^2]$$

$$\text{s.t.} \quad \frac{1}{n} \log M \le R$$

Rate-Distortion Theory: Review



$$J \in \{1, 2, \ldots, M\}$$

$$D^{(\infty)}(R) = \inf_{p_{\hat{X}|X}} \mathbb{E}[\|X - \hat{X}\|^2]$$

$$\text{s.t.} \quad I(X; \hat{X}) \leq R$$

Shannon's rate-distortion theorem

Rate-Distortion Theory: Review

- Test channel $p_{\hat{X}|X}$ satisfying $I(X;\hat{X}) \leq R$
  - Joint distribution $p_{X\hat{X}} = p_X p_{\hat{X}|X}$

$$\mathcal{T}(p_{X\hat{X}}) = \{(x^n, \hat{x}^n) : \text{ joint empirical distribution } \gamma_{x^n\hat{x}^n} \approx p_{X\hat{X}}\}$$

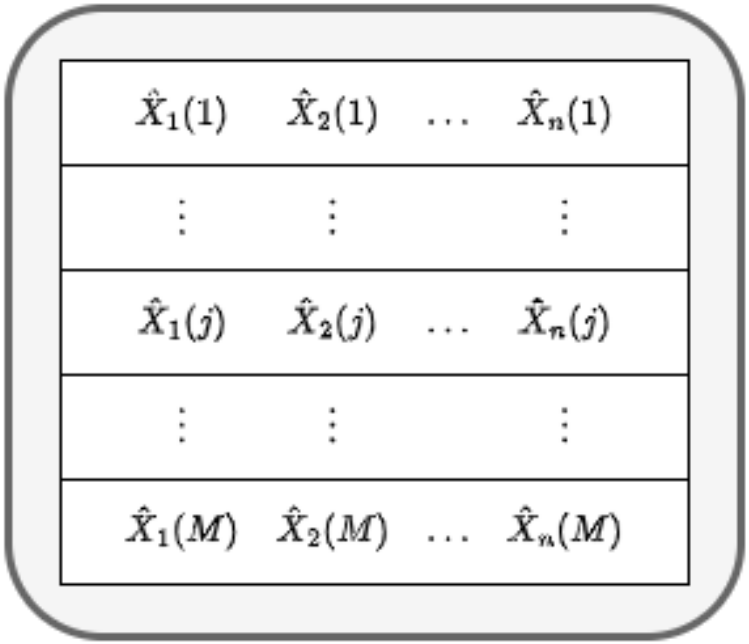
joint-typicality encoding

  - Output distribution $p_{\hat{X}}$

codebook generation

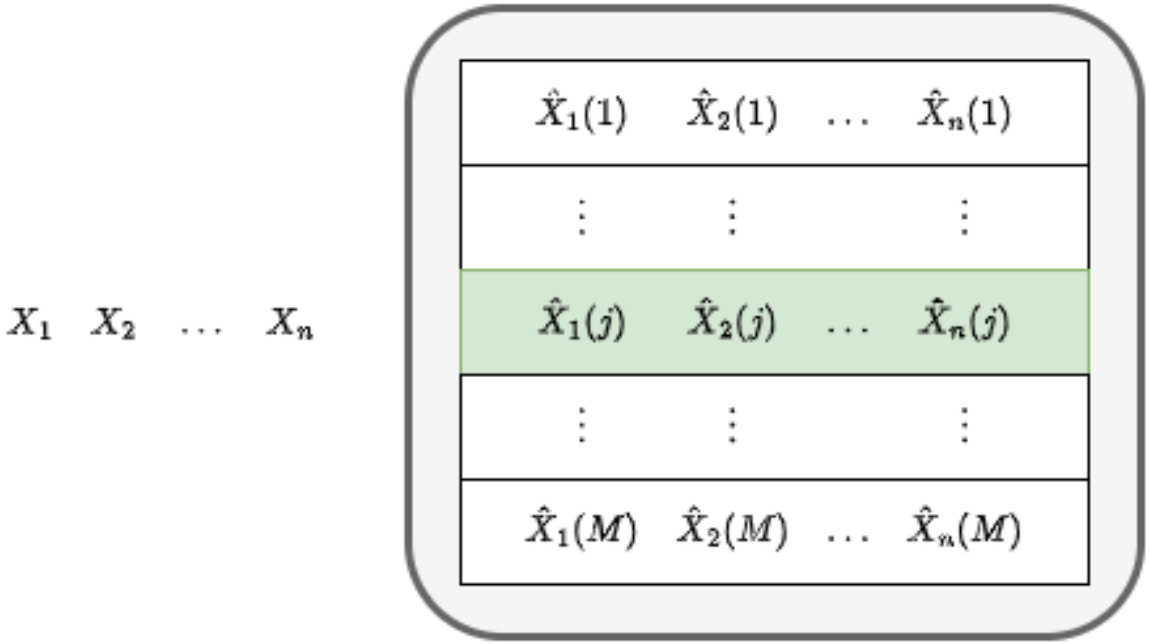
Rate-Distortion Theory: Review

- Codebook generation: $\hat{X}^n(j) \sim p_{\hat{X}}^n$, $j = 1, 2, \ldots, M \approx 2^{nI(X;\hat{X})}$



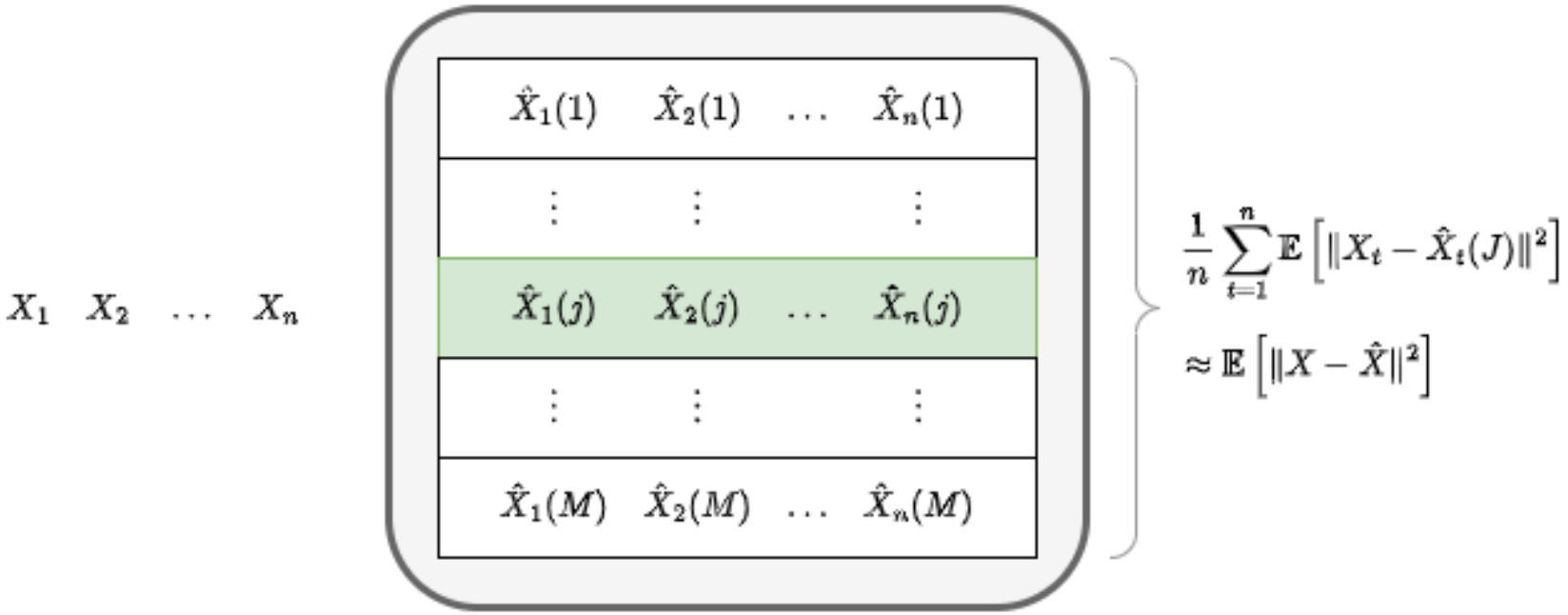
Rate-Distortion Theory: Review

- Encoding: $(X^n, \hat{X}^n(j)) \in \mathcal{T}(p_{X\hat{X}})$



Rate-Distortion Theory: Review

- Decoding:



Distortion-Rate-Perception Function

- Blau and Michaeli 2019

$$D^{(\infty)}(R, P) = \inf_{p_{\hat{X}|X}} \mathbb{E}[\|X - \hat{X}\|^2]$$

$$\text{s.t.} \quad I(X; \hat{X}) \le R$$

$$W_2^2(p_X, p_{\hat{X}}) \le P$$

Problem Formulation

- Test channel $p_{\hat{X}|X}$ satisfying $I(X;\hat{X}) \le R$ and $W_2^2(p_X, p_{\hat{X}}) \le P$
  - Joint distribution $p_{X\hat{X}} = p_X p_{\hat{X}|X}$

$$\mathcal{T}(p_{X\hat{X}}) = \{(x^n, \hat{x}^n) : \text{ joint empirical distribution } \gamma_{x^n\hat{x}^n} \approx p_{X\hat{X}}\}$$

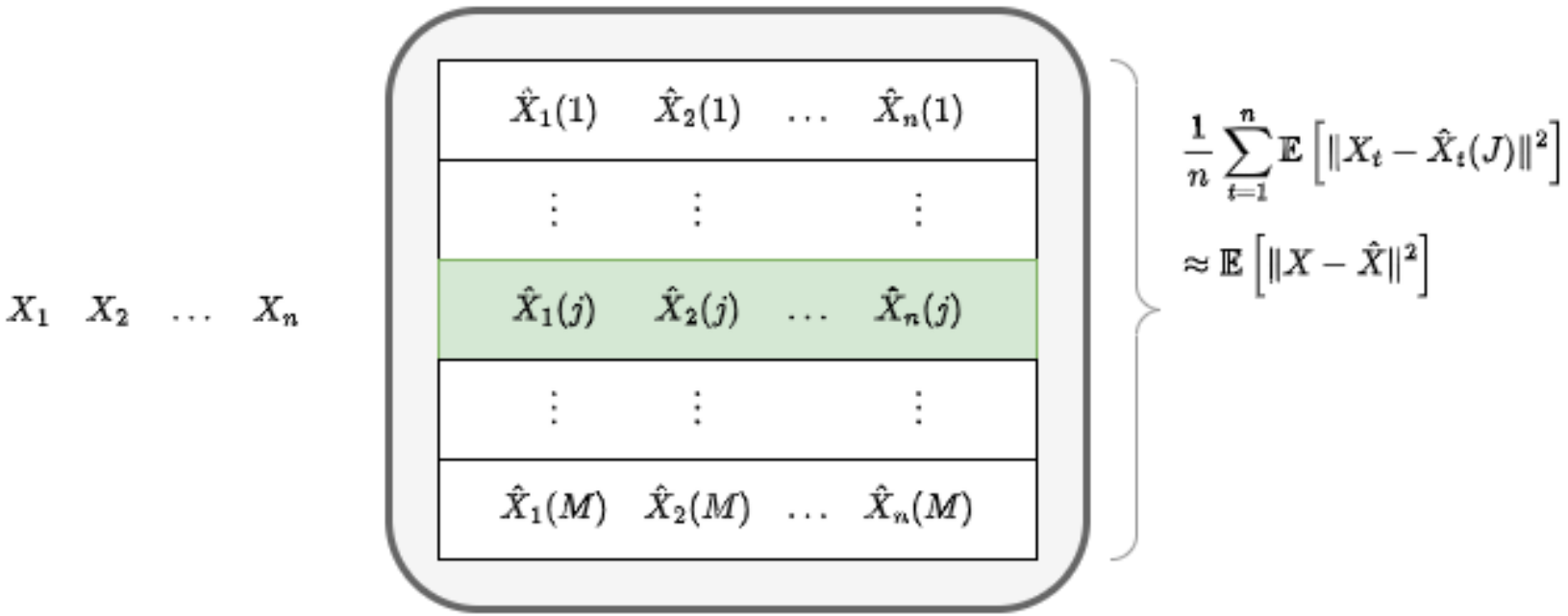
joint-typicality encoding

  - Output distribution $p_{\hat{X}}$

codebook generation

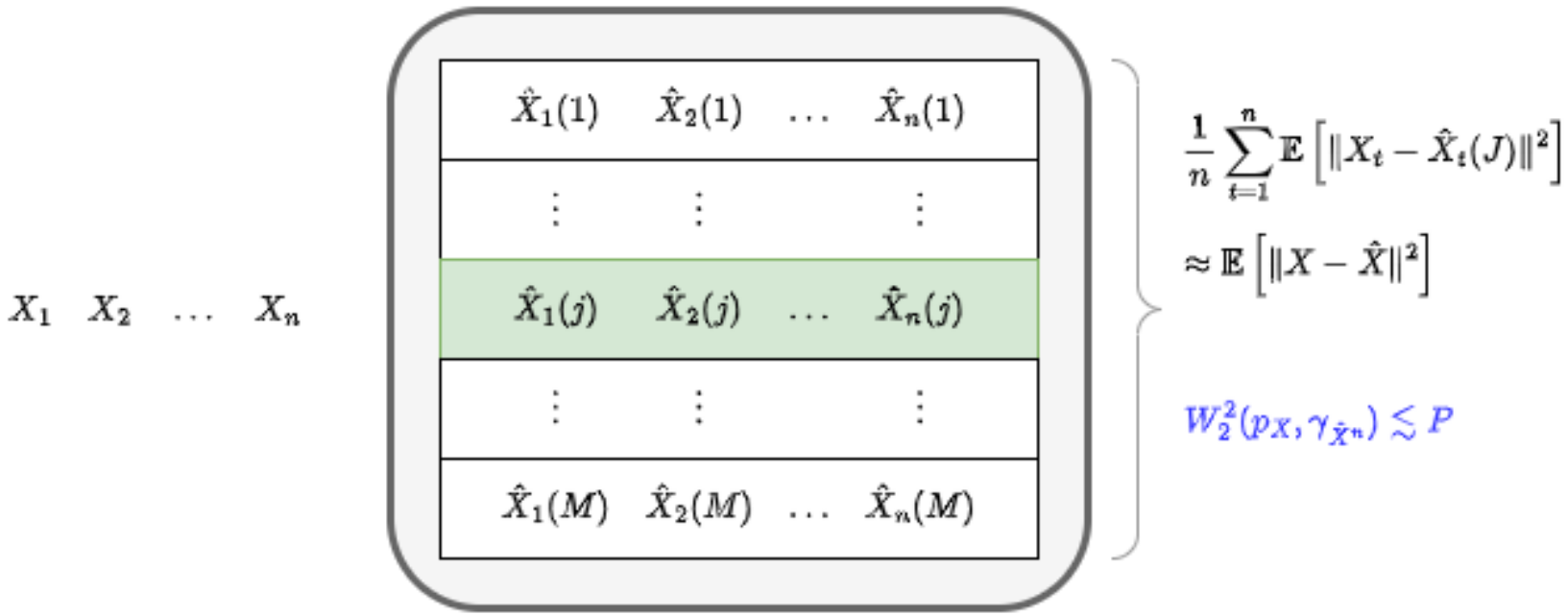
Problem Formulation

- Test channel $p_{\hat{X}|X}$ satisfying $I(X;\hat{X}) \le R$ and $W_2^2(p_X, p_{\hat{X}}) \le P$



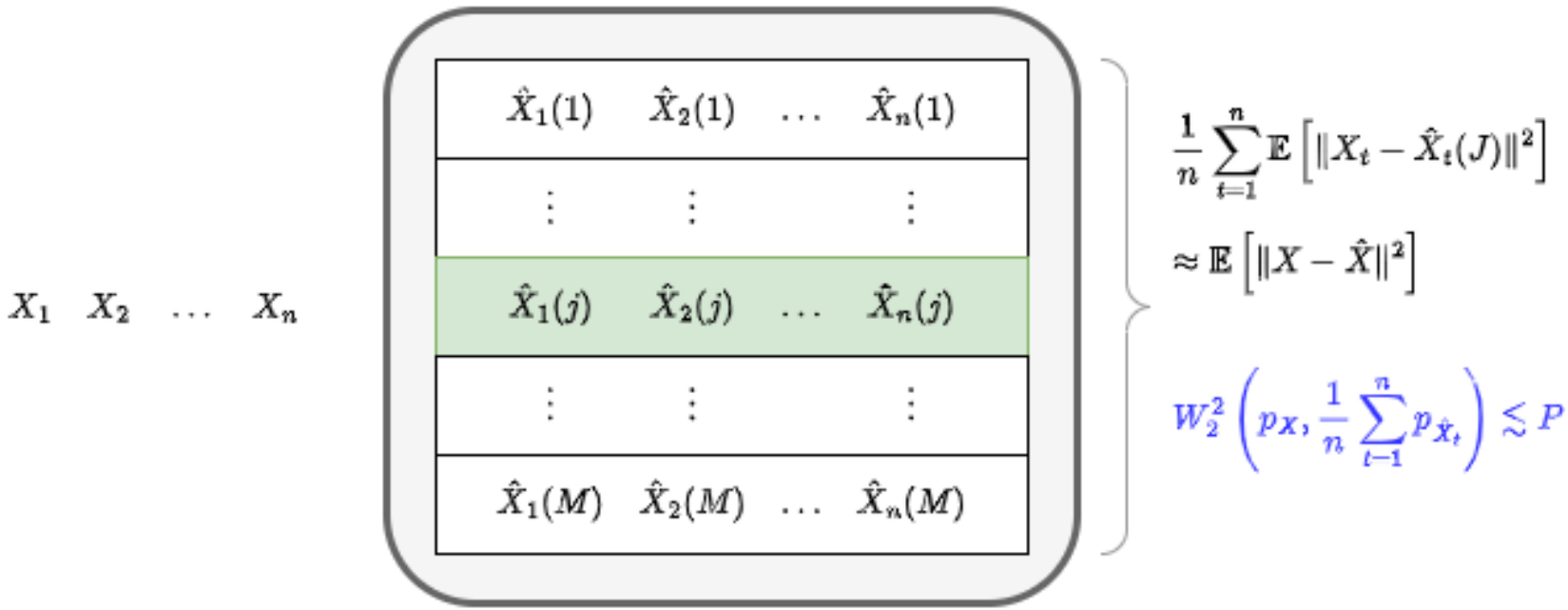
Problem Formulation

- Test channel $p_{\hat{X}|X}$ satisfying $I(X;\hat{X}) \le R$ and $W_2^2(p_X, p_{\hat{X}}) \le P$



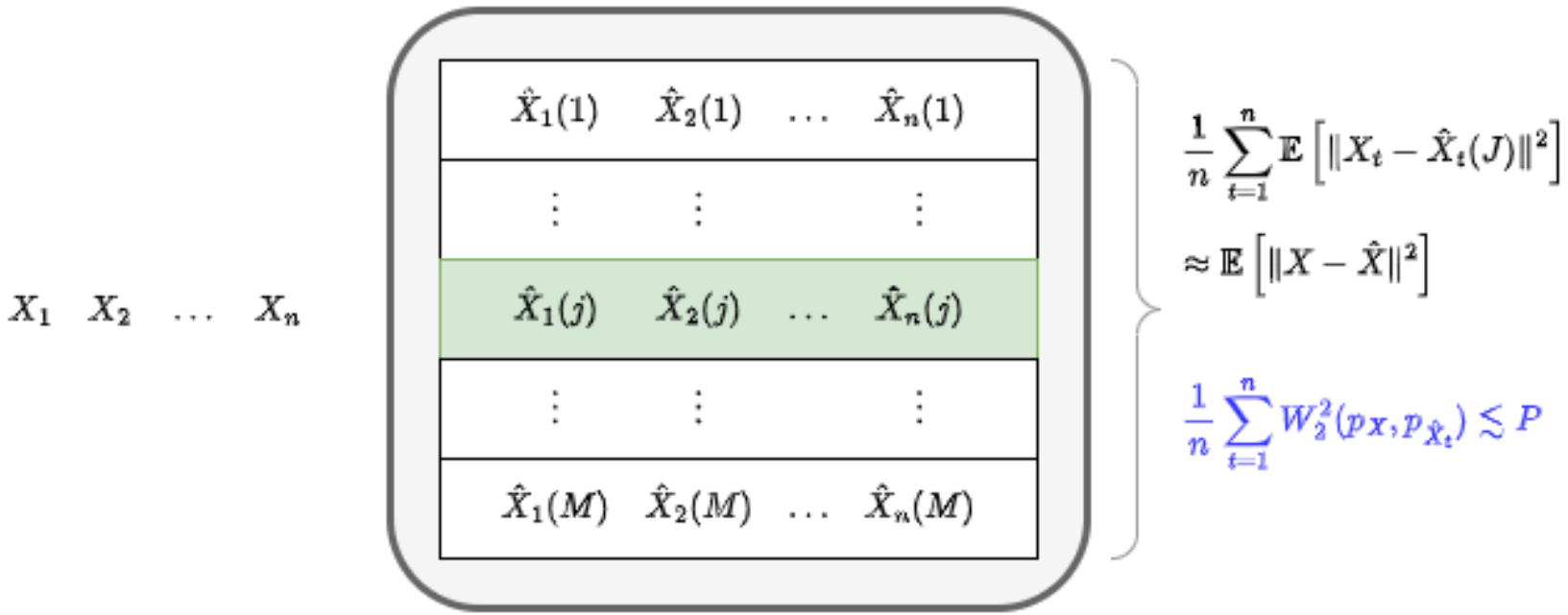
Problem Formulation

- Test channel $p_{\hat{X}|X}$ satisfying $I(X;\hat{X}) \le R$ and $W_2^2(p_X, p_{\hat{X}}) \le P$

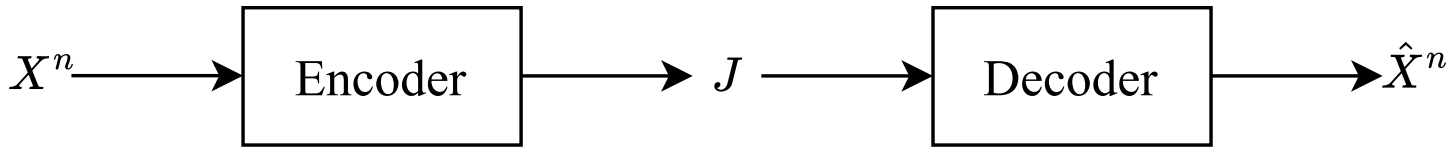


Problem Formulation

- Test channel $p_{\hat{X}|X}$ satisfying $I(X;\hat{X}) \le R$ and $W_2^2(p_X, p_{\hat{X}}) \le P$



Problem Formulation



$$J \in \{1, 2, \ldots, M\}$$

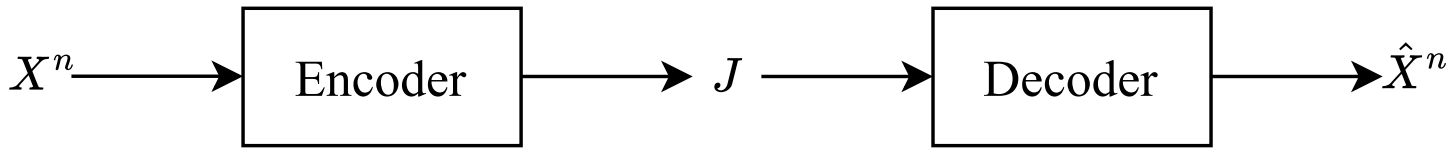
Assume $X^n \sim p_X^n$:

$$D^{(\infty)}(R, P) = \lim_{n\to\infty} \inf_{\mathrm{E,D}} \frac{1}{n} \sum_{t=1}^{n} \mathbb{E}[\|X_t - \hat{X}_t\|^2]$$

$$\text{s.t.} \quad \frac{1}{n} \log M \le R$$

$$\frac{1}{n} \sum_{t=1}^{n} W_2^2(p_X, p_{\hat{X}_t}) \le P$$

Coding Theorem



$$J \in \{1, 2, \ldots, M\}$$

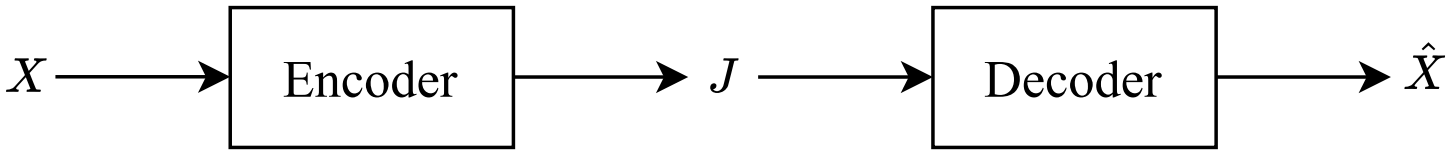
- Blau and Michaeli 2019, Theis and Wagner 2021, Chen et al. 2022

$$D^{(\infty)}(R, P) = \inf_{p_{\hat{X}|X}} \mathbb{E}[\|X - \hat{X}\|^2]$$

$$\text{s.t.} \quad I(X; \hat{X}) \le R$$

$$W_2^2(p_X, p_{\hat{X}}) \le P$$

One-Shot Setting



$$J \in \{1, 2, \ldots, M\}$$

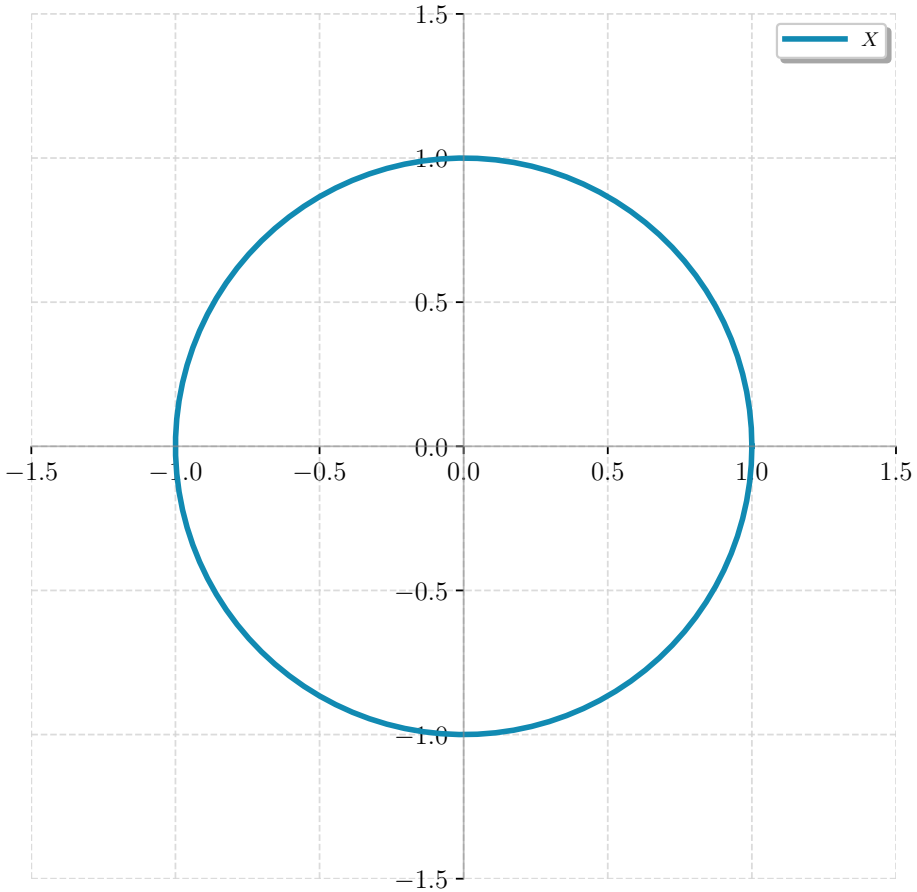
Assume $X \sim p_X$:

$$D^{(1)}(R, P) = \inf_{\mathrm{E,D}} \mathbb{E}[\|X - \hat{X}\|^2]$$

$$\text{s.t.} \quad \log M \leq R$$

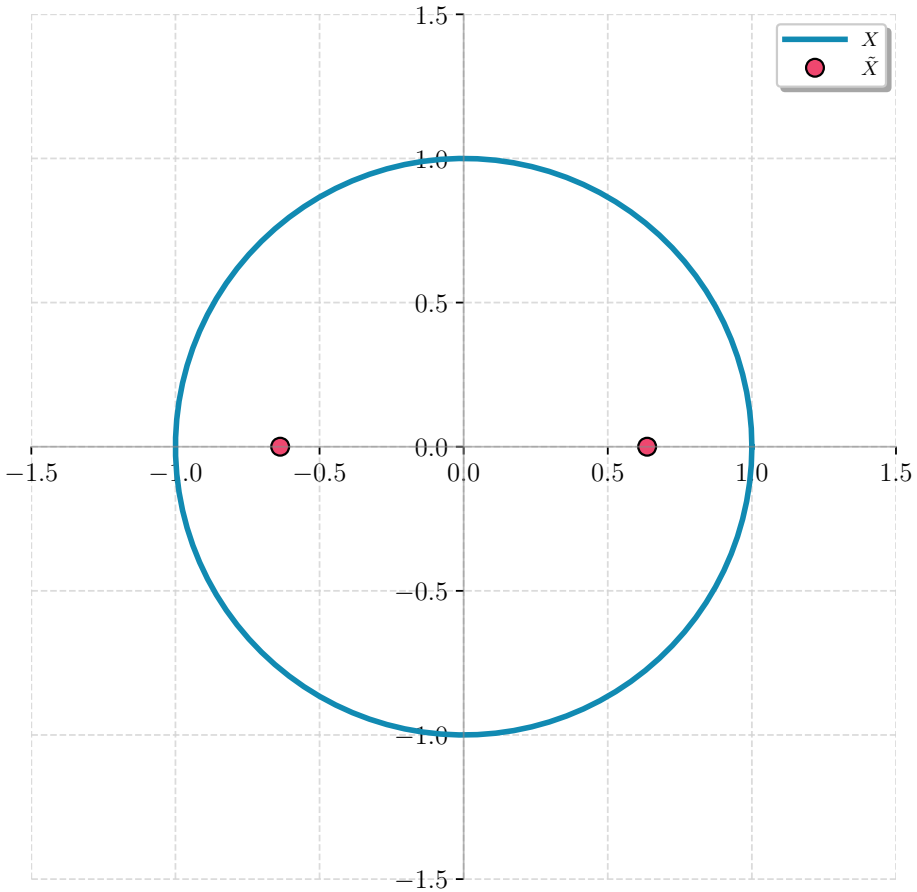
$$W_2^2(p_X, p_{\hat{X}}) \leq P$$

Quantizing the Unit Circle



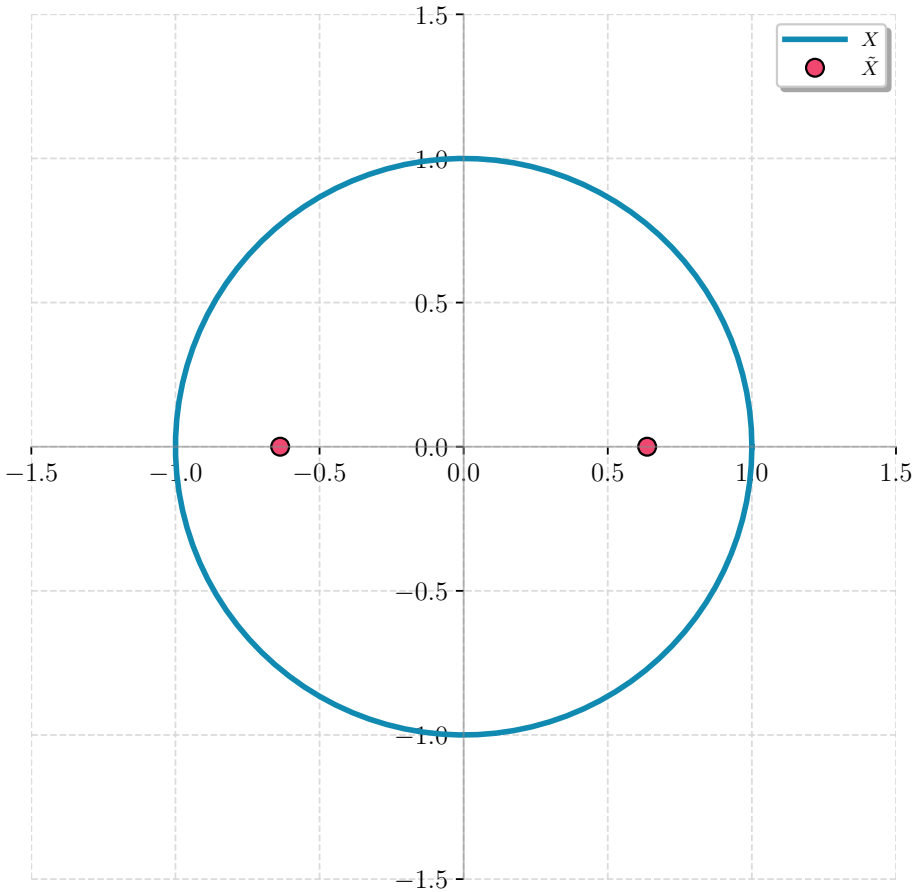
1-bit Quantization

Quantizing the Unit Circle: No Perception Constraint



$$\mathbb{E}[\|X - \tilde{X}\|^2] = 1 - \frac{4}{\pi^2}$$

Quantizing the Unit Circle: No Perception Constraint



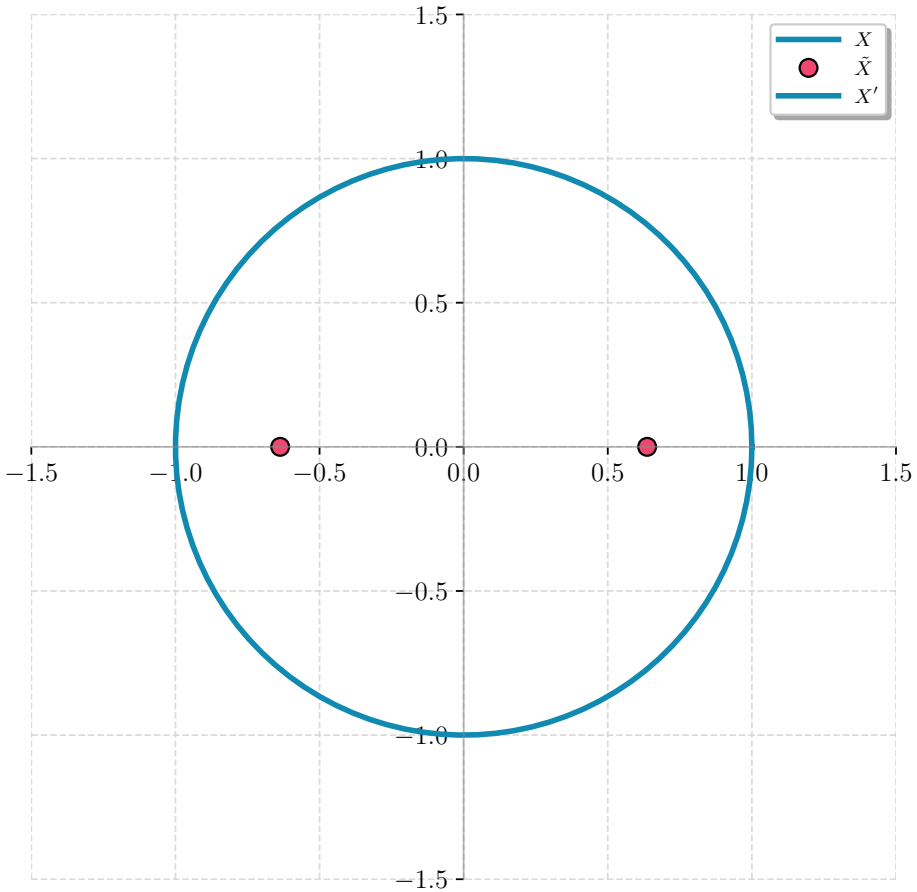
$$W_2^2(p_X, p_{\tilde{X}}) = 1 - \frac{4}{\pi^2}$$

General Observation I

- Yan et al. 2022

$$D^{(1)}(R, P) = D^{(1)}(R, \infty) \text{ if and only if } P \geq D^{(1)}(R, \infty)$$

Quantizing the Unit Circle: Perfect Perception



$$\mathbb{E}[\|X' - \tilde{X}\|^2] = W_2^2(p_X, p_{\tilde{X}}) = \mathbb{E}[\|X - \tilde{X}\|^2]$$

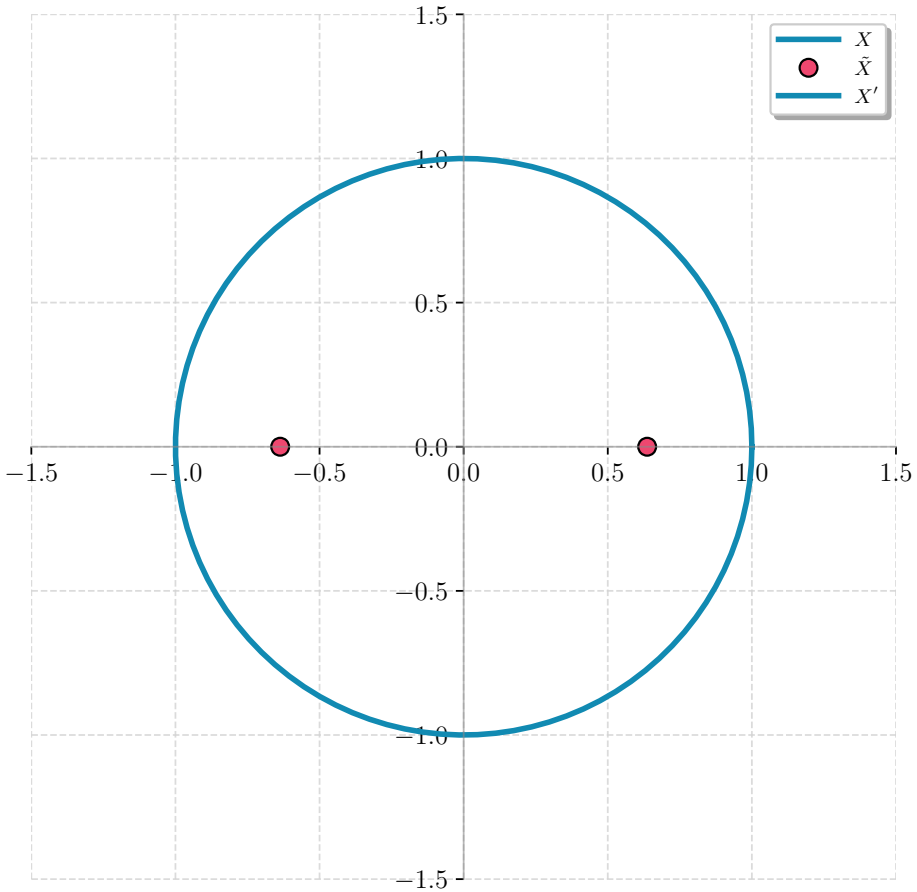
$$\mathbb{E}[\|X - X'\|^2] = \mathbb{E}[\|X - \tilde{X}\|^2] + \mathbb{E}[\|X' - \tilde{X}\|^2] = 2\mathbb{E}[\|X - \tilde{X}\|^2]$$

General Observation II

- Yan et al. 2021, Theis and Agustsson 2021

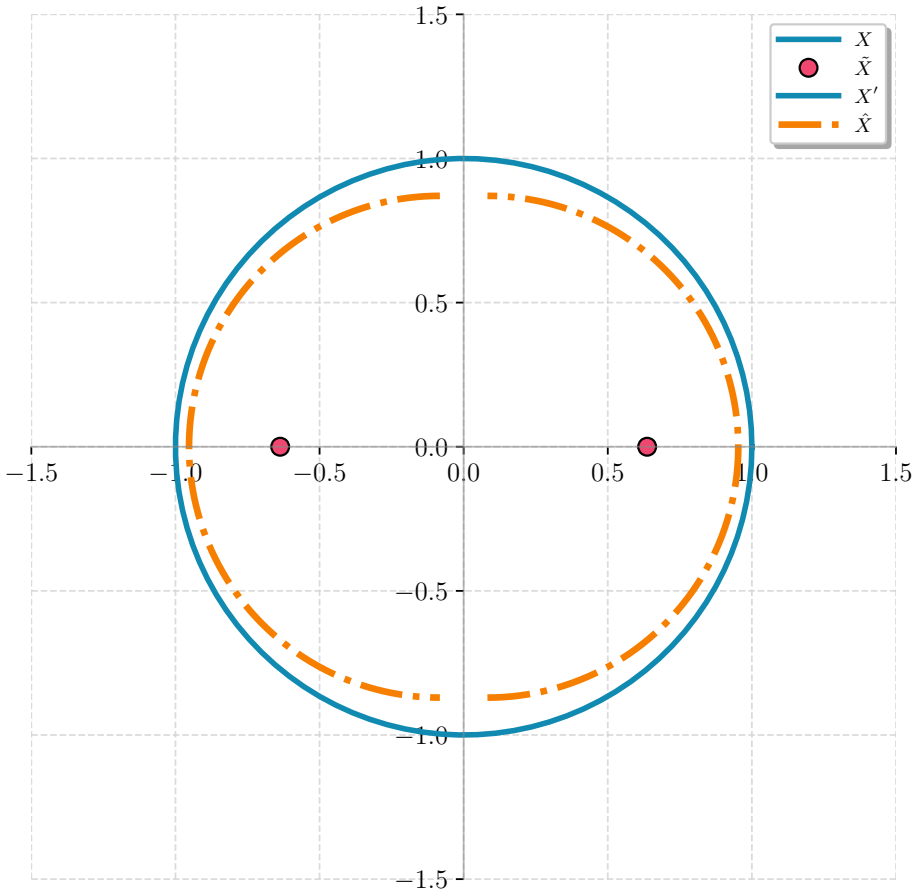
$$D^{(1)}(R, 0) = 2D^{(1)}(R, \infty)$$

Quantizing the Unit Circle: Linear Interpolation



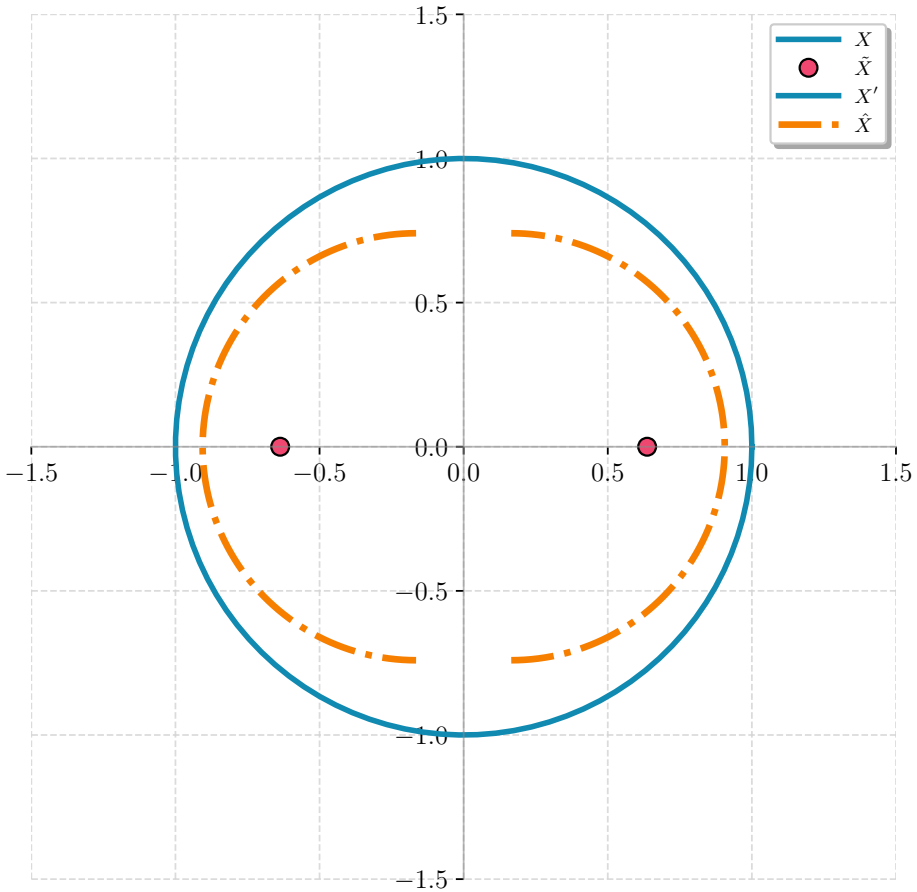
$$\hat{X} = (1-\alpha)X' + \alpha\tilde{X}$$

Quantizing the Unit Circle: Linear Interpolation



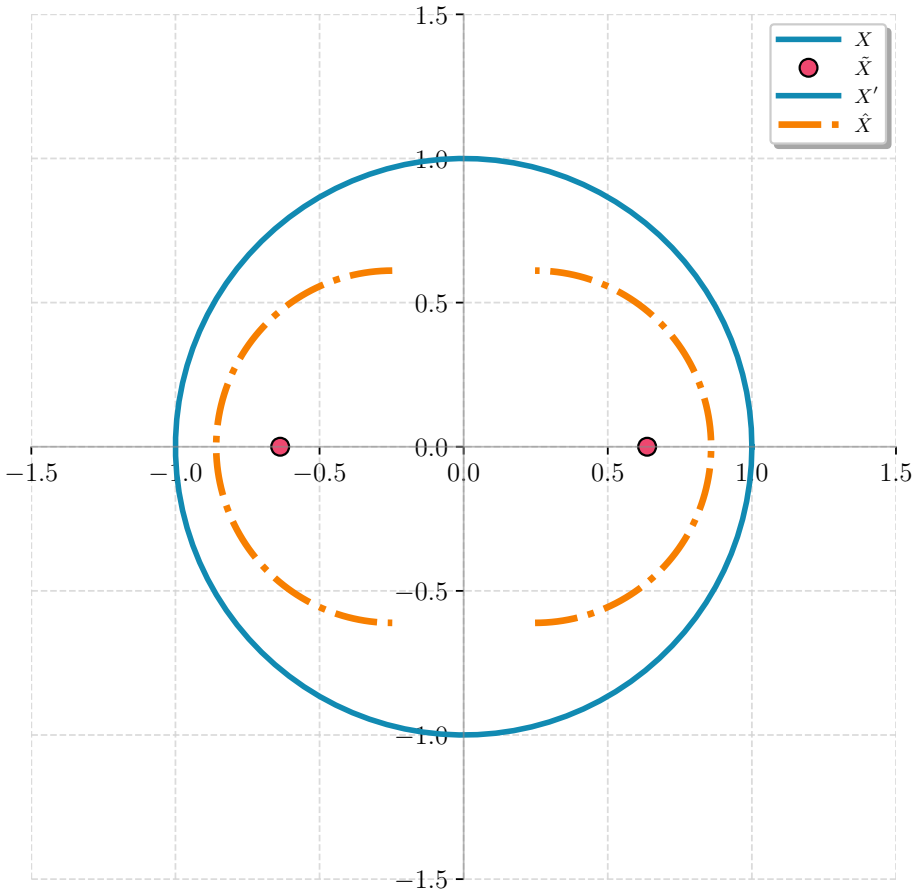
$$P = 0.01$$

Quantizing the Unit Circle: Linear Interpolation



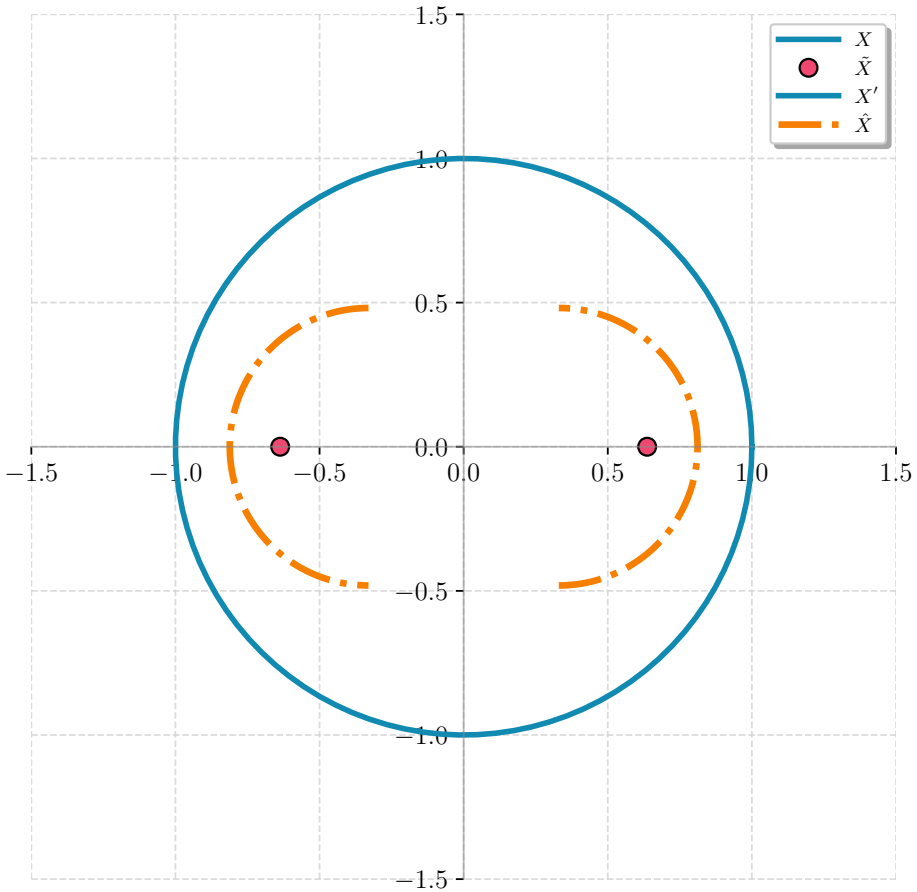
$$P = 0.04$$

Quantizing the Unit Circle: Linear Interpolation



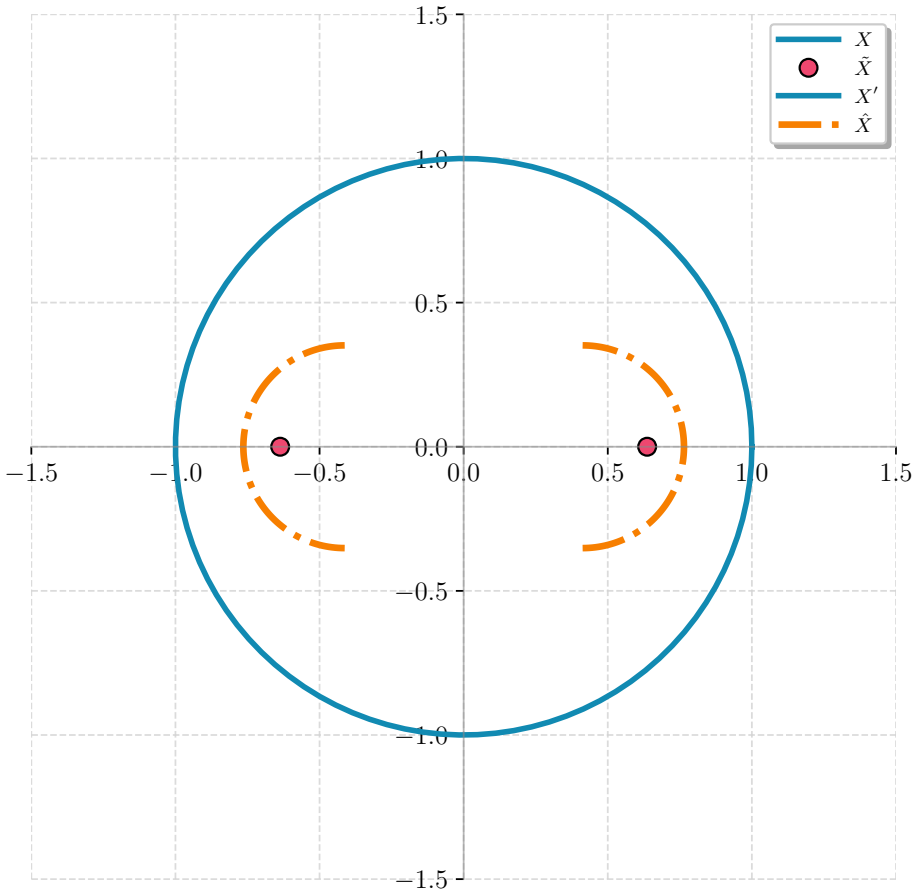
$$P = 0.09$$

Quantizing the Unit Circle: Linear Interpolation



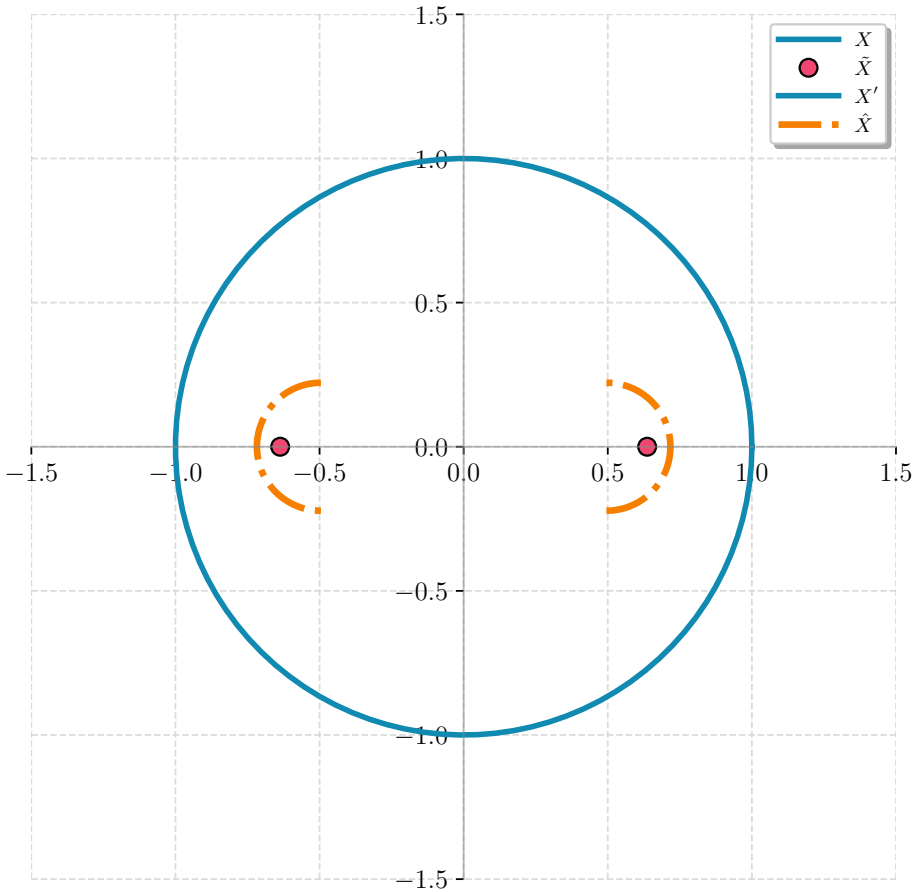
$$P = 0.16$$

Quantizing the Unit Circle: Linear Interpolation



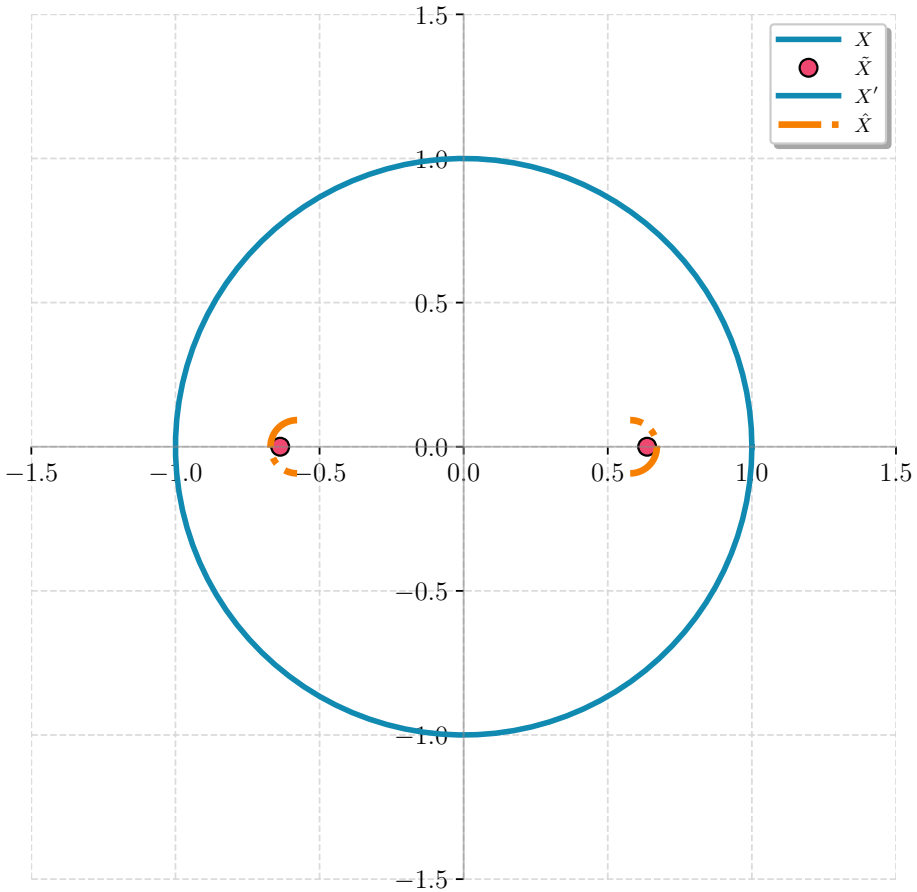
$$P = 0.25$$

Quantizing the Unit Circle: Linear Interpolation



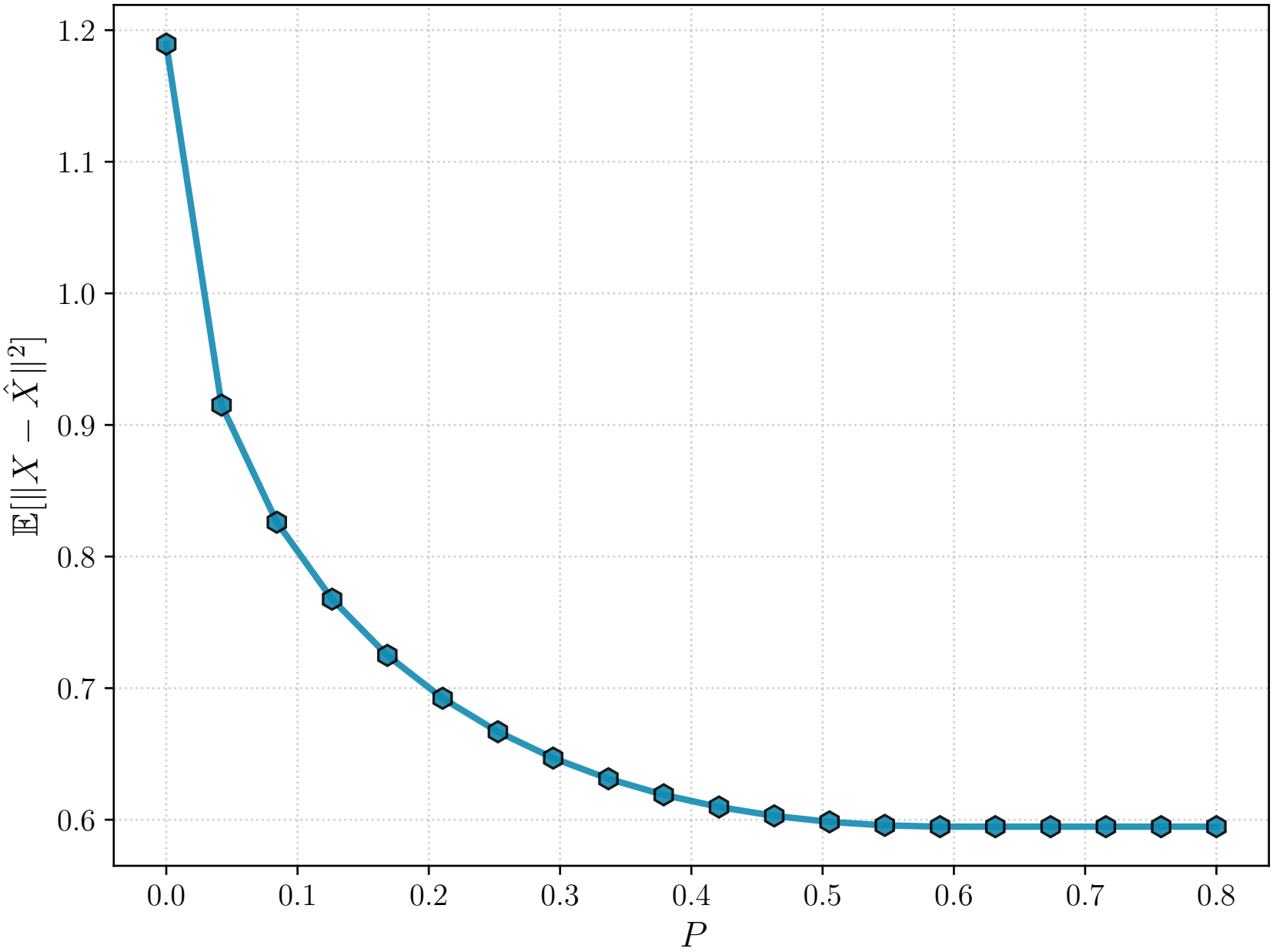
$$P = 0.36$$

Quantizing the Unit Circle: Linear Interpolation

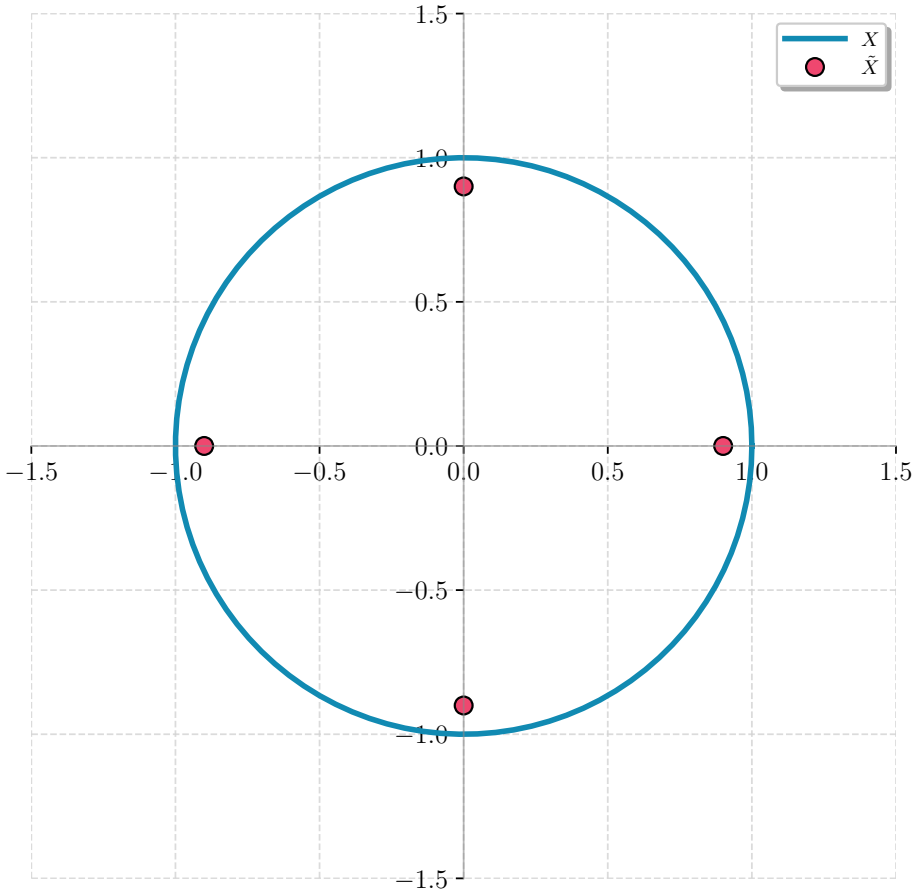


$$P = 0.49$$

Distortion-Perception Tradeoff

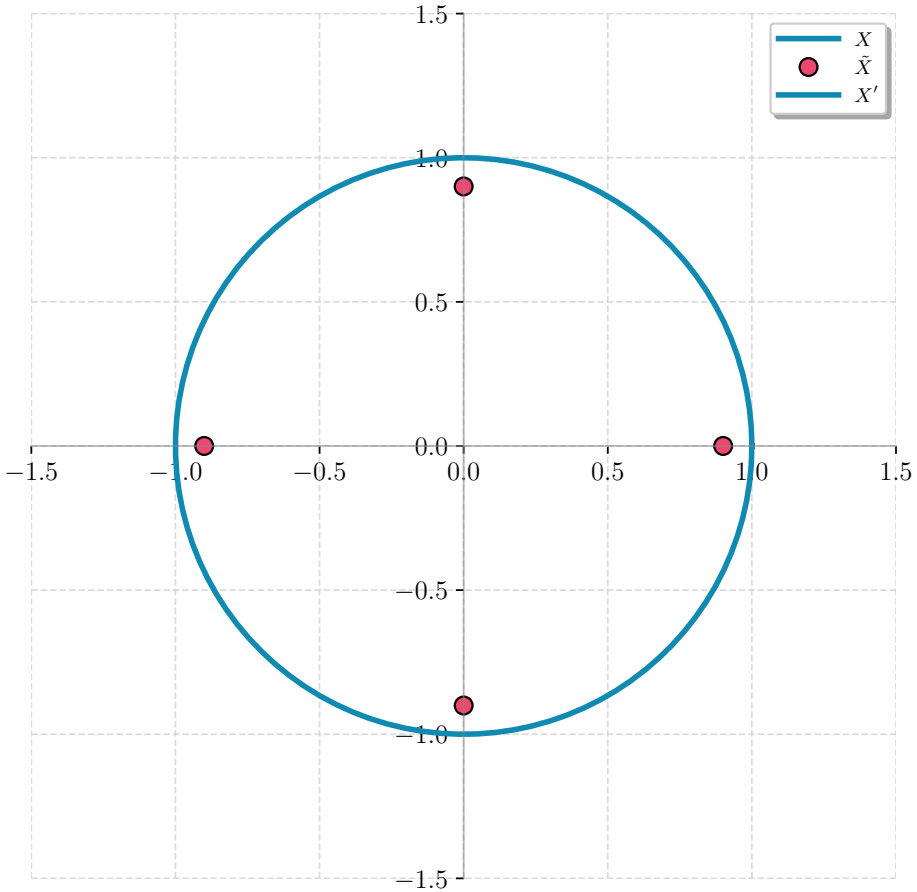


Quantizing the Unit Circle: No Perception Constraint



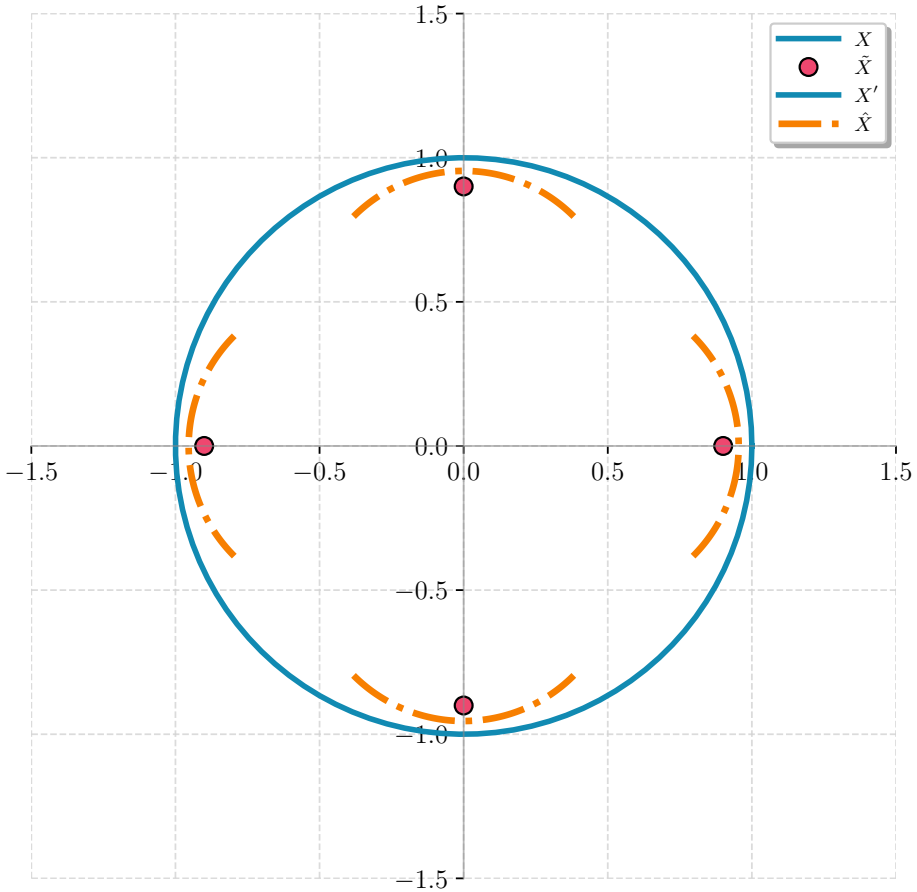
2-bit Quantization

Quantizing the Unit Circle: Perfect Perception



2-bit Quantization

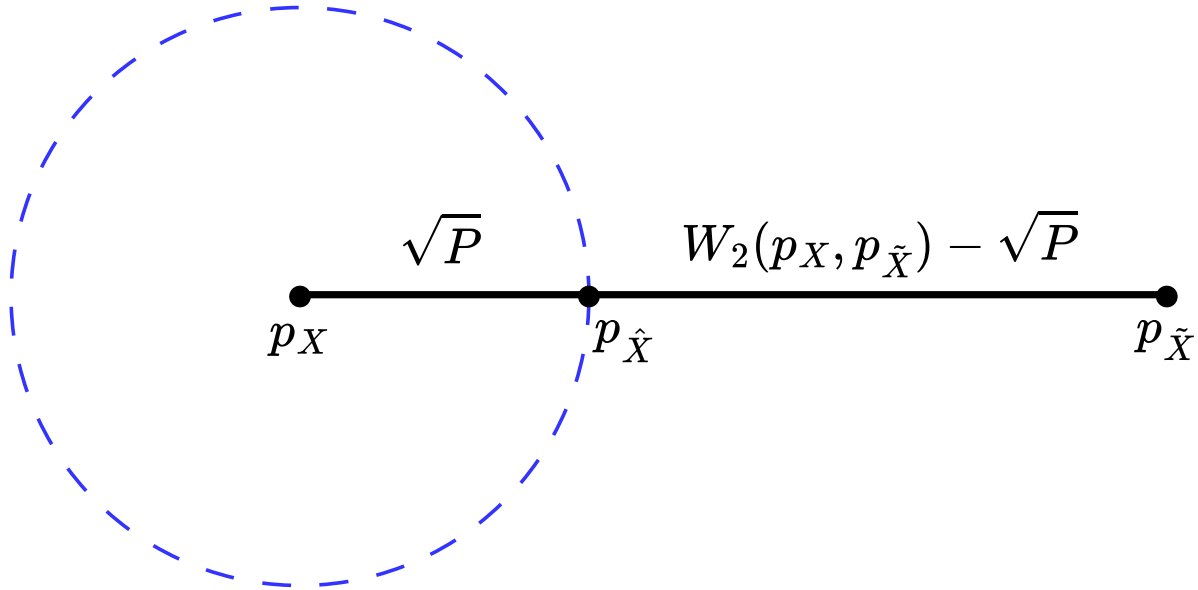
Quantizing the Unit Circle: Linear Interpolation



2-bit Quantization

Optimality of Linear Interpolation

- Zhang et al. 2021, Freirich et al. 2021



$$W_2(p_X, p_{\hat{X}}) \le \sqrt{P}$$

$$\begin{aligned}\mathbb{E}[\|X - \hat{X}\|^2] &= \mathbb{E}[\|X - \tilde{X}\|^2] + \mathbb{E}[\|\tilde{X} - \hat{X}\|^2] \\ &\ge \mathbb{E}[\|X - \tilde{X}\|^2] + W_2^2(p_{\tilde{X}}, p_{\hat{X}}) \\ &\ge \mathbb{E}[\|X - \tilde{X}\|^2] + [(W_2(P_X, p_{\tilde{X}}) - \sqrt{P})_+]^2\end{aligned}$$

Optimality of Linear Interpolation

Consider $\hat{X} = (1-\alpha)X' + \alpha\tilde{X}$ with $\alpha = \frac{\sqrt{P}}{W_2(p_X, p_{\tilde{X}})} \wedge 1$.

- Distortion loss

$$\begin{aligned}
\mathbb{E}[\|X - \hat{X}\|^2] &= \mathbb{E}[\|X - \tilde{X} + (1-\alpha)(\tilde{X} - X')\|^2] \\
&= \mathbb{E}[\|X - \tilde{X}\|^2] + (1-\alpha)^2 \mathbb{E}[\|X' - \tilde{X}\|^2] \\
&= \mathbb{E}[\|X - \tilde{X}\|^2] + (1-\alpha)^2 W_2^2(p_X, p_{\tilde{X}}) \\
&= \mathbb{E}[\|X - \tilde{X}\|^2] + [(W_2(P_X, p_{\tilde{X}}) - \sqrt{P})_+]^2
\end{aligned}$$

- Perception loss

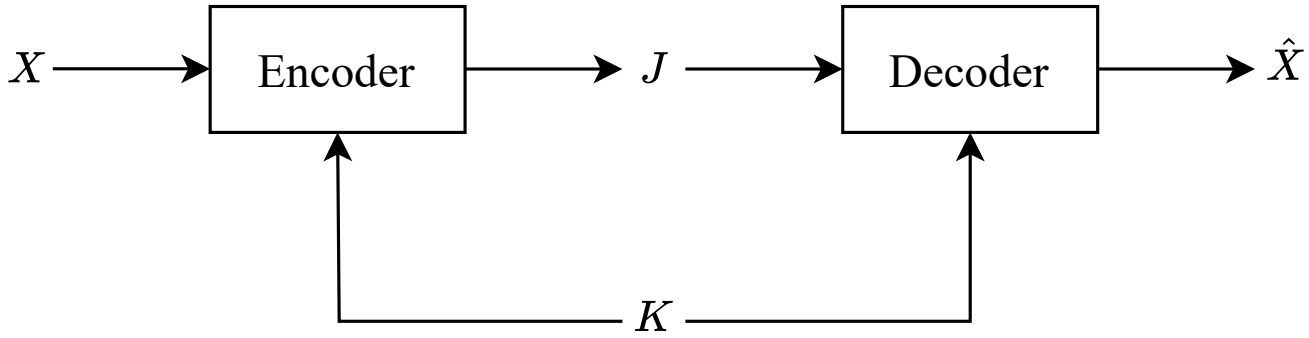
$$\begin{aligned}
W_2^2(p_X, p_{\hat{X}}) &\le \mathbb{E}[\|X' - \hat{X}\|^2] \\
&= \mathbb{E}[\|\alpha(X' - \tilde{X})\|^2] \\
&= \alpha^2 \mathbb{E}[\|X' - \tilde{X}\|^2] \\
&= \alpha^2 W_2^2(p_X, p_{\tilde{X}}) \\
&= P \wedge W_2^2(p_X, p_{\tilde{X}})
\end{aligned}$$

General Observation III

- Yan et al. 2022

$$D^{(1)}(R,P) = D^{(1)}(R,\infty) + [(\sqrt{D^{(1)}(R,\infty)} - \sqrt{P})_+]^2$$

Common Randomness



$$J \in \{1, 2, \ldots, M\} \quad K \in \{1, 2, \ldots, N\}$$

Assume $X \sim p_X$:

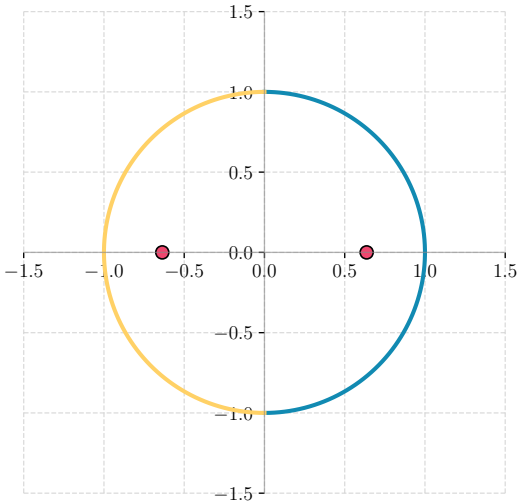
$$D^{(1)}(R, C, P) = \inf_{\mathrm{E,D}} \mathbb{E}[\|X - \hat{X}\|^2]$$

$$\text{s.t.} \quad \log M \leq R$$

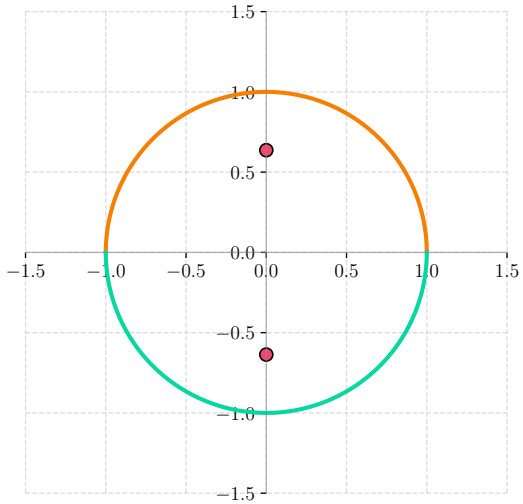
$$\log N \leq C$$

$$W_2^2(p_X, p_{\hat{X}}) \leq P$$

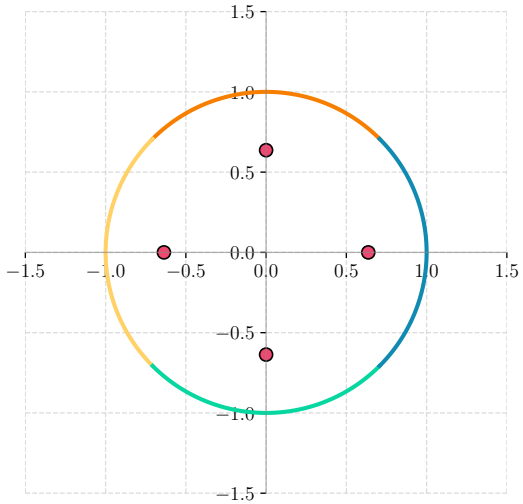
Quantizing the Unit Circle: No Perception Constraint



$\mathbb{E}[\|X - \tilde{X}\|^2 | K = 1]$



$\mathbb{E}[\|X - \tilde{X}\|^2 | K = 2]$

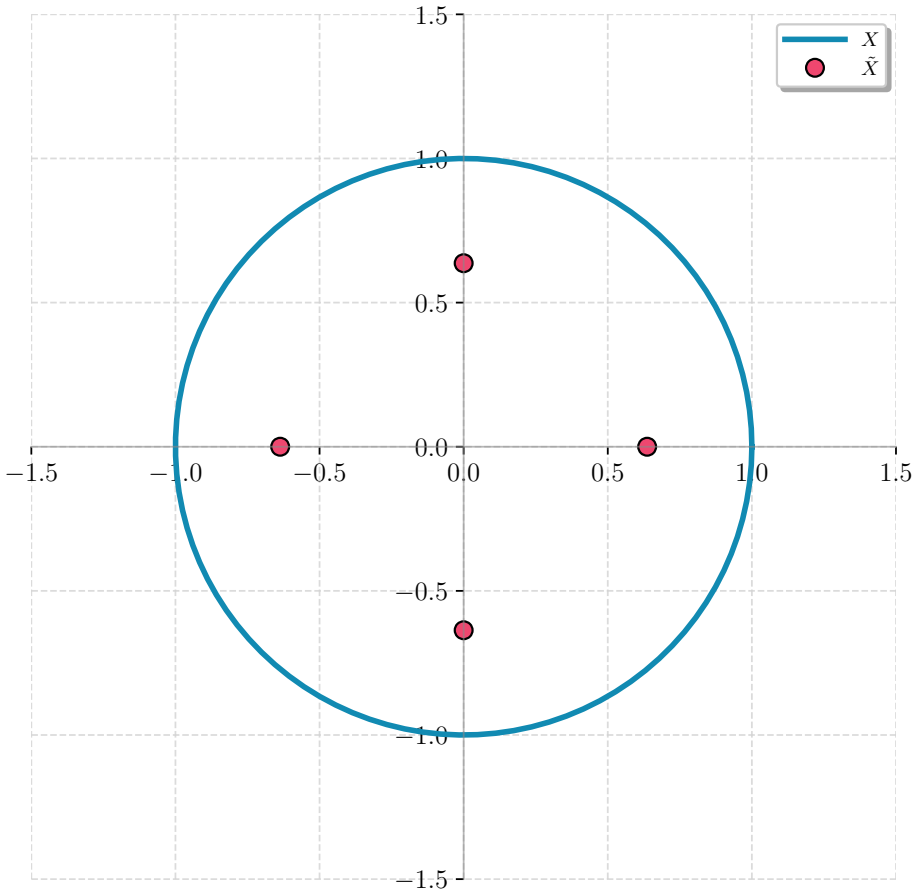


$W_2^2(p_X, p_{\tilde{X}})$

1-bit Quantization with 1-bit of Common Randomness

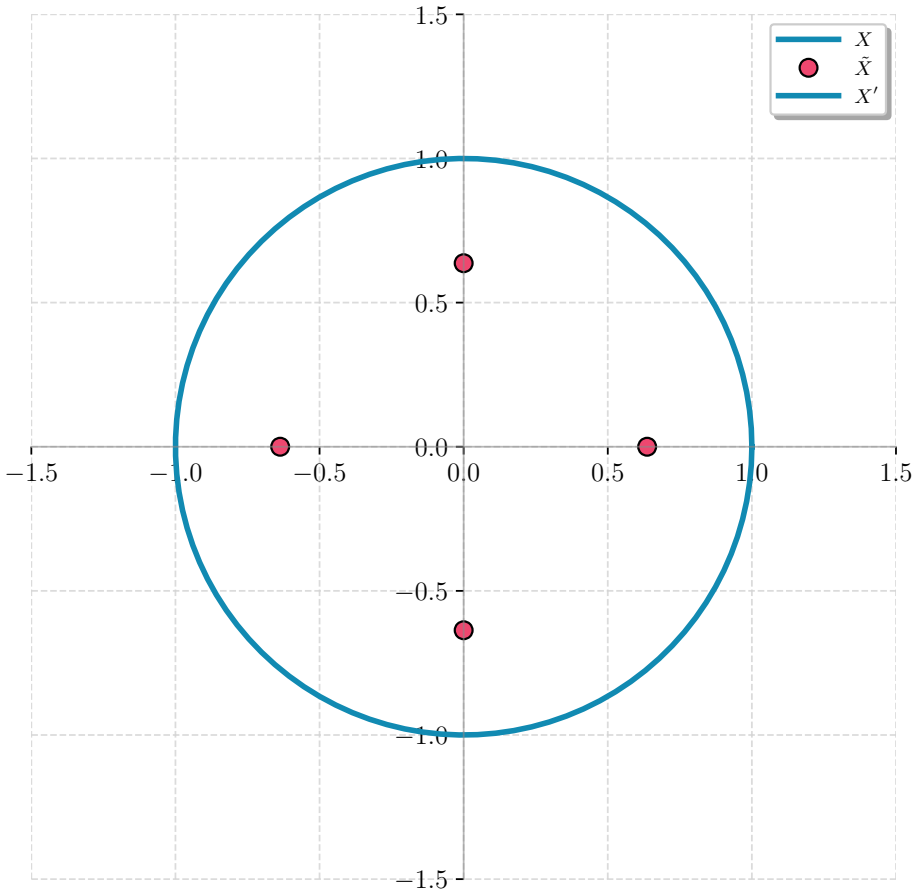
$$\mathbb{E}[\|X - \tilde{X}\|^2] = 1 - \frac{4}{\pi^2} \quad W_2^2(p_X, p_{\tilde{X}}) = 1 - \frac{4(2\sqrt{2} - 1)}{\pi^2}$$

Quantizing the Unit Circle: No Perception Constraint



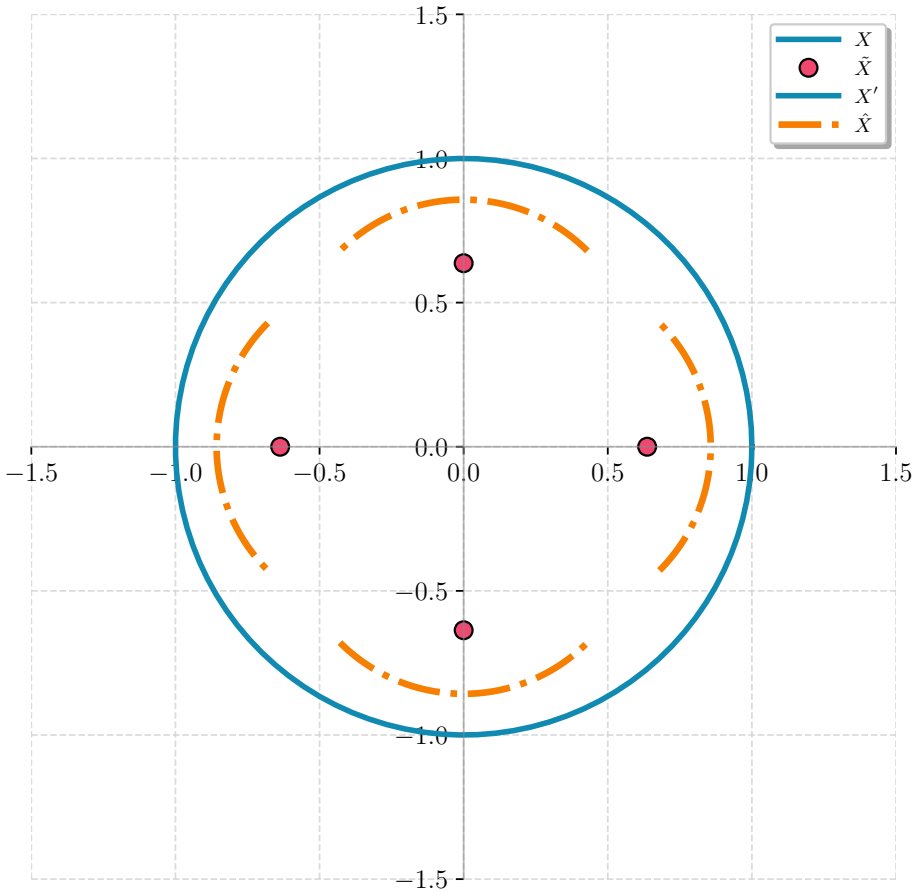
$$\mathbb{E}[\|X - \tilde{X}\|^2] > W_2^2(p_X, p_{\tilde{X}})$$

Quantizing the Unit Circle: Perfect Perception



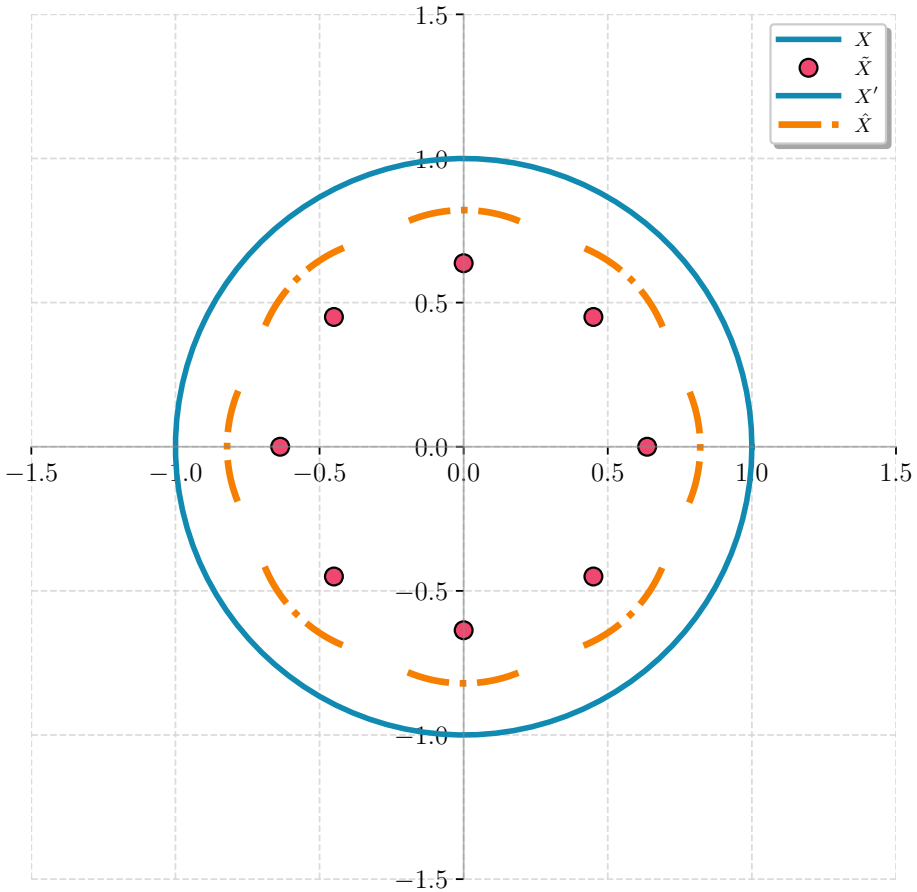
$$\mathbb{E}[\|X - X'\|^2] = \mathbb{E}[\|X - \tilde{X}\|^2] + W_2^2(p_X, p_{\tilde{X}}) < 2\mathbb{E}[\|X - \tilde{X}\|^2]$$

Quantizing the Unit Circle: Linear Interpolation



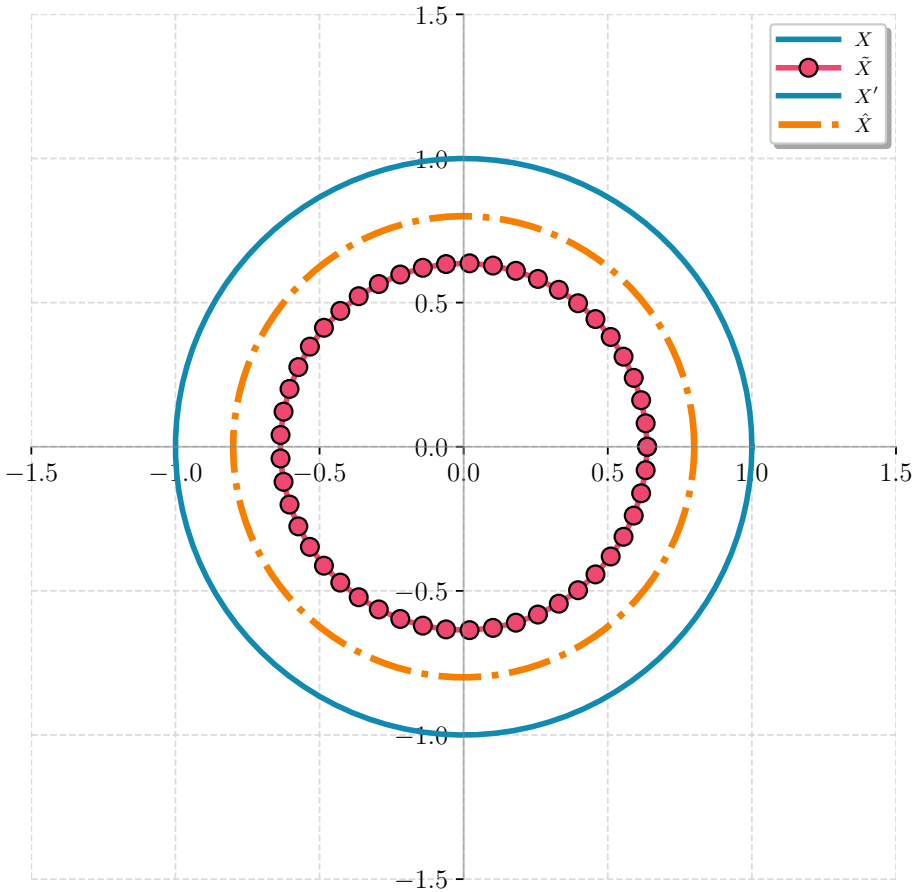
$$\mathbb{E}[\|X - \hat{X}\|^2] = \mathbb{E}[\|X - \tilde{X}\|^2] + [(W_2(p_X, p_{\tilde{X}}) - \sqrt{P})_+]^2$$

2-Bits of Common Randomness



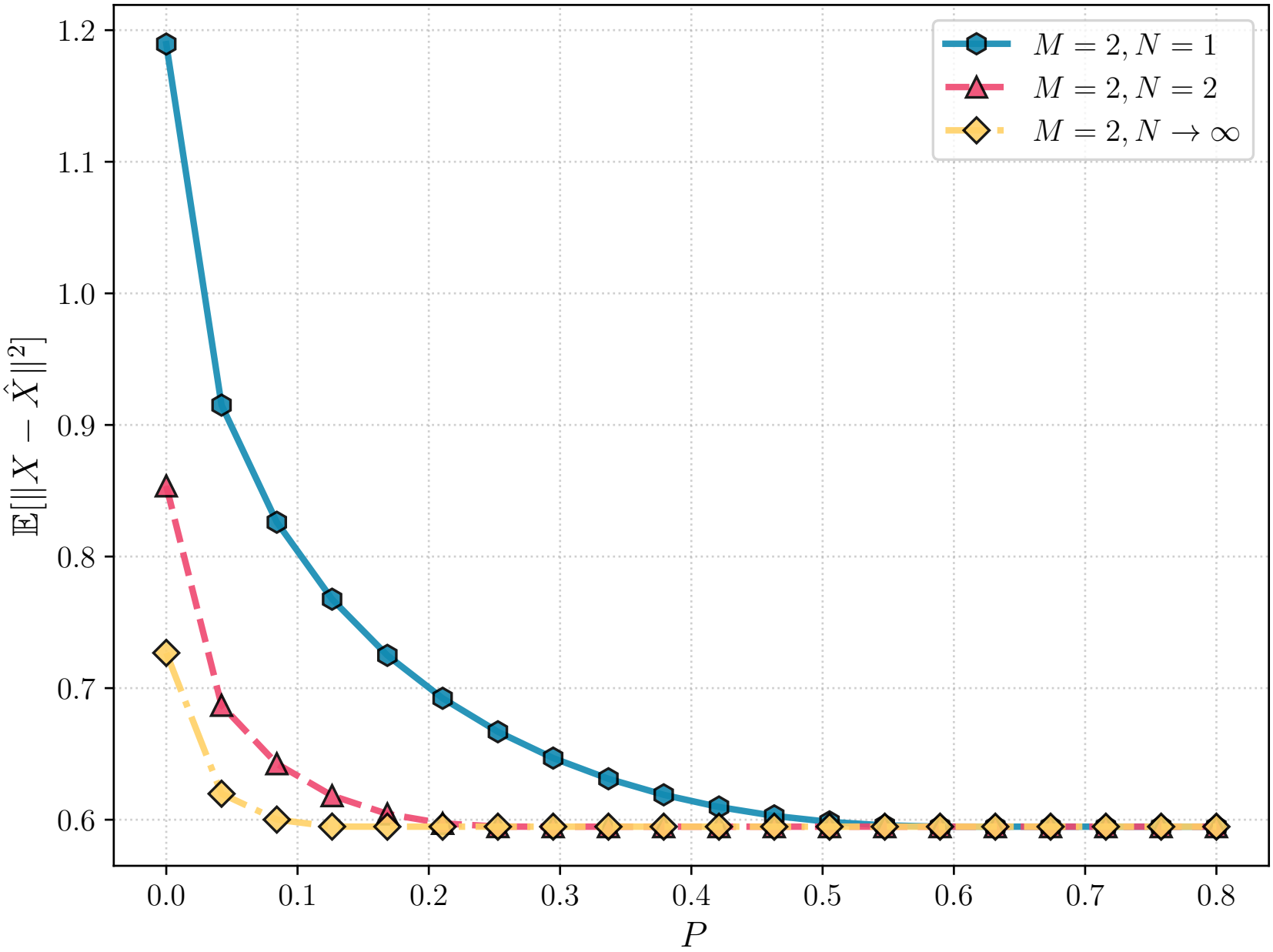
1-bit Quantization with 2-bits of Common Randomness

Unlimited Common Randomness

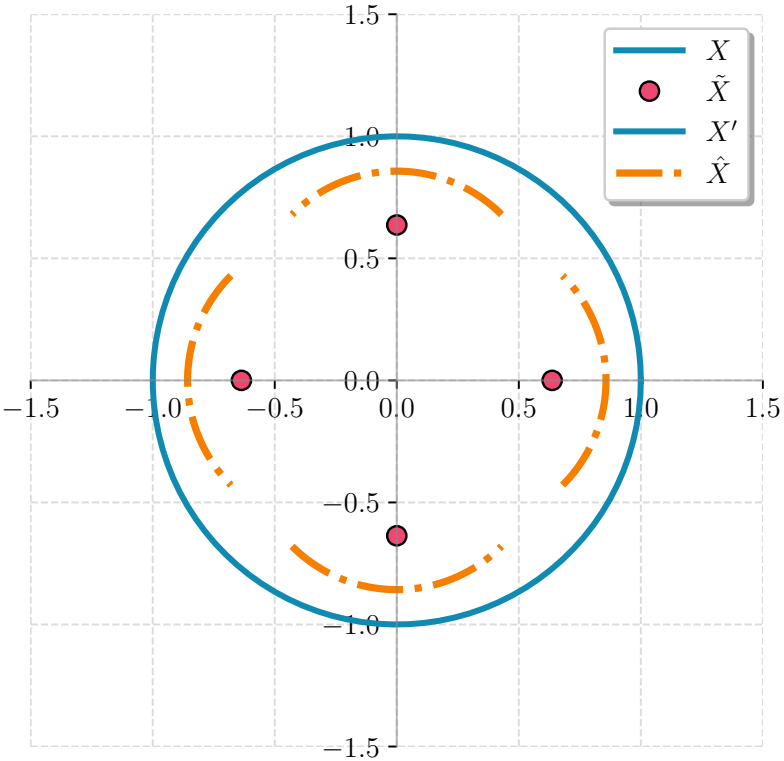


dithered quantization

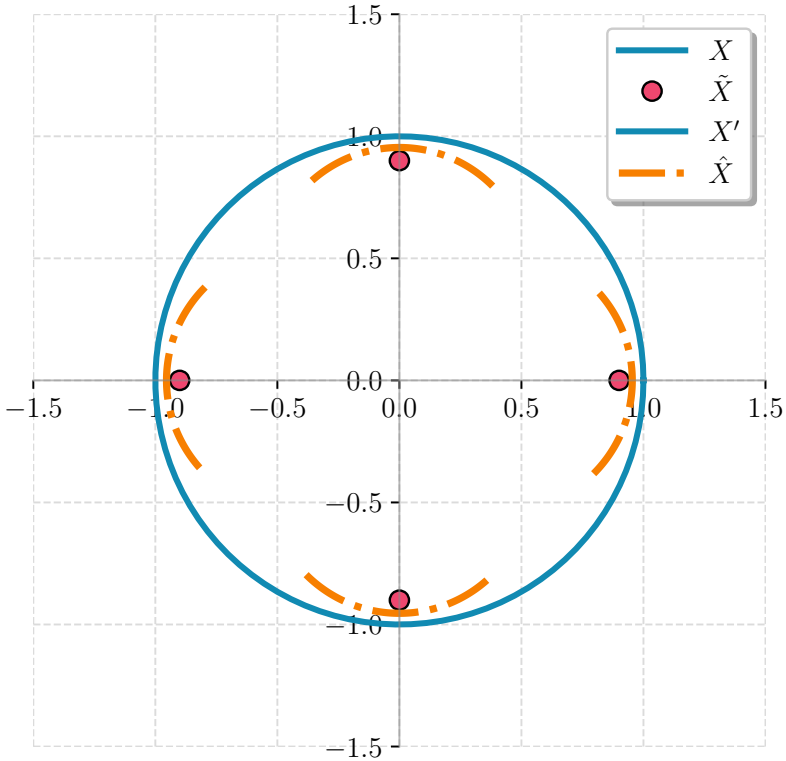
Distortion-Perception Tradeoff



Information Bits vs. Shared Bits

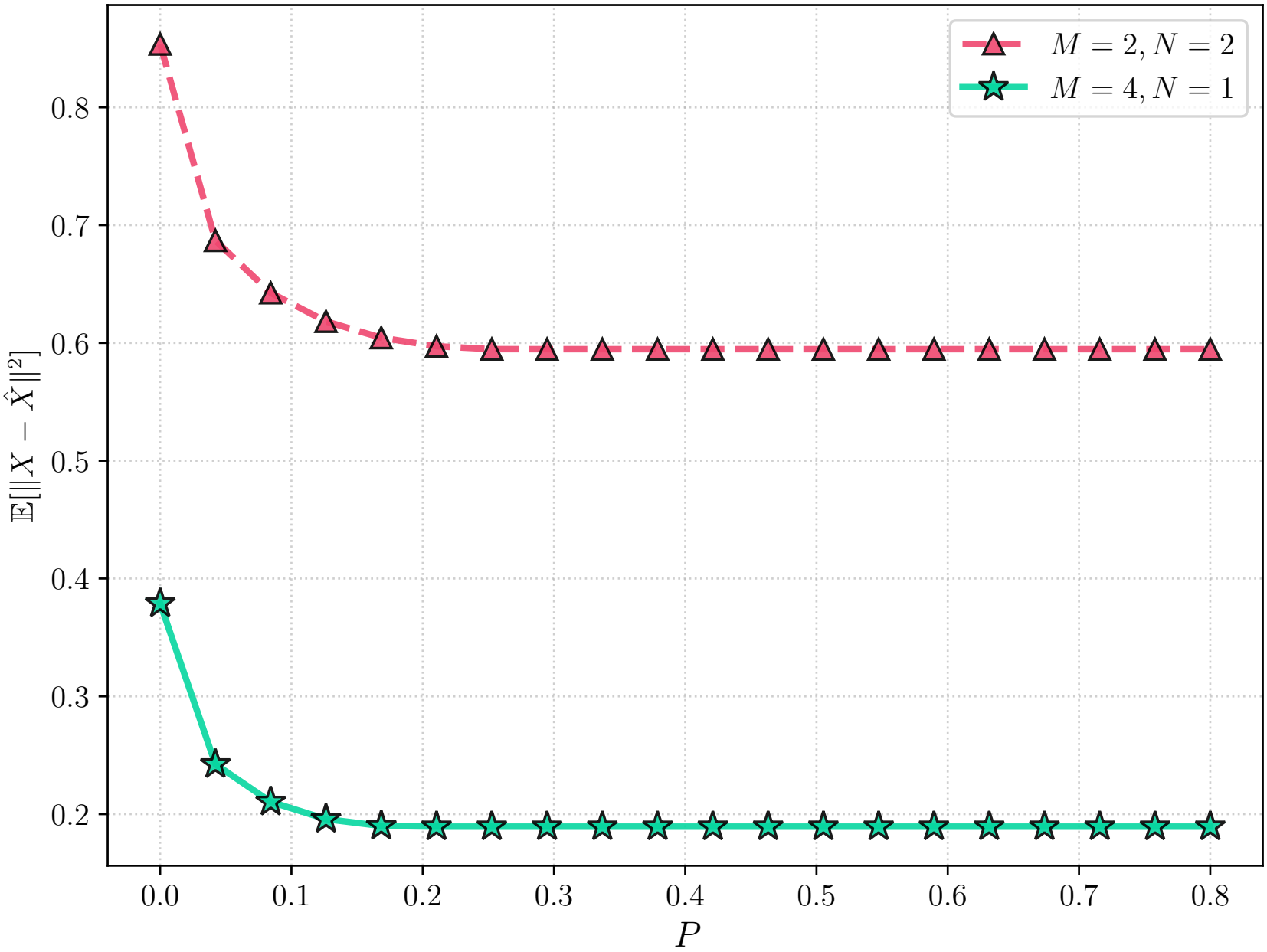


$M = 2, N = 2$

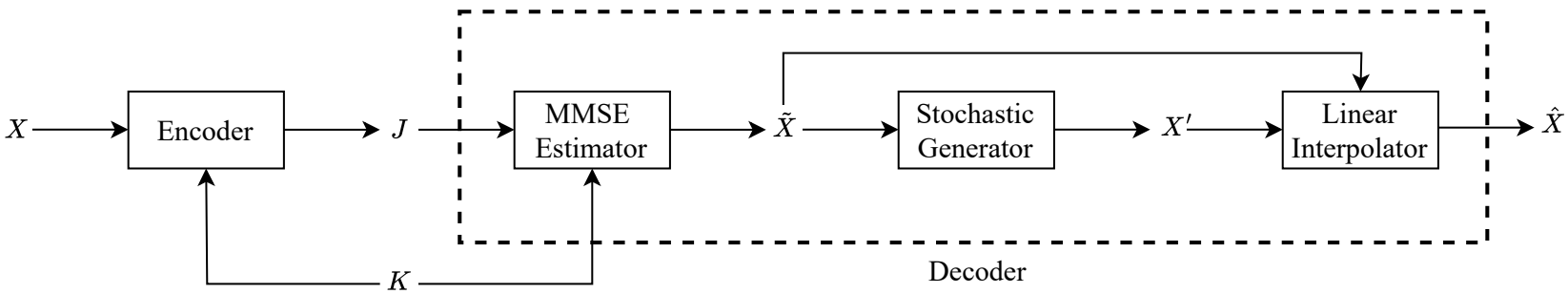


$M = 4, N = 1$

Information Bits vs. Shared Bits



Architectural Principles



$$J \in \{1, 2, \ldots, M\} \quad K \in \{1, 2, \ldots, N\}$$

- Freirich et al. 2021, Chen et al. 2026

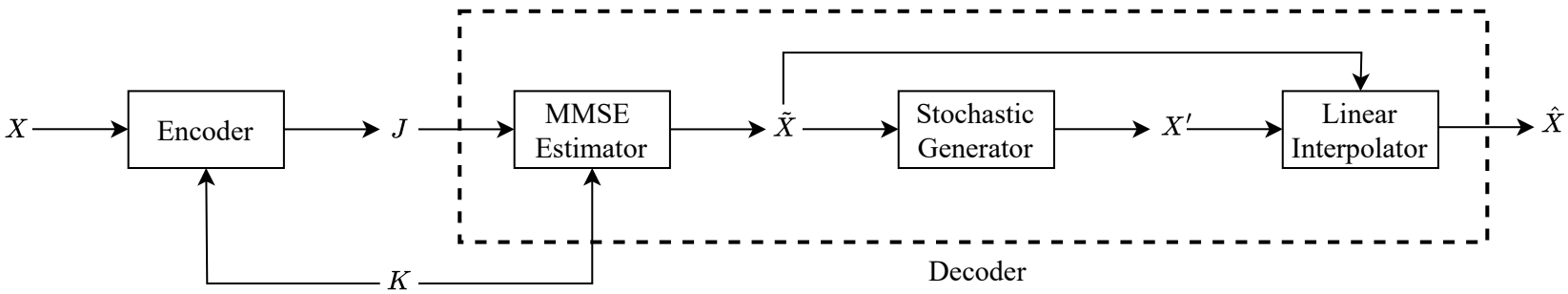
$$D^{(1)}(R, C, P) = \inf_{\mathrm{E}} \mathbb{E}[\|X - \tilde{X}\|^2] + [(W_2(p_X, p_{\tilde{X}}) - \sqrt{P})_+]^2$$

$$\text{s.t.} \quad \log M \leq R$$

$$\log N \leq C$$

$$\tilde{X} = \mathbb{E}[X|J, K]$$

MMSE Estimator

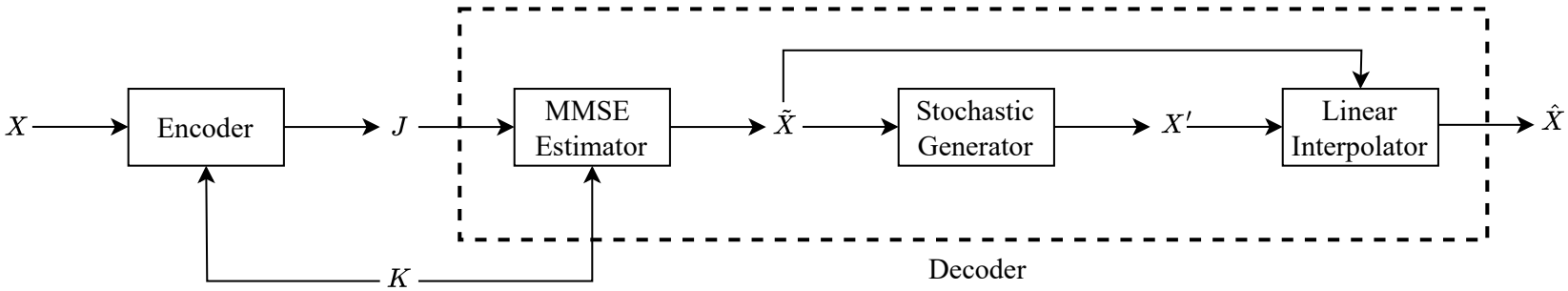


$$J \in \{1, 2, \ldots, M\} \quad K \in \{1, 2, \ldots, N\}$$

- MMSE estimator: $\tilde{X} = \mathbb{E}[X|J, K]$

supervised learning

Stochastic Generator



$$J \in \{1, 2, \ldots, M\} \quad K \in \{1, 2, \ldots, N\}$$

- Stochastic generator: Simulate the optimal transport plan $p_{X'|\tilde{X}}$ that achieves $W_2^2(p_X, p_{\tilde{X}})$

Stochastic Generator

- Without common randomness

$$\mathbb{E}[\|X - \tilde{X}\|^2] = W_2^2(p_X, p_{\tilde{X}})$$
$$\Longrightarrow p_{X'|X} = p_{X|\tilde{X}}$$

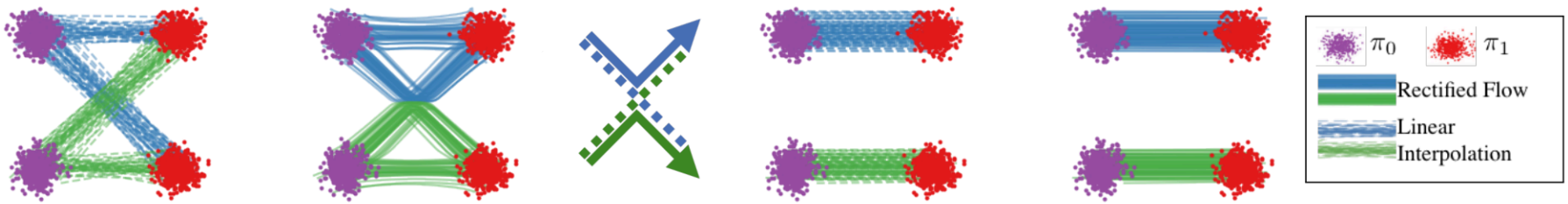
  - Posterior sampling

Stochastic Generator

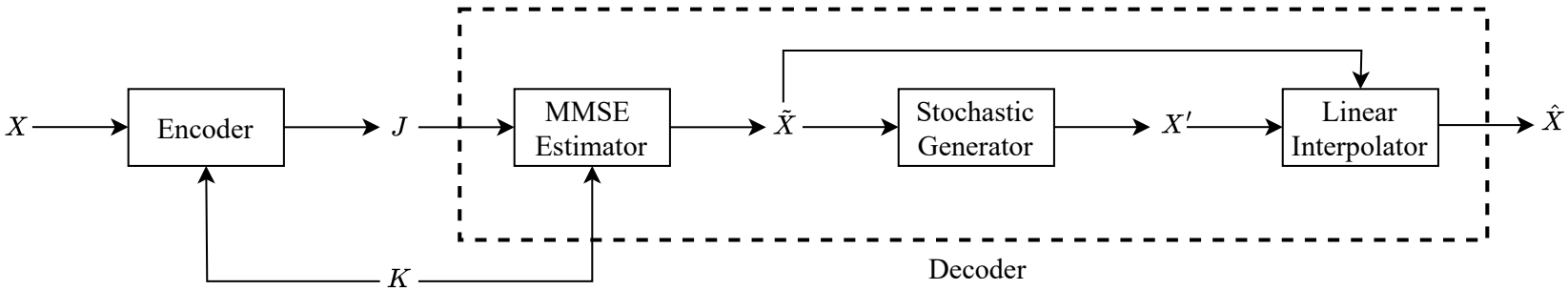
- With common randomness

$$\mathbb{E}[\|X - \tilde{X}\|^2] > W_2^2(p_X, p_{\tilde{X}})$$
$$\Longrightarrow p_{X'|X} \neq p_{X|\tilde{X}}$$

  - Rectified flow



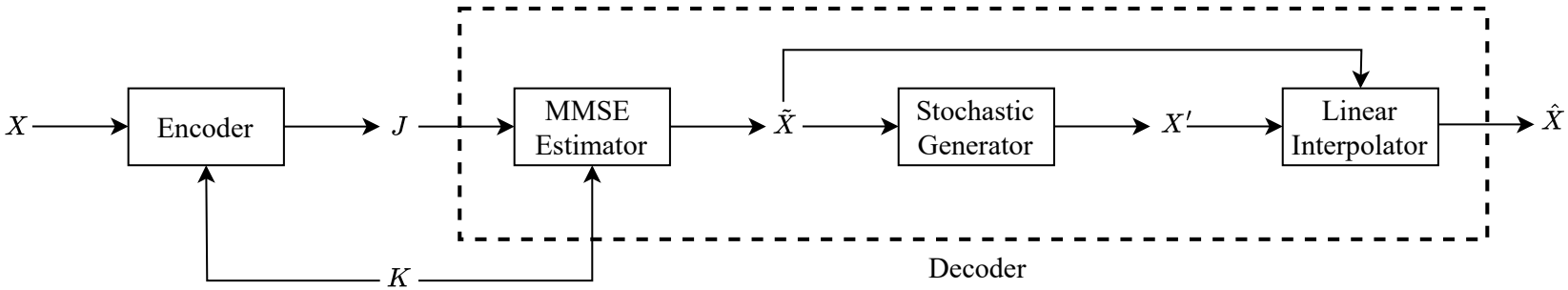
Linear Interpolator



$$J \in \{1, 2, \ldots, M\} \quad K \in \{1, 2, \ldots, N\}$$

- Linear interpolator: $\hat{X} = (1-\alpha)X' + \alpha\tilde{X}$ with $\alpha = \frac{\sqrt{P}}{W_2(p_X, p_{\tilde{X}})} \wedge 1$

Encoder

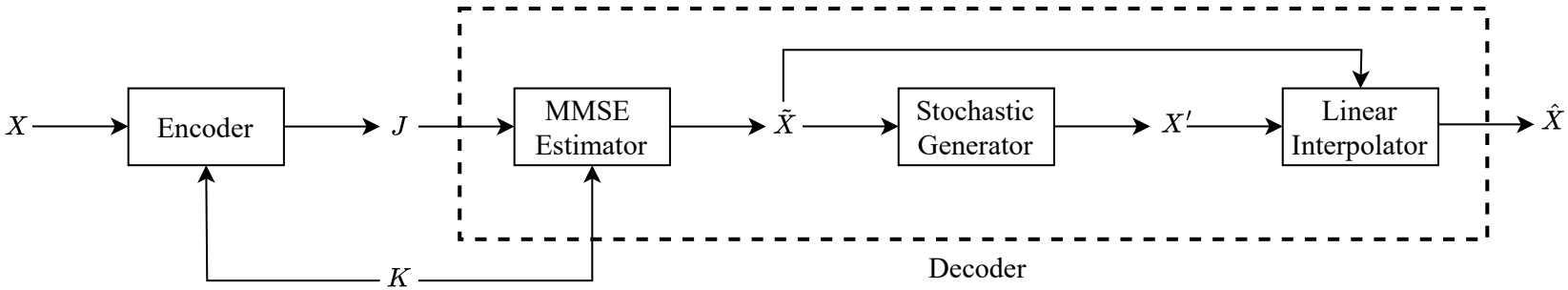


$$J \in \{1, 2, \ldots, M\} \quad K \in \{1, 2, \ldots, N\}$$

- Encoder: $(X, K) \to J$

  nonlinear transformation $+$ quantization

Universal Representation



$$J \in \{1, 2, \ldots, M\} \quad K \in \{1, 2, \ldots, N\}$$

A representation is said to be universal if the choice of the optimal encoder does not depend on the specific perception constraint P.

Universal Representation

$$D^{(1)}(R, C, P) = \inf_{\mathrm{E}} \mathbb{E}[\|X - \tilde{X}\|^2] + [(W_2(p_X, p_{\tilde{X}}) - \sqrt{P})_+]^2$$

$$\text{s.t.} \quad \log M \le R$$

$$\log N \le C$$

$$\tilde{X} = \mathbb{E}[X|J, K]$$

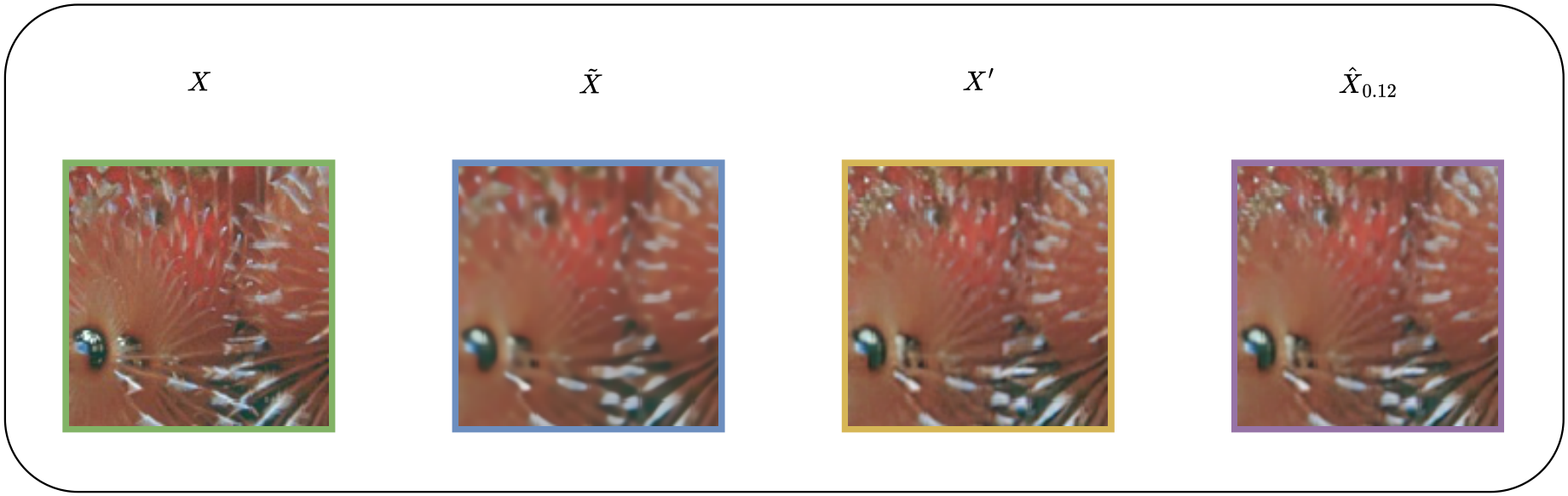
Universal representation exists in the absence of common randomness ($C = 0$):

$$D^{(1)}(R, 0, P) = \inf_{\mathrm{E}} \mathbb{E}[\|X - \tilde{X}\|^2] + [(\sqrt{\mathbb{E}[\|X - \tilde{X}\|^2]} - \sqrt{P})_+]^2$$

$$\text{s.t.} \quad \log M \le R$$

$$\tilde{X} = \mathbb{E}[X|J]$$

Experimental Results

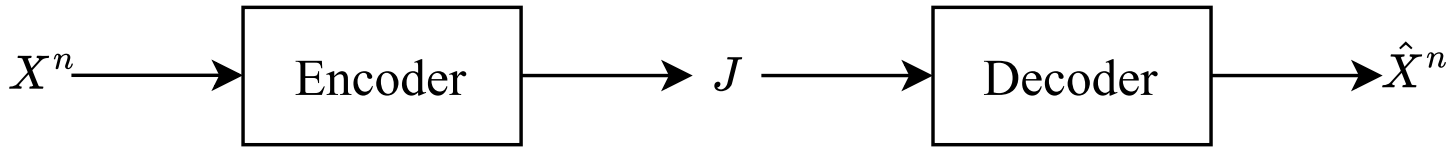


Experimental Results

X $\tilde{X}$ X' $\hat{X}_{0.3}$




Asymptotic Setting



$$J \in \{1, 2, \ldots, M\}$$

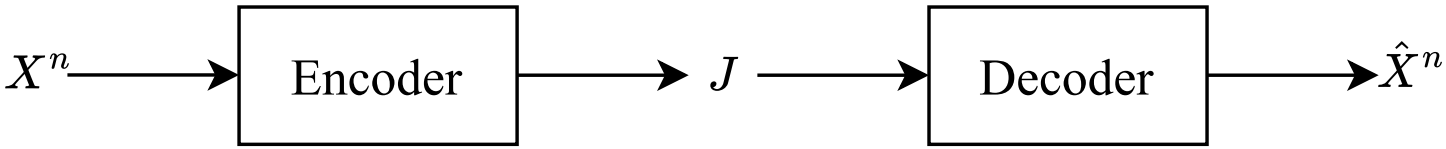
Assume $X^n \sim p_X^n$:

$$D^{(\infty)}(R, P) = \lim_{n\to\infty} \inf_{\mathrm{E,D}} \frac{1}{n} \sum_{t=1}^{n} \mathbb{E}[\|X_t - \hat{X}_t\|^2]$$

$$\text{s.t.} \quad \frac{1}{n} \log M \le R$$

$$\frac{1}{n} \sum_{t=1}^{n} W_2^2(p_X, p_{\hat{X}_t}) \le P$$

Asymptotic Setting



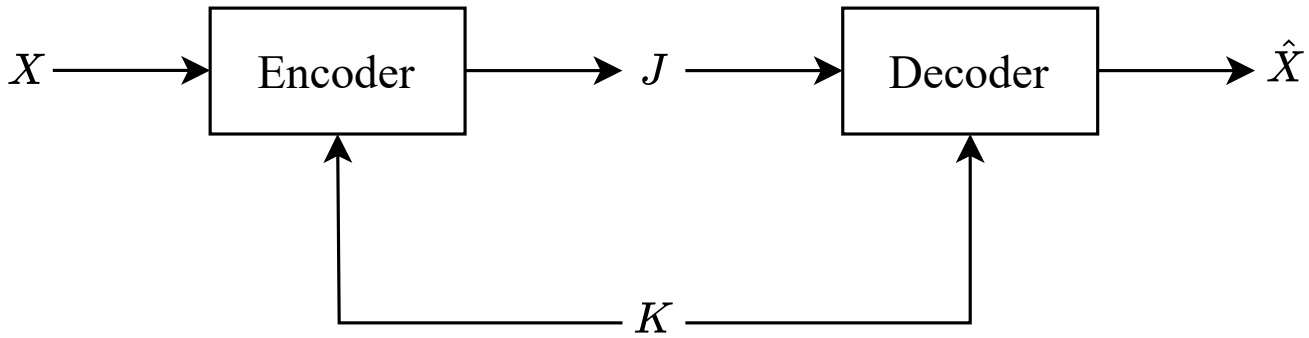
$$J \in \{1, 2, \ldots, M\}$$

- Blau and Michaeli 2019, Theis and Wagner 2021, Chen et al. 2022

$$D^{(\infty)}(R, P) = \inf_{p_{\hat{X}|X}} \mathbb{E}[\|X - \hat{X}\|^2]$$
$$\text{s.t.} \quad I(X; \hat{X}) \le R$$
$$W_2^2(p_X, p_{\hat{X}}) \le P$$

deterministic encoding and decoding suffices!

One-Shot Setting



$$J \in \{1, 2, \ldots, M\} \quad K \in \{1, 2, \ldots, N\}$$

Assume $X \sim p_X$:

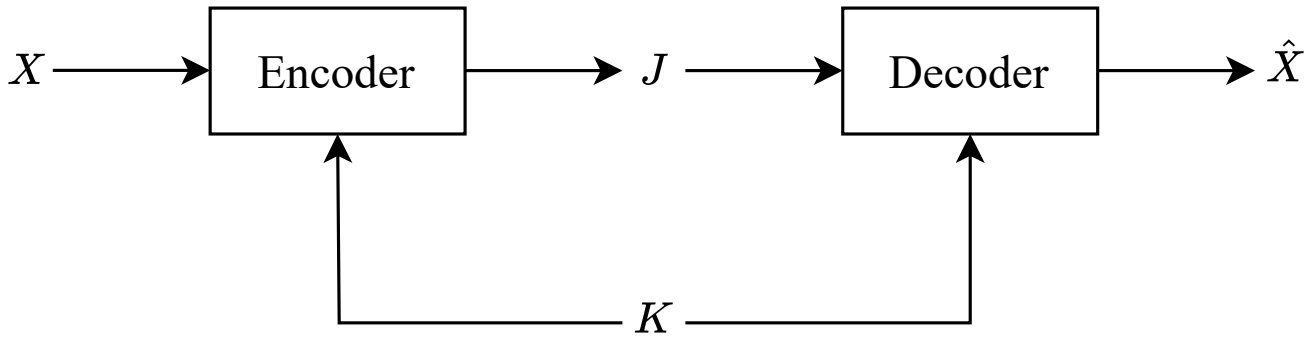
$$D^{(1)}(R, C, P) = \inf_{\mathrm{E,D}} \mathbb{E}[\|X - \hat{X}\|^2]$$

$$\text{s.t.} \quad \log M \leq R$$

$$\log N \leq C$$

$$W_2^2(p_X, p_{\hat{X}}) \leq P$$

One-Shot Setting



$$J \in \{1, 2, \ldots, M\} \quad K \in \{1, 2, \ldots, N\}$$

- Freirich et al. 2021, Chen et al. 2026

$$D^{(1)}(R, C, P) = \inf_{\mathrm{E}} \mathbb{E}[\|X - \tilde{X}\|^2] + [(W_2(p_X, p_{\tilde{X}}) - \sqrt{P})_+]^2$$

$$\text{s.t.} \quad \log M \leq R$$

$$\log N \leq C$$

$$\tilde{X} = \mathbb{E}[X|J, K]$$

(common) randomness is essential!

Random vs. Deterministic

- Asymptotic setting ($R = 1$, $P = 0$)

$$D^{(\infty)}(1,0) = \lim_{n\to\infty} \inf_{\mathrm{E,D}} \frac{1}{n} \sum_{t=1}^{n} \mathbb{E}[\|X_t - \hat{X}_t\|^2]$$

$$\text{s.t.} \quad \frac{1}{n} \log M \le 1$$

$$\frac{1}{n} \sum_{t=1}^{n} W_2^2(p_X, p_{\hat{X}_t}) \le 0$$

- One-shot setting ($R = 1$, $P = 0$)

$$D^{(1)}(1, C, 0) = \inf_{\mathrm{E,D}} \mathbb{E}[\|X - \hat{X}\|^2]$$

$$\text{s.t.} \quad \log M \le 1$$

$$\log N \le C$$

$$W_2^2(p_X, p_{\hat{X}}) \le 0$$

Symbol-Level vs. Sequence-Level

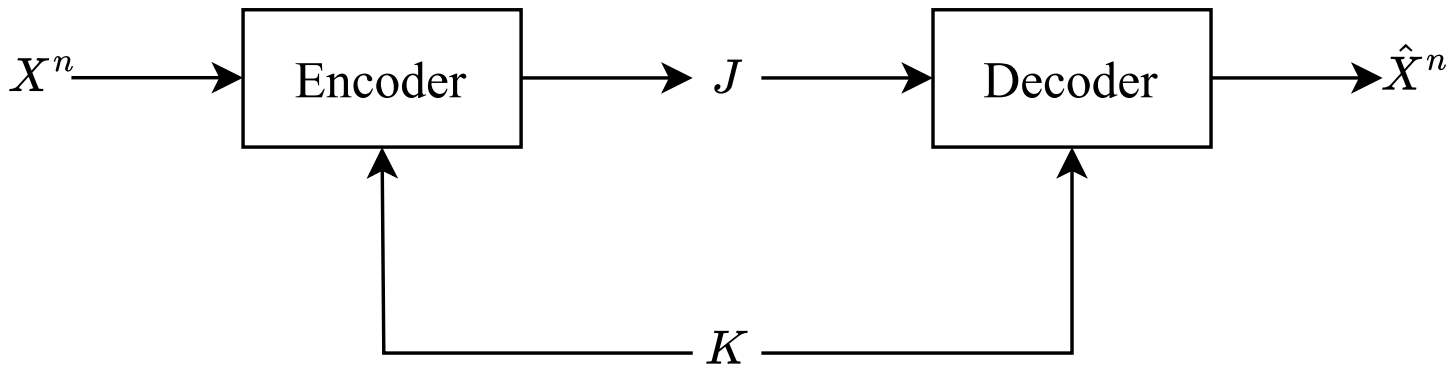
- Symbol-level perception constraint

$$\frac{1}{n}\sum_{t=1}^{n} W_2^2(p_X, p_{\hat{X}_t}) \le P$$

- Sequence-level perception constraint

$$\frac{1}{n} W_2^2(p_{X^n}, p_{\hat{X}^n}) \le P$$

An Alternative Formulation



$$J \in \{1, 2, \ldots, M\} \quad K \in \{1, 2, \ldots, N\}$$

Assume $X^n \sim p_X^n$:

$$D^{(\infty)}(R, C, P) = \lim_{n\to\infty} \inf_{\mathrm{E,D}} \frac{1}{n} \sum_{t=1}^{n} \mathbb{E}[\|X_t - \hat{X}_t\|^2]$$

$$\text{s.t.} \quad \frac{1}{n} \log M \le R$$

$$\frac{1}{n} \log N \le C$$

$$\frac{1}{n} W_2^2(p_{X^n}, p_{\hat{X}^n}) \le P$$

Fundamental Limits

- Asymptotic setting

$$D^{(\infty)}(R,C,P) = \inf_{\substack{p_{\tilde{X}} \\ \mu,\nu\in\Pi(p_X,p_{\tilde{X}})}} \mathbb{E}_\mu[\|X-\tilde{X}\|^2] + [\sqrt{\mathbb{E}_\nu[\|X-\tilde{X}\|^2]} - \sqrt{P})_+]^2$$

$$\text{s.t.} \quad I_\mu(X;\tilde{X}) \le R$$

$$I_\nu(X;\tilde{X}) \le R + C$$

$$\tilde{X} = \mathbb{E}_\mu[X|\tilde{X}]$$

- One-shot setting

$$D^{(1)}(R,C,P) = \inf_{\mathrm{E}} \mathbb{E}[\|X-\tilde{X}\|^2] + [(W_2(p_X,p_{\tilde{X}}) - \sqrt{P})_+]^2$$

$$\text{s.t.} \quad \log M \le R$$

$$\log N \le C$$

$$\tilde{X} = \mathbb{E}[X|J,K]$$

Fundamental Limits

- Asymptotic setting

$$D^{(\infty)}(R,C,P) = \inf_{\substack{p_{\tilde{X}} \\ \mu,\nu\in\Pi(p_X,p_{\tilde{X}})}} \mathbb{E}_\mu[\|X-\tilde{X}\|^2] + [\sqrt{\mathbb{E}_\nu[\|X-\tilde{X}\|^2]} - \sqrt{P})_+]^2$$

$$\text{s.t.} \quad I_\mu(X;\tilde{X}) \le R$$

$$I_\nu(X;\tilde{X}) \le R + C$$

$$\tilde{X} = \mathbb{E}_\mu[X|\tilde{X}]$$

- One-shot setting

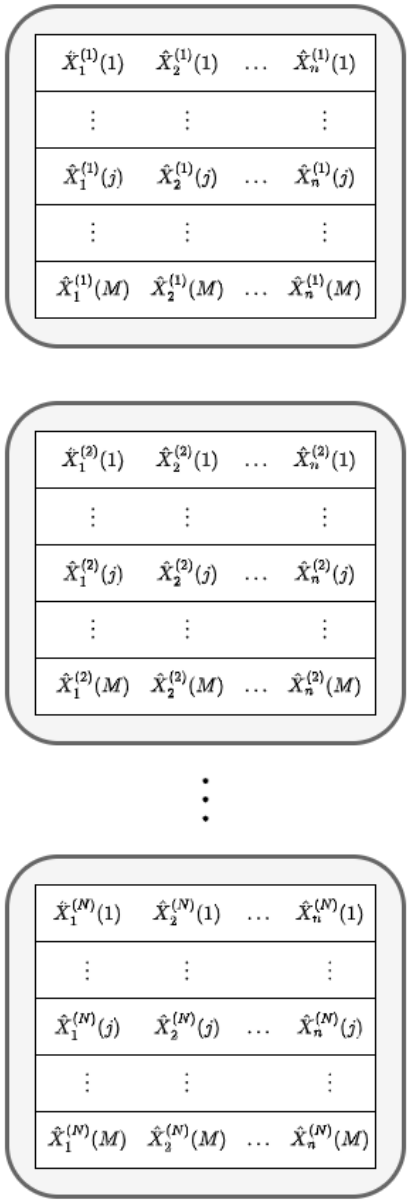
$$D^{(1)}(R,C,P) = \inf_{\mathrm{E}} \mathbb{E}[\|X-\tilde{X}\|^2] + [(W_2(p_X,p_{\tilde{X}}) - \sqrt{P})_+]^2$$

$$\text{s.t.} \quad \log M \le R$$

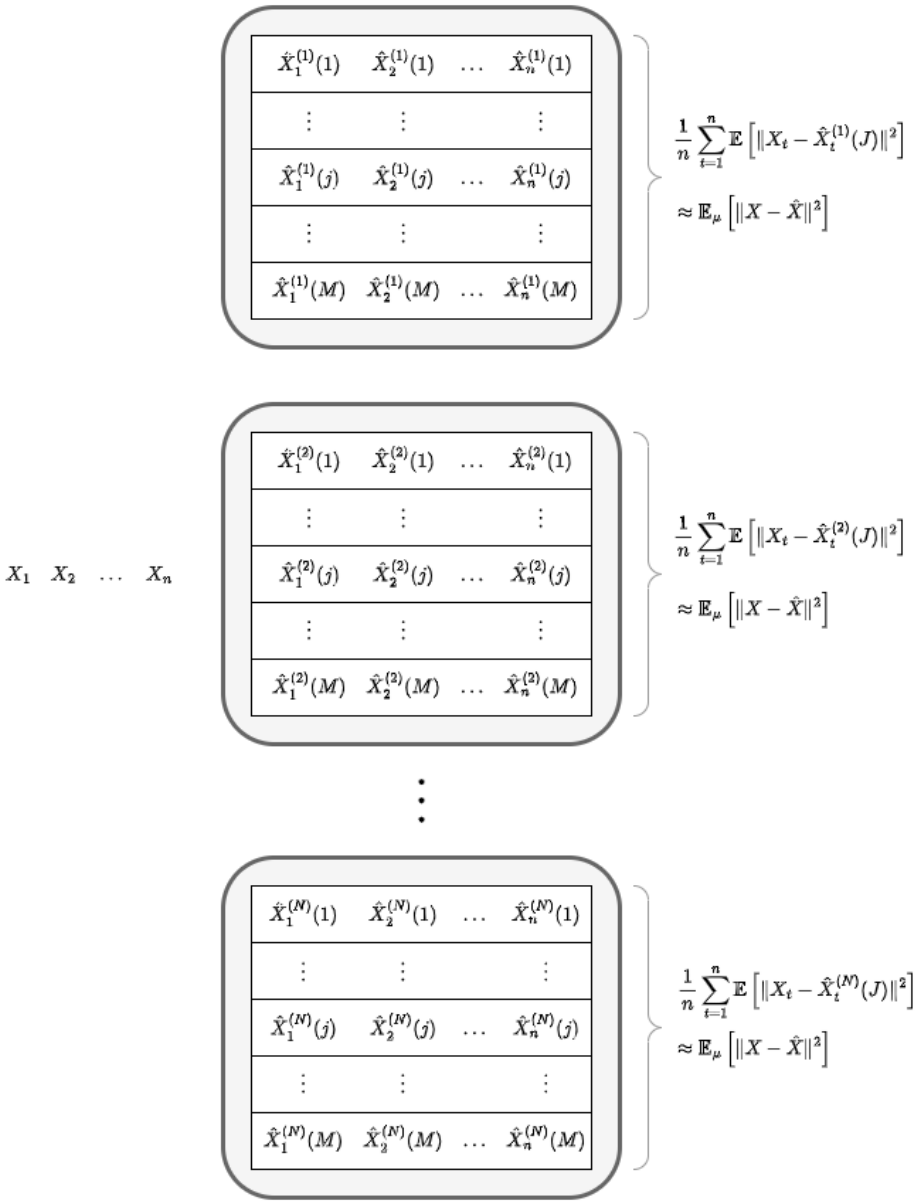
$$\log N \le C$$

$$\tilde{X} = \mathbb{E}[X|J,K]$$

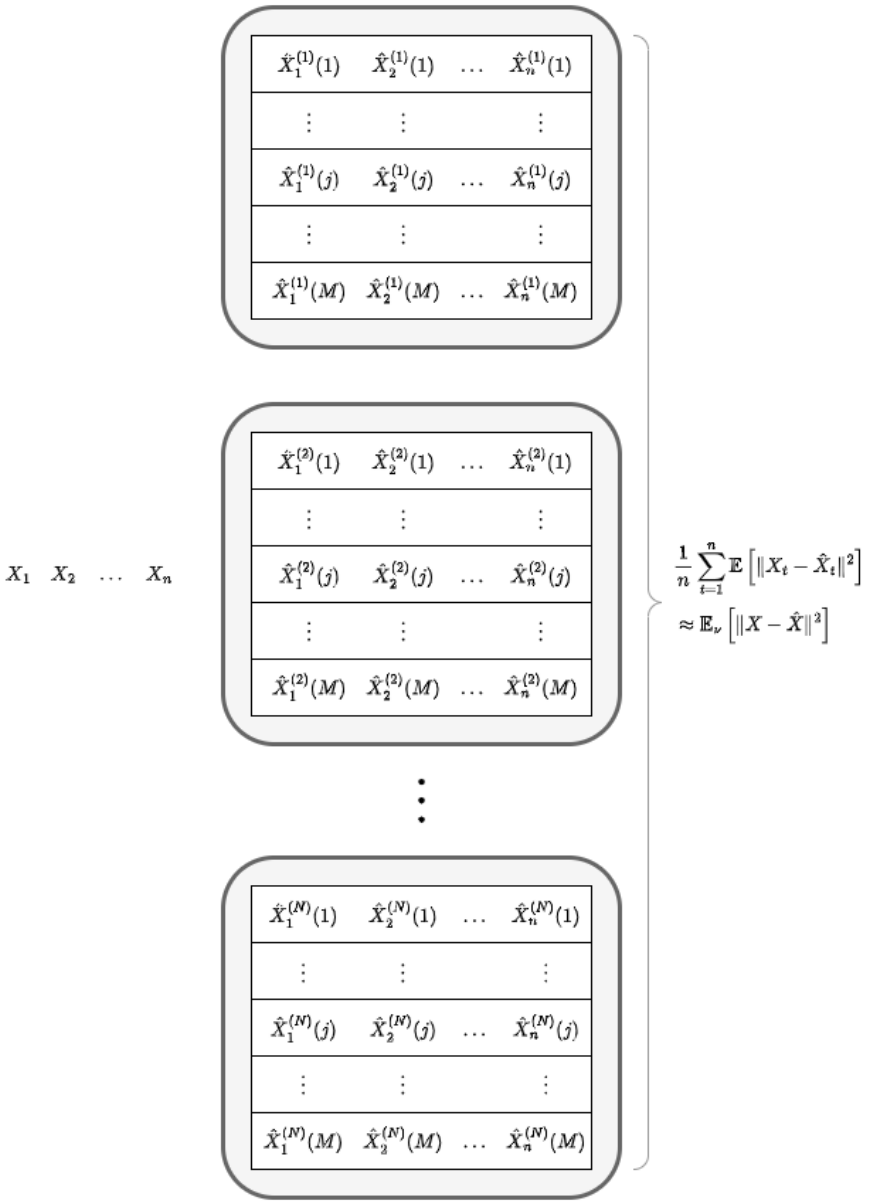
Operational Encoding vs. Virtual Encoding



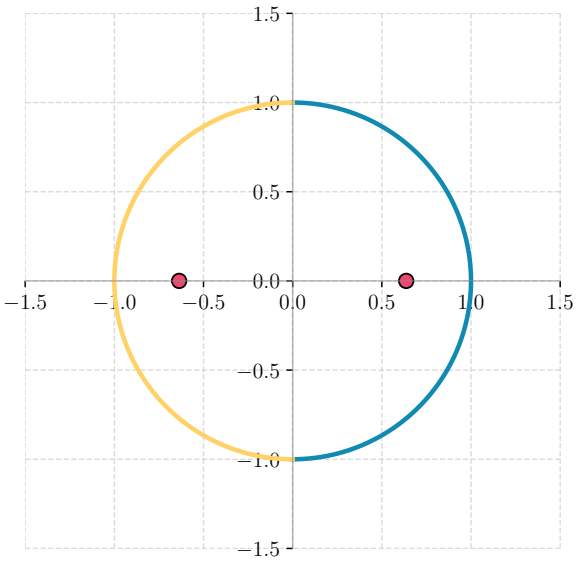
Operational Encoding vs. Virtual Encoding



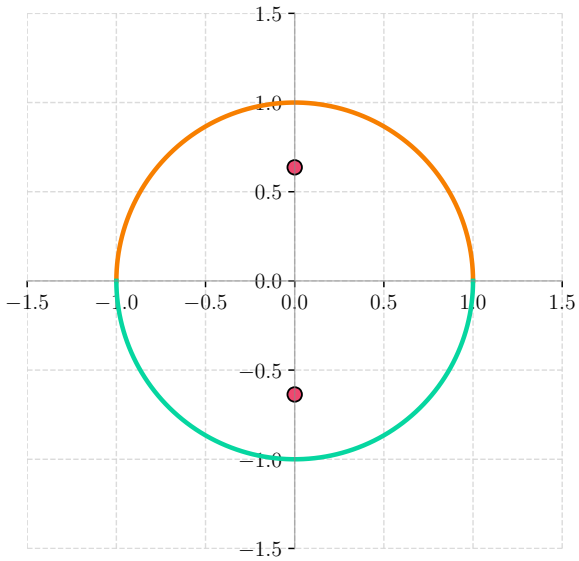
Operational Encoding vs. Virtual Encoding



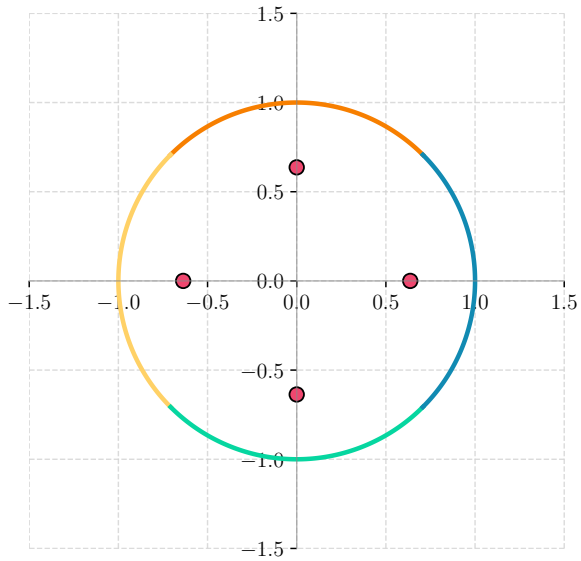
Operational Encoding vs. Virtual Encoding



$\mathbb{E}[\|X - \tilde{X}\|^2 | K = 1]$

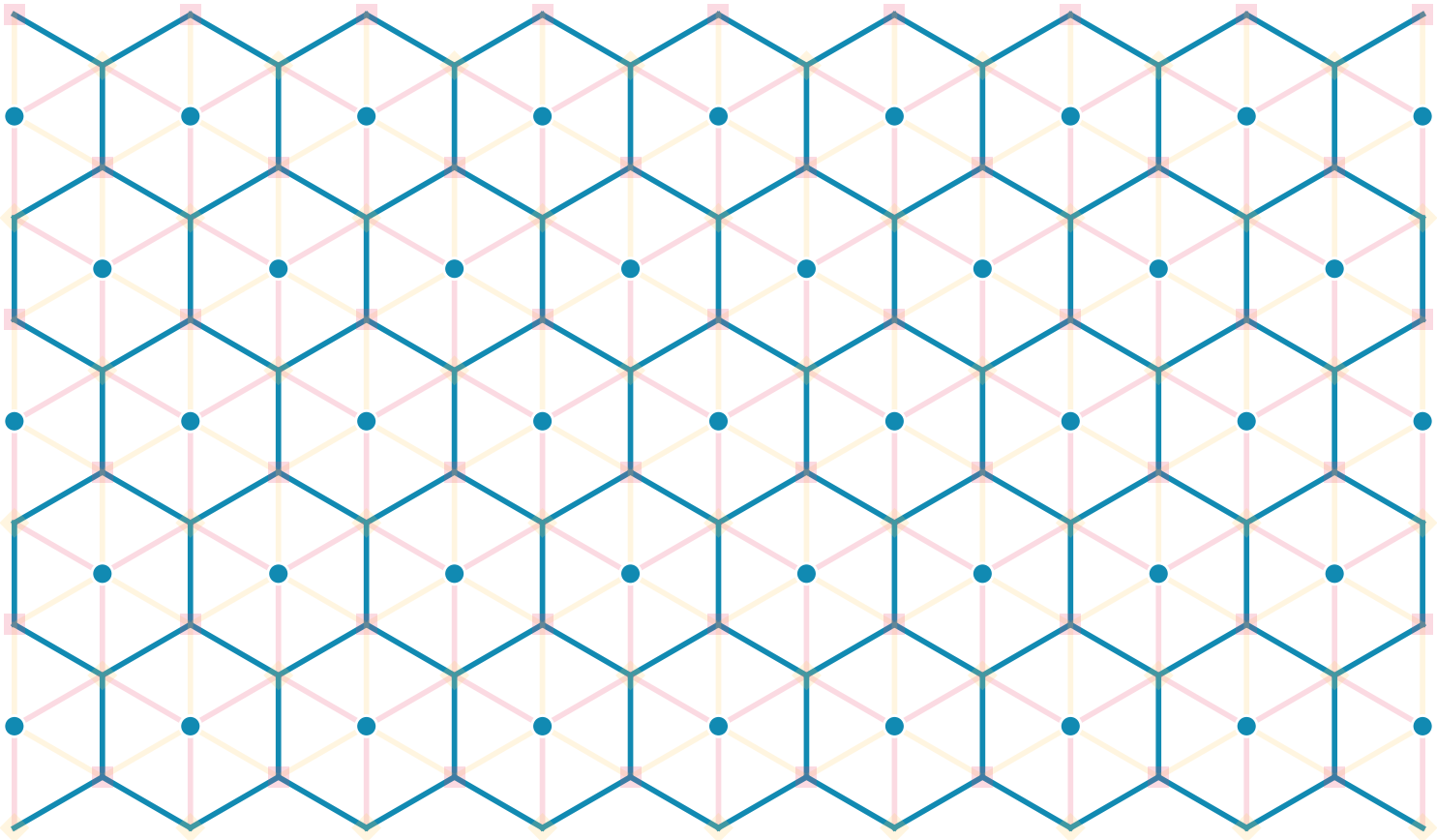


$\mathbb{E}[\|X - \tilde{X}\|^2 | K = 2]$

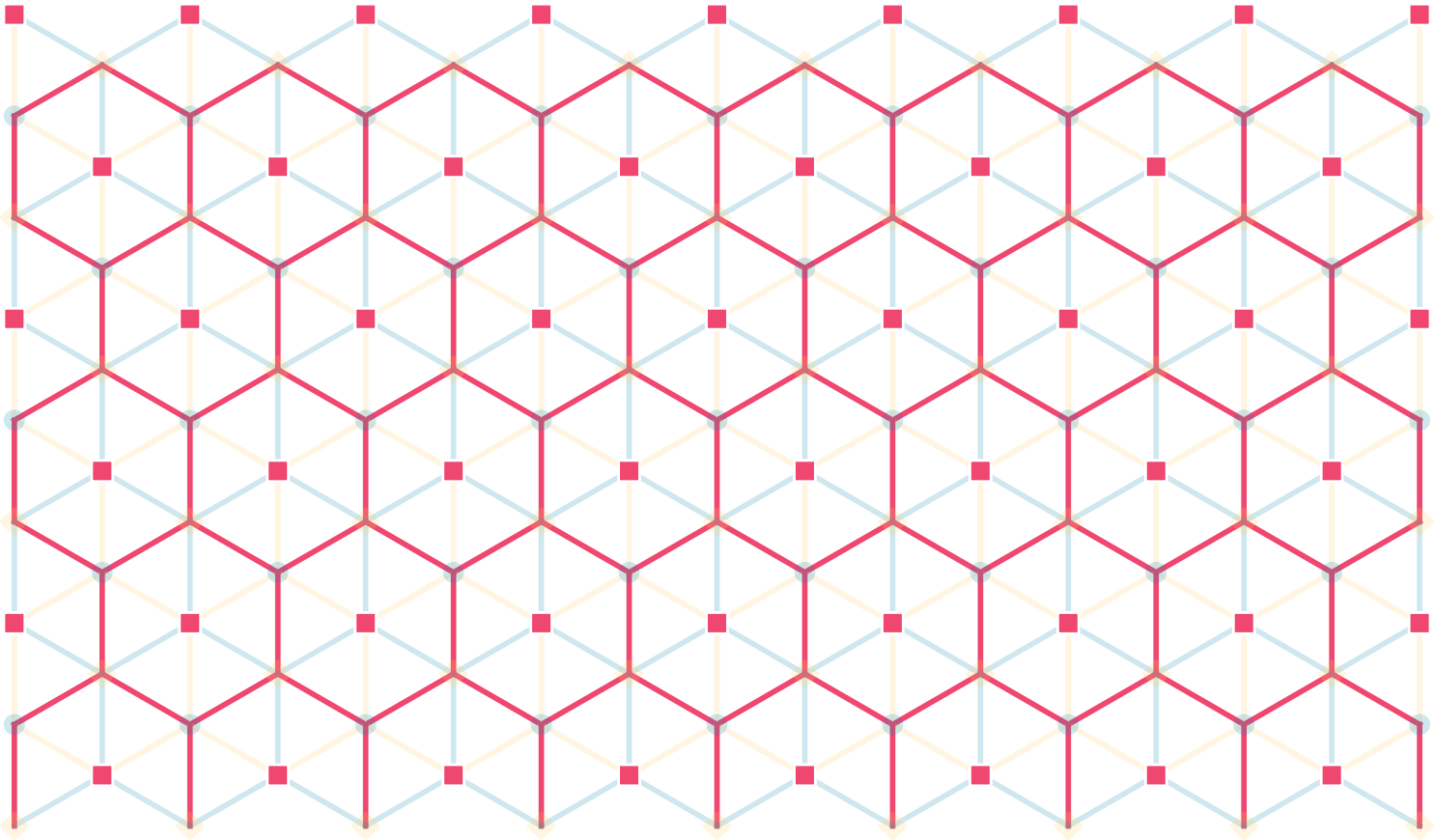


$W_2^2(p_X, p_{\tilde{X}})$

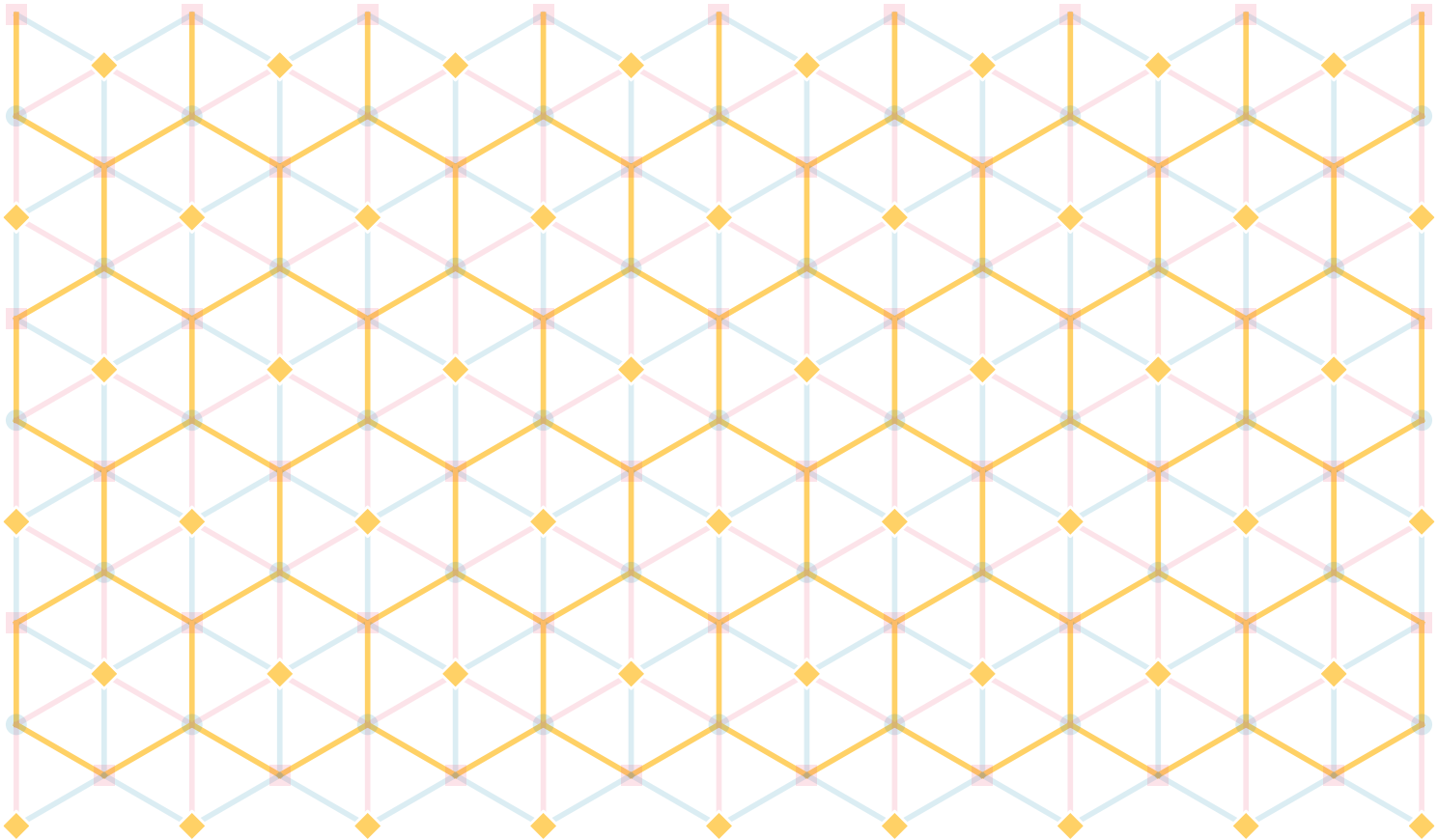
Operational Encoding vs. Virtual Encoding



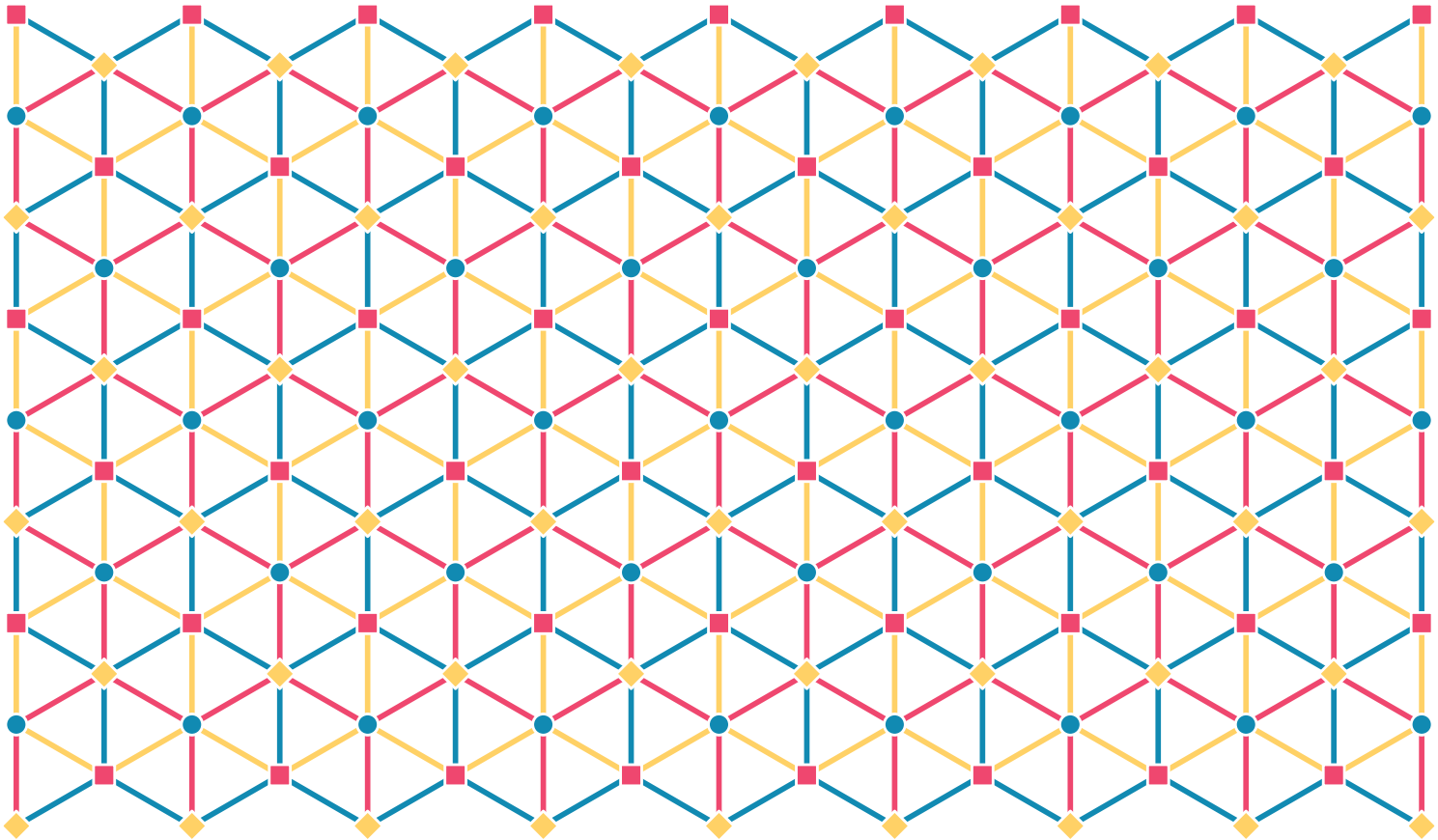
Operational Encoding vs. Virtual Encoding



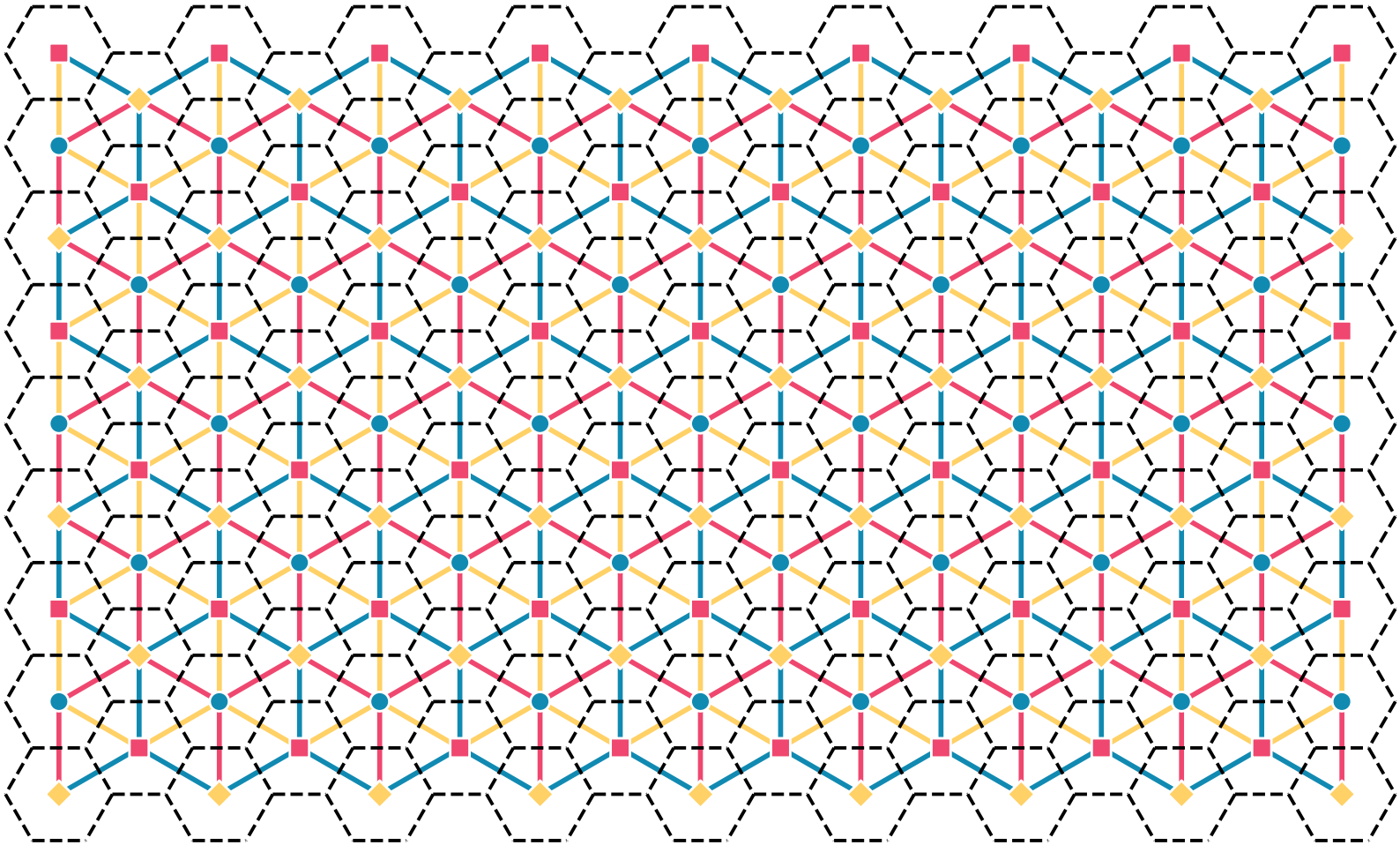
Operational Encoding vs. Virtual Encoding



Operational Encoding vs. Virtual Encoding



Operational Encoding vs. Virtual Encoding



Symbol-Level vs. Sequence-Level

- Symbol-level perception constraint: $\frac{1}{n}\sum_{t=1}^{n} W_2^2(p_X, p_{\hat{X}_t}) \le P$

$$D^{(\infty)}(R,P) = \inf_{p_{\hat{X}|X}} \mathbb{E}_\mu[\|X-\tilde{X}\|^2]$$

$$\text{s.t.} \quad I(X;\tilde{X}) \le R \qquad \text{Blau-Michaeli}$$

$$W_2^2(p_X, p_{\tilde{X}}) \le P$$

- Sequence-level perception constraint: $\frac{1}{n} W_2^2(p_{X^n}, p_{\hat{X}^n}) \le P$

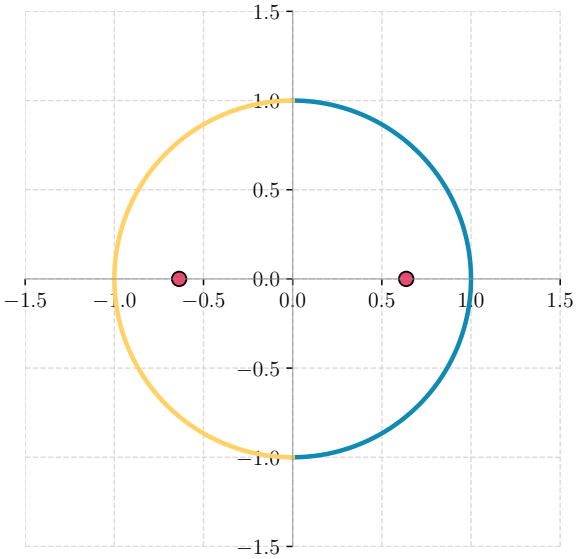
$$D^{(\infty)}(R,C,P) = \inf_{\substack{p_{\tilde{X}} \\ \mu,\nu \in \Pi(p_X, p_{\tilde{X}})}} \mathbb{E}_\mu[\|X-\tilde{X}\|^2] + [\sqrt{\mathbb{E}_\nu[\|X-\tilde{X}\|^2]} - \sqrt{P})_+]^2$$

$$\text{s.t.} \quad I_\mu(X;\tilde{X}) \le R \qquad \text{reduces to Blau-Michaeli}$$

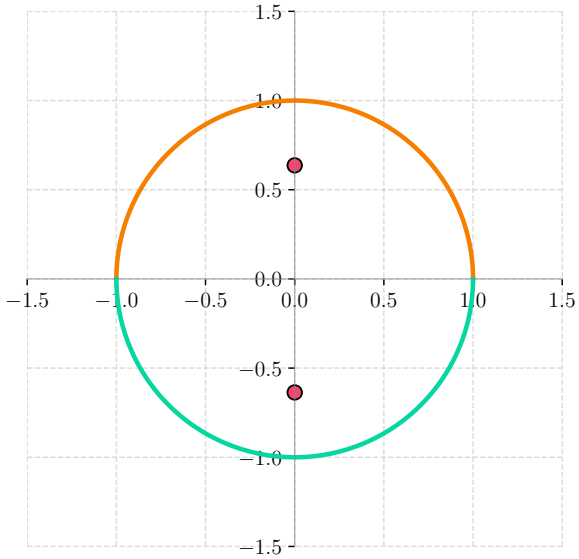
$$(X;\tilde{X}) \le R + C \quad \text{when } C = \infty!$$

$$\tilde{X} = \mathbb{E}_\mu[X|\tilde{X}]$$

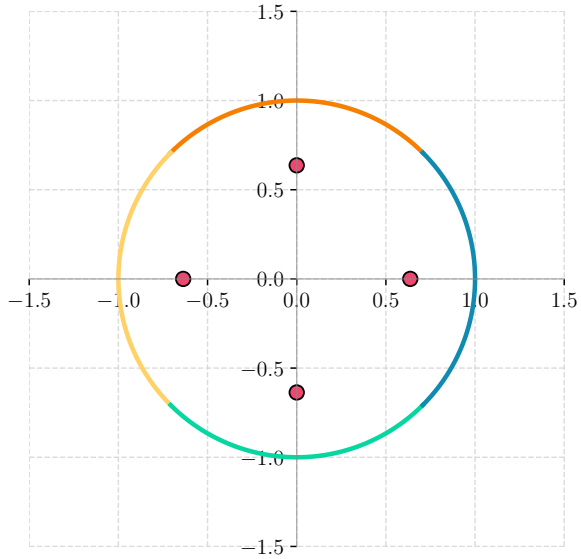
A New Perspective



$\mathbb{E}[\|X - \tilde{X}\|^2 | K = 1]$

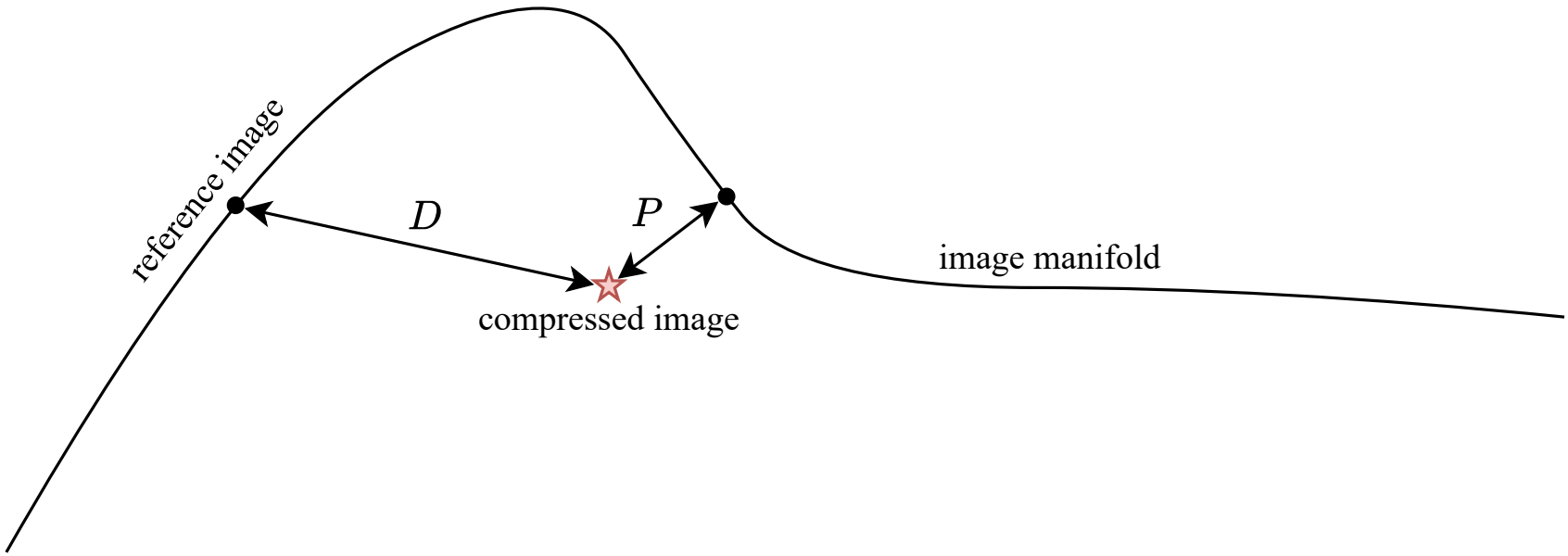


$\mathbb{E}[\|X - \tilde{X}\|^2 | K = 2]$

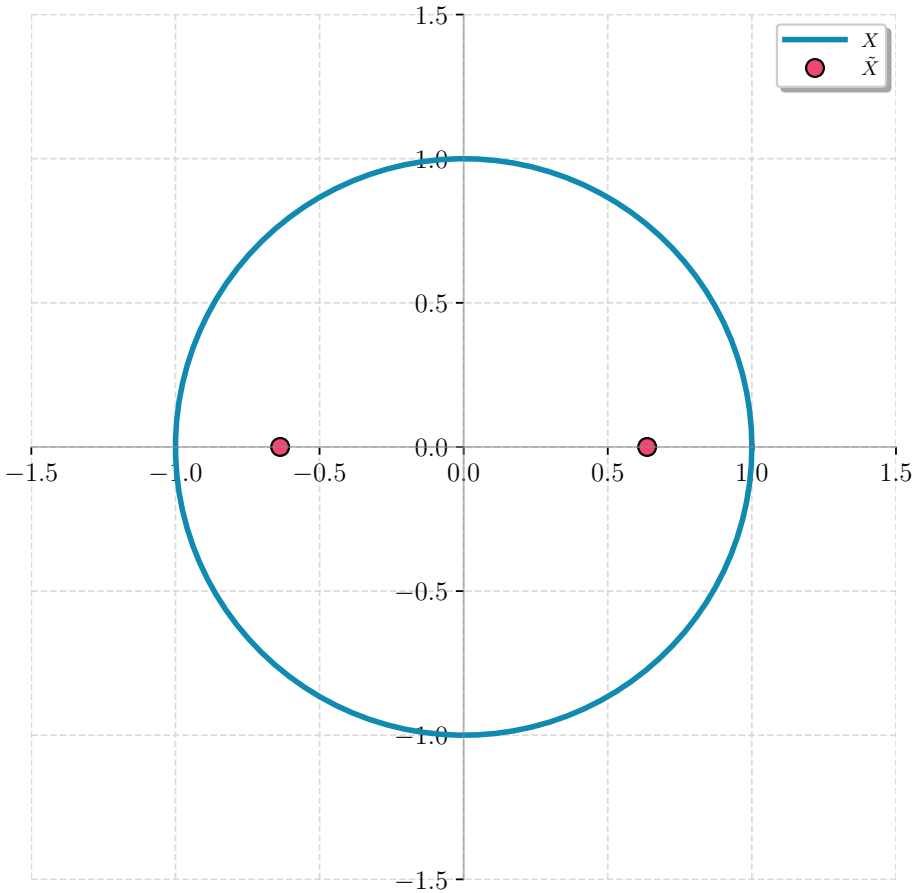


$W_2^2(p_X, p_{\tilde{X}})$

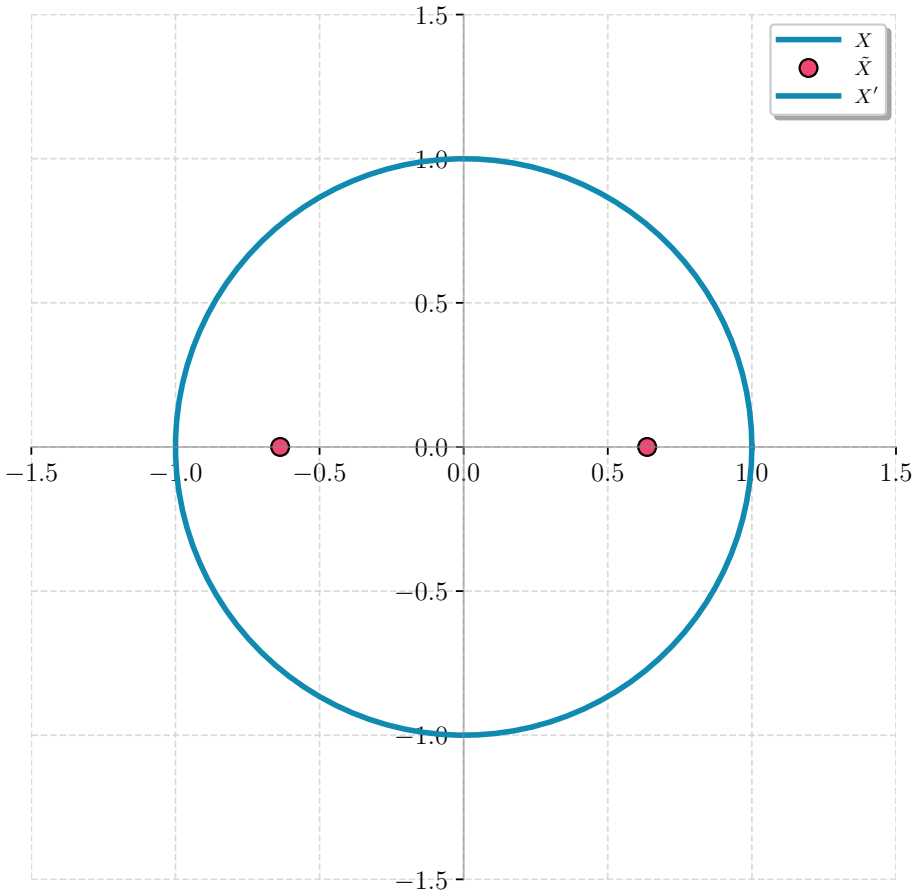
A New Perspective



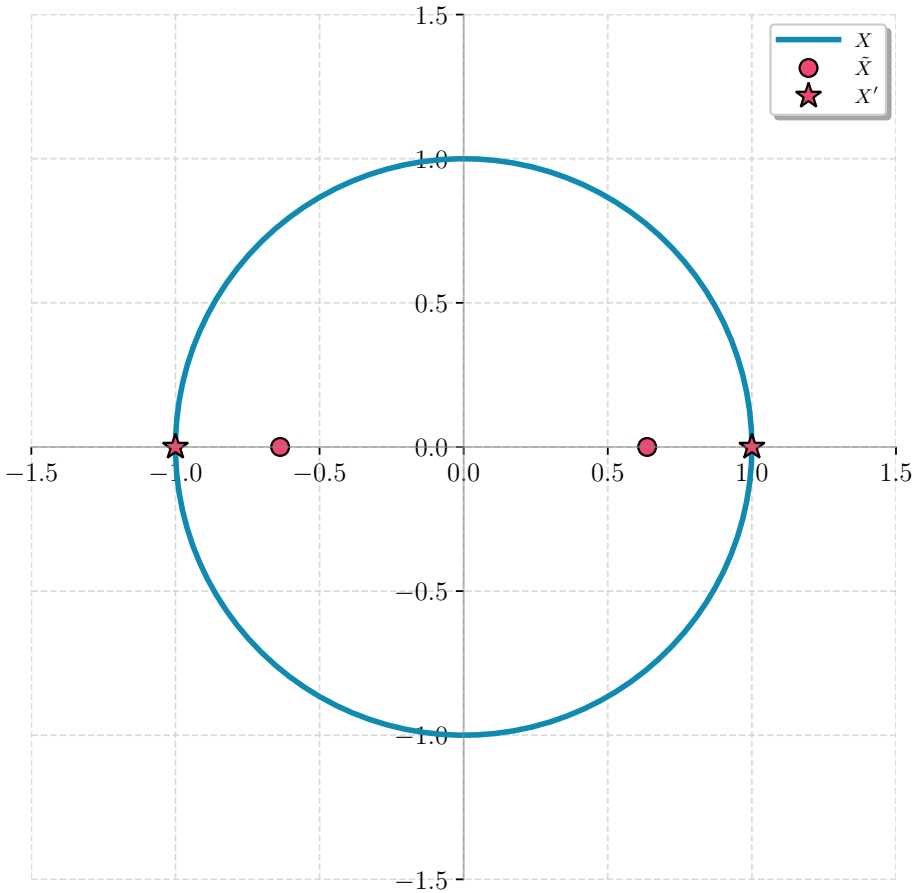
Quantizing the Unit Circle: No Perception Constraint



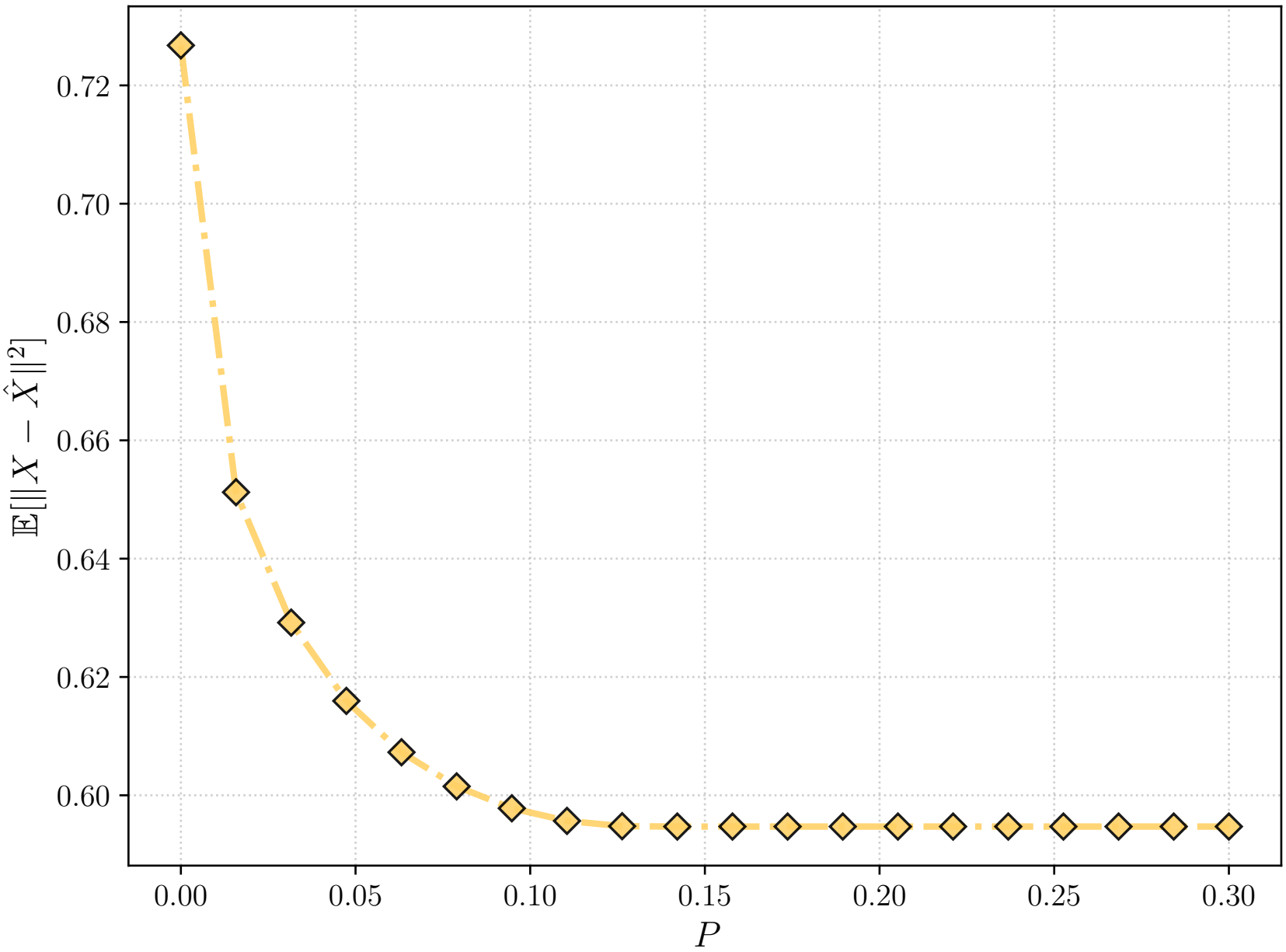
Quantizing the Unit Circle: Perfect Perception (I)



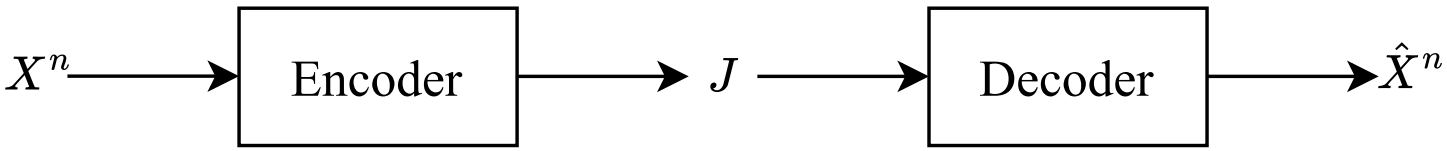
Quantizing the Unit Circle: Perfect Perception (II)



Distortion-Perception Tradeoff



Yet Another Formulation



$$J \in \{1, 2, \ldots, M\}$$

Assume X^n is uniformly distributed over image manifold $\mathcal{M}_I$:

$$D^{(\infty)}(R, P) = \lim_{n\to\infty} \inf_{\mathrm{E,D}} \frac{1}{n} \sum_{t=1}^{n} \mathbb{E}[\|X_t - \hat{X}_t\|^2]$$

$$\text{s.t.} \quad \frac{1}{n} \log M \le R$$

$$\frac{1}{n} \|\hat{X}^n - \mathcal{M}_I\|^2 \le P$$

Coding Theorem

Assume $\mathcal{M}_I = \mathcal{T}(p_X)$:

$$D^{(\infty)}(R,P) = \inf_{p_{\hat{X}|X}} \mathbb{E}[\|X-\hat{X}\|^2]$$

$$\text{s.t.} \quad I(X;\hat{X}) \le R$$

$$W_2^2(p_X, p_{\hat{X}}) \le P$$

Blau-Michaeli again!

Is the Image Manifold a Typical Set?

- No
  Permuting a typical sequence yields another typical sequence; however, permuting the pixels of a natural image does not guarantee a natural image.

- Yes
  Natural images reside on a low-dimensional manifold in a high-dimensional space. This represents a concentration phenomenon, analogous to the behavior of a typical set.

Conclusion

- Rate-distortion-perception theory: An evolving framework
  - Nature of perception measures
  - Role of common randomness
  - Computation of fundamental limits
  - Principles of system design

References

- A. Hussein, J. Chen, C. Tian, S. S. Pradhan, “Constructive approaches to perception-aware lossy source coding: Information-theoretic guidelines,” arXiv:2604.19515.

The End

Questions? Comments?