How to Measure Side-Channel Leakage

Aaron Wagner
School of Electrical and Computer Engineering
Cornell University

Collaborators

Ibrahim Issa
Cornell → EPFL → AUB

Sudeep Kamath
Princeton → $$$

Ed Suh
Cornell

Ben Wu
Cornell

Packet-Timing Side Channel

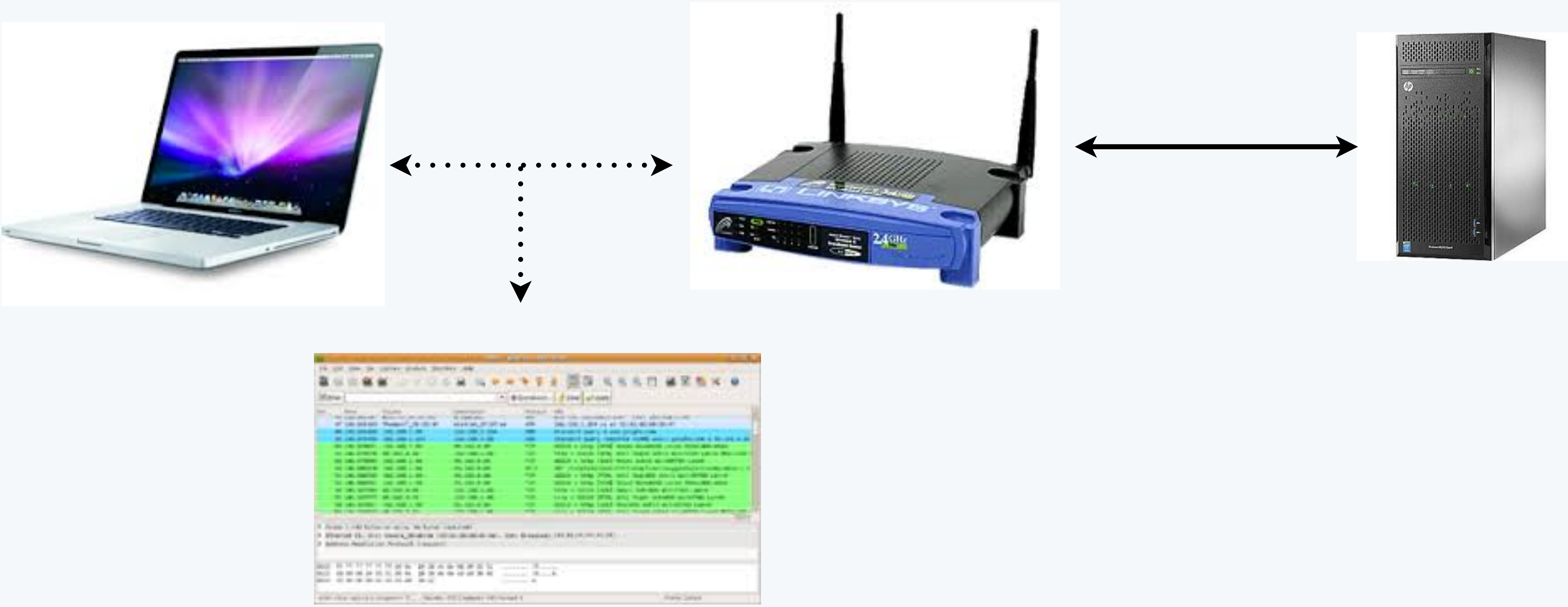
Packet-Timing Side Channel

- ssh: keystrokes are sent as separate packets.

Packet-Timing Side Channel

- ssh: keystrokes are sent as separate packets.
- Packet timing ↔ keystroke timing ↔ typed letters

Packet-Timing Side Channel



- ssh: keystrokes are sent as separate packets.
- Packet timing ↔ keystroke timing ↔ typed letters
- Packet-sniffing eavesdropper can acquire information about typed characters (e.g. passwords).

[Song, Wagner, and Tian ’01]

Side Channels

Side Channels

- *Side channel*: a mechanism that conveys information inadvertently

Side Channels

- *Side channel*: a mechanism that conveys information inadvertently
- Examples:

Side Channels

- *Side channel*: a mechanism that conveys information inadvertently
- Examples:
  - Packet-timing based:

Side Channels

- *Side channel*: a mechanism that conveys information inadvertently
- Examples:
  - Packet-timing based:
    - typed characters (ssh) [Song *et al.* ’01]

Side Channels

- *Side channel*: a mechanism that conveys information inadvertently
- Examples:
  - Packet-timing based:
    - typed characters (ssh) [Song *et al.* ’01]
    - routing [Chaum ’81]

Side Channels

- *Side channel*: a mechanism that conveys information inadvertently
- Examples:
  - Packet-timing based:
    - typed characters (ssh) [Song *et al.* ’01]
    - routing [Chaum ’81]
    - spoken phrases (VoIP) [Wright *et al.* ’08]

Side Channels

- *Side channel*: a mechanism that conveys information inadvertently
- Examples:
  - Packet-timing based:
    - typed characters (ssh) [Song *et al.* ’01]
    - routing [Chaum ’81]
    - spoken phrases (VoIP) [Wright *et al.* ’08]
  - RSA decryption time [Kocher ’96]

Side Channels

- *Side channel*: a mechanism that conveys information inadvertently
- Examples:
  - Packet-timing based:
    - typed characters (ssh) [Song *et al.* ’01]
    - routing [Chaum ’81]
    - spoken phrases (VoIP) [Wright *et al.* ’08]
  - RSA decryption time [Kocher ’96]
  - Cache/memory contention [Ferraiuolo *et al.* ’16]

Side Channels

- *Side channel*: a mechanism that conveys information inadvertently
- Examples:
  - Packet-timing based:
    - typed characters (ssh) [Song *et al.* ’01]
    - routing [Chaum ’81]
    - spoken phrases (VoIP) [Wright *et al.* ’08]
  - RSA decryption time [Kocher ’96]
  - Cache/memory contention [Ferraiuolo *et al.* ’16]
  - CPU power consumption [Kocher *et al.* ’99]

Side Channels

- *Side channel*: a mechanism that conveys information inadvertently
- Examples:
  - Packet-timing based:
    - typed characters (ssh) [Song *et al.* ’01]
    - routing [Chaum ’81]
    - spoken phrases (VoIP) [Wright *et al.* ’08]
  - RSA decryption time [Kocher ’96]
  - Cache/memory contention [Ferraiuolo *et al.* ’16]
  - CPU power consumption [Kocher *et al.* ’99]

Side Channels

TECH

TECH | MOBILE | SOCIAL MEDIA | ENTERPRISE | CYBERSECURITY | TECH GUIDE

Intel sells off for a second day as massive security exploit shakes the stock

- Newly discovered vulnerabilities could theoretically allow a hacker to steal information stored in the memory of chips themselves.
- Although the exploits affected leading processors in many devices, Intel is bearing most of the fallout.
- Some on Wall Street think that Intel's loss could mean gains for rivals.

Anita Balakrishnan | @MsABalakrishnan

Published 11:02 AM ET Thu, 4 Jan 2018 | Updated 1:31 PM ET Thu, 4 Jan 2018

CNBC

- *Meltdown* (Lipp *et al.*, '18)
- *Spectre* (Kocher *et al.*, '18)

Side Channels

- *Side channel*: a mechanism that conveys information inadvertently
- Examples:
  - Packet-timing based:
    - typed characters (ssh) [Song *et al.* ’01]
    - routing [Chaum ’81]
    - spoken phrases (VoIP) [Wright *et al.* ’08]
  - RSA decryption time [Kocher ’96]
  - Cache/memory contention [Ferraiuolo *et al.* ’16]
  - CPU power consumption [Kocher *et al.* ’99]

Side Channels

- *Side channel*: a mechanism that conveys information inadvertently
- Examples:
  - Packet-timing based:
    - typed characters (ssh) [Song *et al.* '01]
    - routing [Chaum '81]
    - spoken phrases (VoIP) [Wright *et al.* '08]
  - RSA decryption time [Kocher '96]
  - Cache/memory contention [Ferraiuolo *et al.* '16]
  - CPU power consumption [Kocher *et al.* '99]

How to measure leakage in this context?

Side Channels

- *Side channel*: a mechanism that conveys information inadvertently
- Examples:
  - Packet-timing based:
    - typed characters (ssh) [Song *et al.* '01]
    - routing [Chaum '81]
    - spoken phrases (VoIP) [Wright *et al.* '08]
  - RSA decryption time [Kocher '96]
  - Cache/memory contention [Ferraiuolo *et al.* '16]
  - CPU power consumption [Kocher *et al.* '99]

Given RVs X and Y, how much does Y leak about X?

Existing Possibilities

Existing Possibilities

- Mutual information (or equivocation) between X and Y

Existing Possibilities

- Mutual information (or equivocation) between X and Y
- Eavesdroppers expected distortion in reproducing X

Existing Possibilities

- Mutual information (or equivocation) between X and Y
- Eavesdroppers expected distortion in reproducing X
- Probability of (approximately) guessing X

Existing Possibilities

- Mutual information (or equivocation) between X and Y
- Eavesdroppers expected distortion in reproducing X
- Probability of (approximately) guessing X
- Expected number of guesses to guess X correctly

Existing Possibilities

- Mutual information (or equivocation) between X and Y
- Eavesdroppers expected distortion in reproducing X
- Probability of (approximately) guessing X
- Expected number of guesses to guess X correctly
- Maximal correlation between X and Y

Existing Possibilities

- Mutual information (or equivocation) between X and Y
- Eavesdroppers expected distortion in reproducing X
- Probability of (approximately) guessing X
- Expected number of guesses to guess X correctly
- Maximal correlation between X and Y
- k-correlation between X and Y

Existing Possibilities

- Mutual information (or equivocation) between X and Y
- Eavesdroppers expected distortion in reproducing X
- Probability of (approximately) guessing X
- Expected number of guesses to guess X correctly
- Maximal correlation between X and Y
- k-correlation between X and Y
- Cryptographic advantage

Existing Possibilities

- Mutual information (or equivocation) between X and Y
- Eavesdroppers expected distortion in reproducing X
- Probability of (approximately) guessing X
- Expected number of guesses to guess X correctly
- Maximal correlation between X and Y
- k-correlation between X and Y
- Cryptographic advantage
- Entropic security

Existing Possibilities

- Mutual information (or equivocation) between X and Y
- Eavesdroppers expected distortion in reproducing X
- Probability of (approximately) guessing X
- Expected number of guesses to guess X correctly
- Maximal correlation between X and Y
- k-correlation between X and Y
- Cryptographic advantage
- Entropic security
- (Local) differential privacy

Existing Possibilities

- Mutual information (or equivocation) between X and Y
- Eavesdroppers expected distortion in reproducing X
- Probability of (approximately) guessing X
- Expected number of guesses to guess X correctly
- Maximal correlation between X and Y
- k-correlation between X and Y
- Cryptographic advantage
- Entropic security
- (Local) differential privacy
- ...

Existing Possibilities

- Mutual information (or equivocation) between X and Y
- Eavesdroppers expected distortion in reproducing X
- Probability of (approximately) guessing X
- Expected number of guesses to guess X correctly
- Maximal correlation between X and Y
- k-correlation between X and Y
- Cryptographic advantage
- Entropic security
- (Local) differential privacy
- ...

Wagner and Eckhoff ('15):

Existing Possibilities

- Mutual information (or equivocation) between X and Y
- Eavesdroppers expected distortion in reproducing X
- Probability of (approximately) guessing X
- Expected number of guesses to guess X correctly
- Maximal correlation between X and Y
- k-correlation between X and Y
- Cryptographic advantage
- Entropic security
- (Local) differential privacy
- ...

Wagner and Eckhoff ('15): 81 metrics

The Threat Model

The Threat Model

1. The eavesdropper is interested in a possibly randomized function of X called U.

The Threat Model

1. The eavesdropper is interested in a possibly randomized function of *X* called *U*.

nominal packet timings

The Threat Model

1. The eavesdropper is interested in a possibly randomized function of *X* called *U*.

password

The Threat Model

1. The eavesdropper is interested in a possibly randomized function of X called U.

2. The eavesdropper observes Y.

The Threat Model

1. The eavesdropper is interested in a possibly randomized function of X called U.

2. The eavesdropper observes Y.

randomized/blurred version of X

The Threat Model

1. The eavesdropper is interested in a possibly randomized function of X called U.

2. The eavesdropper observes Y.

3. The eavesdropper wants to guess, and we want to prevent the eavesdropper from guessing, U.

The Threat Model

1. The eavesdropper is interested in a possibly randomized function of X called U.

2. The eavesdropper observes Y.

3. The eavesdropper wants to guess, and we want to prevent the eavesdropper from guessing, U.

brute-force attack

The Threat Model

1. The eavesdropper is interested in a possibly randomized function of X called U.

2. The eavesdropper observes Y.

3. The eavesdropper wants to guess, and we want to prevent the eavesdropper from guessing, U.

4. The distribution $P_{U|X}(u|x)$ is unknown to us (but known to the eavesdropper)

The Threat Model

1. The eavesdropper is interested in a possibly randomized function of X called U.

2. The eavesdropper observes Y.

3. The eavesdropper wan… we want to prevent the eavesdrop… g, U.

(U,X) joint distribution is complicated; "future-proof"

4. The distribution $P_{U|X}(u|x)$ is unknown to us (but known to the eavesdropper)

Maximal Leakage

U	X	Y
[sensitive info]	[nominal process]	[revealed process]

Maximal Leakage

Markov chain: $U \longleftrightarrow X \longleftrightarrow Y$

[sensitive info] [nominal process] [revealed process]

Maximal Leakage

Markov chain: $U \longleftrightarrow X \longleftrightarrow Y$

[sensitive info] [nominal process] [revealed process]

Def (Issa-Kamath-Wagner)**:** Given P_{XY}, the *maximal leakage* from X to Y is

$$\sup_{\tilde{u}(\cdot)} \Pr(U = \tilde{u}(Y))$$

Maximal Leakage

Markov chain: $U \longleftrightarrow X \longleftrightarrow Y$

[sensitive info] [nominal process] [revealed process]

Def (Issa-Kamath-Wagner)**:** Given P_{XY}, the *maximal leakage* from X to Y is

$$\frac{\sup_{\tilde{u}(\cdot)} \Pr(U = \tilde{u}(Y))}{\sup_{\tilde{u}} \Pr(U = \tilde{u})}$$

Maximal Leakage

Markov chain: $U \longleftrightarrow X \longleftrightarrow Y$

[sensitive info] [nominal process] [revealed process]

Def (Issa-Kamath-Wagner)**:** Given P_{XY}, the *maximal leakage* from X to Y is

$$\log \frac{\sup_{\tilde{u}(\cdot)} \Pr(U = \tilde{u}(Y))}{\sup_{\tilde{u}} \Pr(U = \tilde{u})}$$

Maximal Leakage

Markov chain: $U \longleftrightarrow X \longleftrightarrow Y$

[sensitive info] [nominal process] [revealed process]

Def (Issa-Kamath-Wagner)**:** Given P_{XY}, the *maximal leakage* from X to Y is

$$\sup_{U: U \leftrightarrow X \leftrightarrow Y} \log \frac{\sup_{\tilde{u}(\cdot)} \Pr(U = \tilde{u}(Y))}{\sup_{\tilde{u}} \Pr(U = \tilde{u})}$$

Maximal Leakage

Markov chain: $U \longleftrightarrow X \longleftrightarrow Y$

[sensitive info] [nominal process] [revealed process]

Def (Issa-Kamath-Wagner)**:** Given P_{XY}, the *maximal leakage* from X to Y is

$$\mathcal{L}(X \to Y) = \sup_{U: U \longleftrightarrow X \longleftrightarrow Y} \log \frac{\sup_{\tilde{u}(\cdot)} \Pr(U = \tilde{u}(Y))}{\sup_{\tilde{u}} \Pr(U = \tilde{u})}$$

Maximal Leakage

Markov chain: $U \longleftrightarrow X \longleftrightarrow Y$

[sensitive info] [nominal process] [revealed process]

Def (Issa-Kamath-Wagner)**:** Given P_{XY}, the *maximal leakage* from X to Y is

$$\mathcal{L}(X \to Y) = \sup_{U:U\longleftrightarrow X\longleftrightarrow Y} \log \frac{\sup_{\tilde{u}(\cdot)} \Pr(U = \tilde{u}(Y))}{\sup_{\tilde{u}} \Pr(U = \tilde{u})}$$

[operationally interpretable]

Maximal Leakage

Markov chain: $U \longleftrightarrow X \longleftrightarrow Y$

[sensitive info] [nominal process] [revealed process]

Def (Issa-Kamath-Wagner)**:** Given P_{XY}, the *maximal leakage* from X to Y is

$$\mathcal{L}(X \to Y) = \sup_{U: U \leftrightarrow X \leftrightarrow Y} \log \frac{\sup_{\tilde{u}(\cdot)} \Pr(U = \tilde{u}(Y))}{\sup_{\tilde{u}} \Pr(U = \tilde{u})}$$

[not evidently computable; Carathéodory?]

Maximal Leakage

Theorem (Issa-Kamath-Wagner): For any joint distribution P_{XY} on finite alphabets

Maximal Leakage

Theorem (Issa-Kamath-Wagner): For any joint distribution P_{XY} on finite alphabets

$$\mathcal{L}(X \to Y) = \log \sum_{y \in \mathcal{Y}} \max_{\substack{x \in \mathcal{X}: \\ P_X(x) > 0}} P_{Y|X}(y|x)$$

Maximal Leakage

Theorem (Issa-Kamath-Wagner): For any joint distribution P_{XY} on finite alphabets

$$\mathcal{L}(X \to Y) = \log \sum_{y \in \mathcal{Y}} \max_{\substack{x \in \mathcal{X}: \\ P_X(x) > 0}} P_{Y|X}(y|x)$$

$$= I_\infty(X;Y) \quad \text{[Sibson MI of order } \infty\text{]}$$

Maximal Leakage

Theorem (Issa-Kamath-Wagner): For any joint distribution P_{XY} on finite alphabets

$$\mathcal{L}(X \to Y) = \log \sum_{y \in \mathcal{Y}} \max_{\substack{x \in \mathcal{X}: \\ P_X(x) > 0}} P_{Y|X}(y|x)$$

$$= I_\infty(X;Y) \quad \text{[Sibson MI of order } \infty]$$

[depends on P_X only through its support]

The Worst-Case U

$$\mathcal{L}(X \to Y) = \sup_{U \leftrightarrow X \leftrightarrow Y} \log \frac{\sup_{\tilde{u}(\cdot)} \Pr(U = \tilde{u}(Y))}{\sup_{\tilde{u}} \Pr(U = \tilde{u})}$$

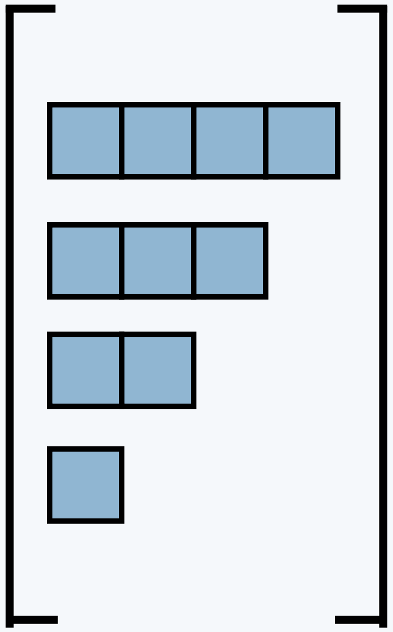
The Worst-Case U

$$\mathcal{L}(X \to Y) = \sup_{U \leftrightarrow X \leftrightarrow Y} \log \frac{\sup_{\tilde{u}(\cdot)} \Pr(U = \tilde{u}(Y))}{\sup_{\tilde{u}} \Pr(U = \tilde{u})}$$

$$\underset{P_X}{\begin{bmatrix} \\ \\ \\ \end{bmatrix}}$$

The Worst-Case U

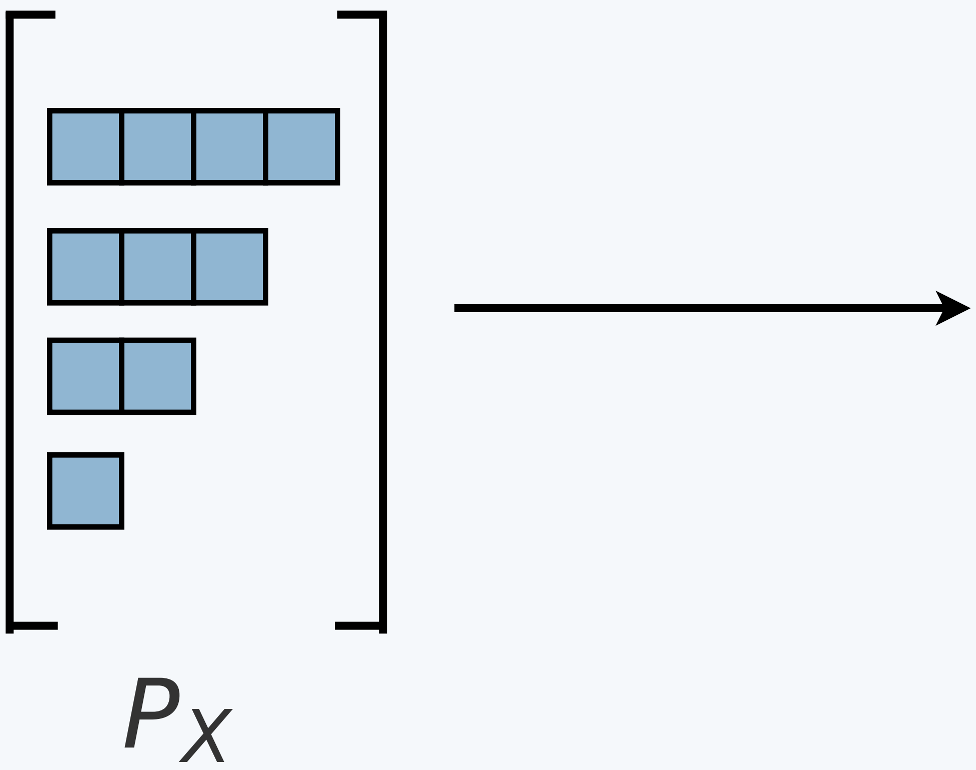
$$\mathcal{L}(X \to Y) = \sup_{U \leftrightarrow X \leftrightarrow Y} \log \frac{\sup_{\tilde{u}(\cdot)} \Pr(U = \tilde{u}(Y))}{\sup_{\tilde{u}} \Pr(U = \tilde{u})}$$



P_X

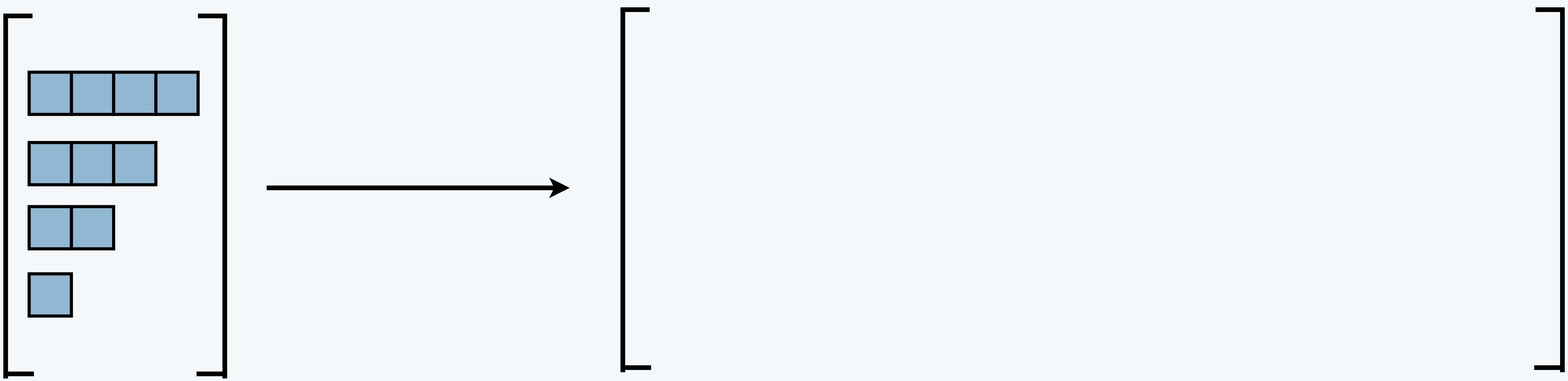
The Worst-Case U

$$\mathcal{L}(X \to Y) = \sup_{U \leftrightarrow X \leftrightarrow Y} \log \frac{\sup_{\tilde{u}(\cdot)} \Pr(U = \tilde{u}(Y))}{\sup_{\tilde{u}} \Pr(U = \tilde{u})}$$



The Worst-Case U

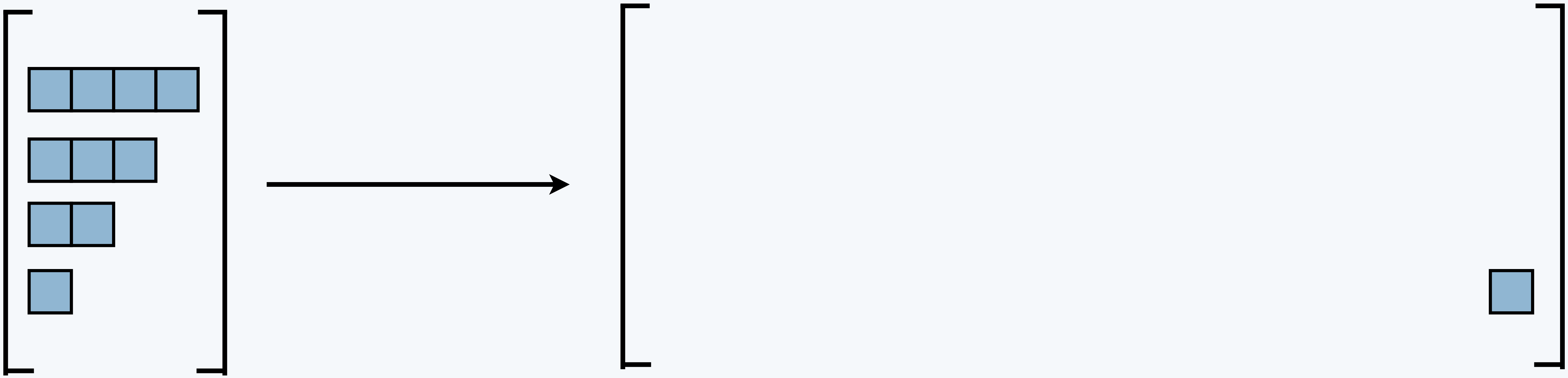
$$\mathcal{L}(X \to Y) = \sup_{U \leftrightarrow X \leftrightarrow Y} \log \frac{\sup_{\tilde{u}(\cdot)} \Pr(U = \tilde{u}(Y))}{\sup_{\tilde{u}} \Pr(U = \tilde{u})}$$



P_X P_{XU}

The Worst-Case U

$$\mathcal{L}(X \to Y) = \sup_{U \leftrightarrow X \leftrightarrow Y} \log \frac{\sup_{\tilde{u}(\cdot)} \Pr(U = \tilde{u}(Y))}{\sup_{\tilde{u}} \Pr(U = \tilde{u})}$$



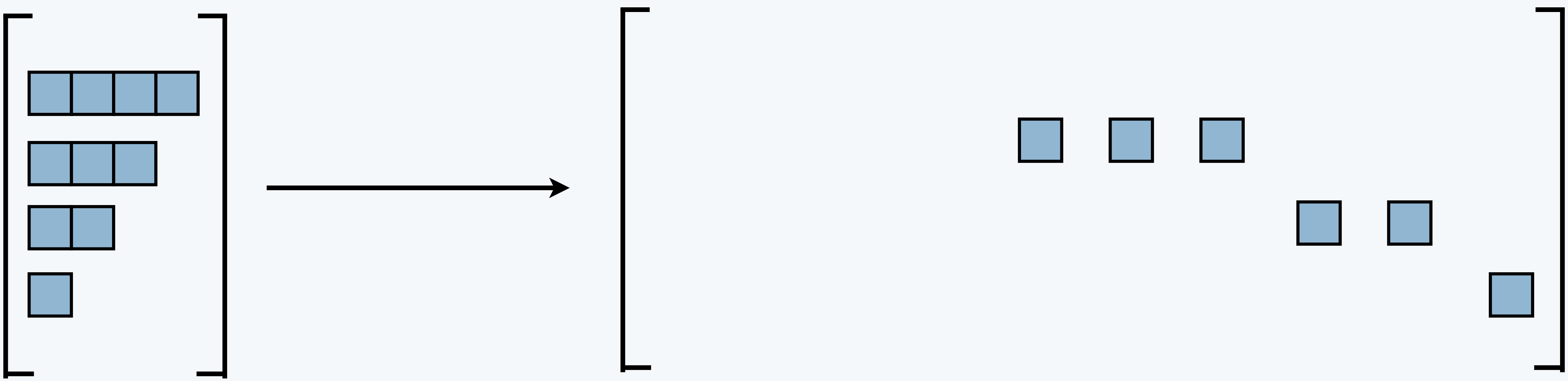
P_X P_{XU}

The Worst-Case U

$$\mathcal{L}(X \to Y) = \sup_{U \leftrightarrow X \leftrightarrow Y} \log \frac{\sup_{\tilde{u}(\cdot)} \Pr(U = \tilde{u}(Y))}{\sup_{\tilde{u}} \Pr(U = \tilde{u})}$$

The Worst-Case U

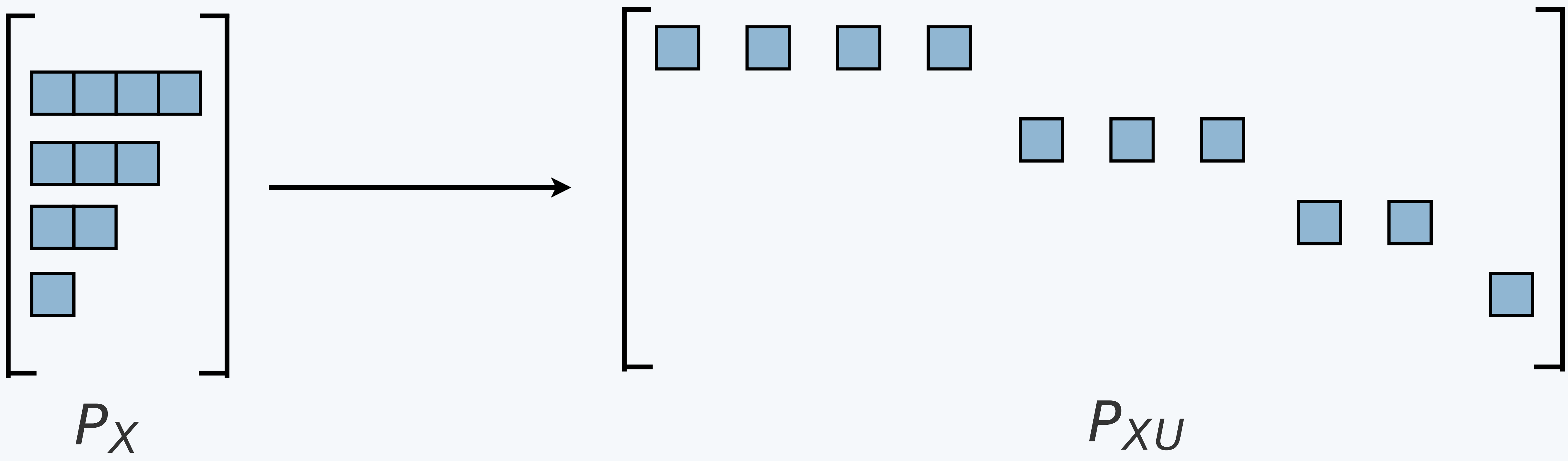
$$\mathcal{L}(X \to Y) = \sup_{U \leftrightarrow X \leftrightarrow Y} \log \frac{\sup_{\tilde{u}(\cdot)} \Pr(U = \tilde{u}(Y))}{\sup_{\tilde{u}} \Pr(U = \tilde{u})}$$



P_X P_{XU}

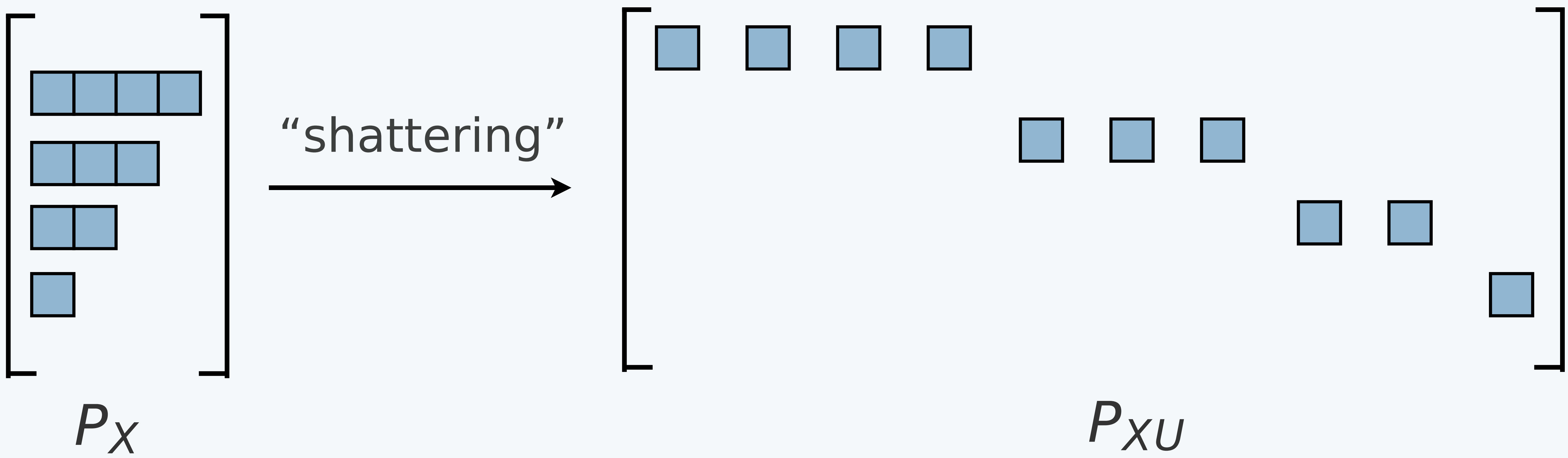
The Worst-Case U

$$\mathcal{L}(X \to Y) = \sup_{U \leftrightarrow X \leftrightarrow Y} \log \frac{\sup_{\tilde{u}(\cdot)} \Pr(U = \tilde{u}(Y))}{\sup_{\tilde{u}} \Pr(U = \tilde{u})}$$



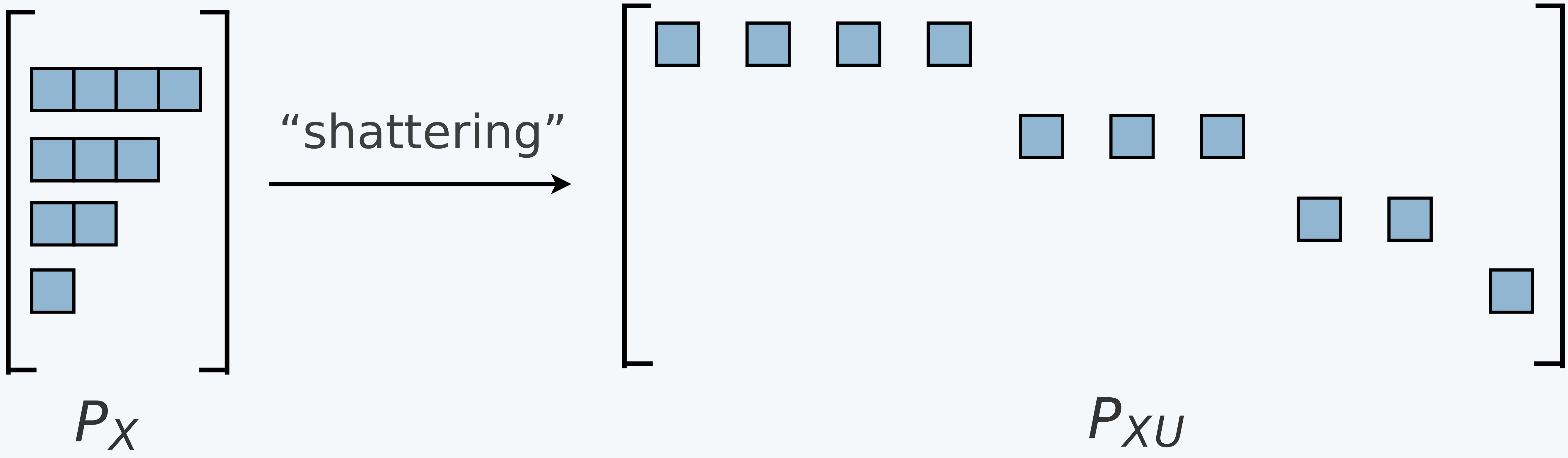
The Worst-Case U

$$\mathcal{L}(X \to Y) = \sup_{U \leftrightarrow X \leftrightarrow Y} \log \frac{\sup_{\tilde{u}(\cdot)} \Pr(U = \tilde{u}(Y))}{\sup_{\tilde{u}} \Pr(U = \tilde{u})}$$



The Worst-Case U

$$\mathcal{L}(X \to Y) = \sup_{U \leftrightarrow X \leftrightarrow Y} \log \frac{\sup_{\tilde{u}(\cdot)} \Pr(U = \tilde{u}(Y))}{\sup_{\tilde{u}} \Pr(U = \tilde{u})}$$



[U is uniform and s.t. X is a deterministic function of U]

Upper Bound

$$
\begin{aligned}
&\sum_{y\in\mathcal{Y}} P_Y(y) \max_{u\in\mathcal{U}} P_{U|Y}(u|y) \\
&= \sum_{y\in\mathcal{Y}} \max_{u\in\mathcal{U}} P_{UY}(u, y) \\
&= \sum_{y\in\mathcal{Y}} \max_{u\in\mathcal{U}} \sum_{x\in\mathcal{X}} P_X(x) P_{U|X}(u|x) P_{Y|X}(y|x) \\
&\leq \sum_{y\in\mathcal{Y}} \max_{u\in\mathcal{U}} \sum_{x\in\mathcal{X}} P_X(x) P_{U|X}(u|x) \max_{x'\in\mathcal{X}} P_{Y|X}(y|x') \\
&= \sum_{y\in\mathcal{Y}} \left(\max_{x'\in\mathcal{X}} P_{Y|X}(y|x') \right) \max_{u\in\mathcal{U}} \sum_{x\in\mathcal{X}} P_X(x) P_{U|X}(u|x) \\
&= \sum_{y\in\mathcal{Y}} \max_{x\in\mathcal{X}} P_{Y|X}(y|x) \max_{u\in\mathcal{U}} P_U(u).
\end{aligned}
$$

Maximal Leakage

Theorem (Issa-Kamath-Wagner): For any joint distribution P_{XY} on finite alphabets

$$\mathcal{L}(X \to Y) = \log \sum_{y \in \mathcal{Y}} \max_{\substack{x \in \mathcal{X}: \\ P_X(x) > 0}} P_{Y|X}(y|x)$$

$$= I_\infty(X;Y) \quad \text{[Sibson MI of order } \infty\text{]}$$

Properties of Max. Leakage

Properties of Max. Leakage

Corollary: For any joint distribution P_{XY} on finite alphabets

Properties of Max. Leakage

Corollary: For any joint distribution P_{XY} on finite alphabets

- Data processing inequality: If $X \leftrightarrow Y \leftrightarrow Z$ then

$$\mathcal{L}(X \to Z) \le \min\{\mathcal{L}(X \to Y), \mathcal{L}(Y \to Z)\}$$

Properties of Max. Leakage

Corollary: For any joint distribution P_{XY} on finite alphabets

- Data processing inequality: If $X \leftrightarrow Y \leftrightarrow Z$ then

$$\mathcal{L}(X \to Z) \leq \min\{\mathcal{L}(X \to Y), \mathcal{L}(Y \to Z)\}$$

- Self-leakage

$$\mathcal{L}(X \to X) = \log |\{x : P_X(x) > 0\}|$$

Properties of Max. Leakage

Corollary: For any joint distribution P_{XY} on finite alphabets

- Data processing inequality: If $X \leftrightarrow Y \leftrightarrow Z$ then

$$\mathcal{L}(X \to Z) \leq \min\{\mathcal{L}(X \to Y), \mathcal{L}(Y \to Z)\}$$

- Self-leakage

$$\mathcal{L}(X \to X) = \log |\{x : P_X(x) > 0\}|$$

- Cardinality bound

$$\mathcal{L}(X \to Y) \leq \min\{\log |\mathcal{X}|, \log |\mathcal{Y}|\}$$

Properties of Max. Leakage

Properties of Max. Leakage

- Independence: $\mathcal{L}(X \rightarrow Y) = 0$ iff X and Y are indep.

Properties of Max. Leakage

- Independence: $\mathcal{L}(X \to Y) = 0$ iff X and Y are indep.
- Asymmetry: $\mathcal{L}(X \to Y) \neq \mathcal{L}(Y \to X)$ in general.

Properties of Max. Leakage

- Independence: $\mathcal{L}(X \to Y) = 0$ iff X and Y are indep.
- Asymmetry: $\mathcal{L}(X \to Y) \neq \mathcal{L}(Y \to X)$ in general.
- Additivity: if $(X_i, Y_i)_{i=1}^n$ are independent over i

$$\mathcal{L}(X^n \to Y^n) = \sum_{i=1}^{n} \mathcal{L}(X_i \to Y_i)$$

Properties of Max. Leakage

- Independence: $\mathcal{L}(X \to Y) = 0$ iff X and Y are indep.
- Asymmetry: $\mathcal{L}(X \to Y) \neq \mathcal{L}(Y \to X)$ in general.
- Additivity: if $(X_i, Y_i)_{i=1}^n$ are independent over i

$$\mathcal{L}(X^n \to Y^n) = \sum_{i=1}^{n} \mathcal{L}(X_i \to Y_i)$$

- Convexity: $\exp(\mathcal{L}(X \to Y))$ is convex in $P_{Y|X}$

Properties of Max. Leakage

- Independence: $\mathcal{L}(X \to Y) = 0$ iff X and Y are indep.
- Asymmetry: $\mathcal{L}(X \to Y) \neq \mathcal{L}(Y \to X)$ in general.
- Additivity: if $(X_i, Y_i)_{i=1}^n$ are independent over i

$$\mathcal{L}(X^n \to Y^n) = \sum_{i=1}^{n} \mathcal{L}(X_i \to Y_i)$$

- Convexity: $\exp(\mathcal{L}(X \to Y))$ is convex in $P_{Y|X}$
- Maximal leakage upper bounds mutual info.

$$\mathcal{L}(X \to Y) \geq I(X;Y) = \sum_{x,y} p(x,y) \log \frac{p(x,y)}{p(x)p(y)}$$

Variations and Extensions

- Multiple guesses
- Approximate guesses
- General gains
- Opportunistic choice of U
- Conditional version
- Formula for general measure spaces
- Guessing X itself

Extension: Multiple Guesses

Def (Issa-Kamath-Wagner)**:** For any positive integer k,

$$\mathcal{L}_k(X \to Y) = \sup_{U \longleftrightarrow X \longleftrightarrow Y} \log \frac{\sup_{\tilde{u}_1(\cdot),\ldots,\tilde{u}_k(\cdot)} P(\cup_i \{U = \tilde{u}_i(Y)\})}{\sup_{\tilde{u}_1,\ldots,\tilde{u}_k} P(\cup_i \{U = \tilde{u}_i\})}$$

Extension: Multiple Guesses

Def (Issa-Kamath-Wagner)**:** For any positive integer k,

$$\mathcal{L}_k(X \to Y) = \sup_{U \leftrightarrow X \leftrightarrow Y} \log \frac{\sup_{\tilde{u}_1(\cdot),\ldots,\tilde{u}_k(\cdot)} P(\cup_i \{U = \tilde{u}_i(Y)\})}{\sup_{\tilde{u}_1,\ldots,\tilde{u}_k} P(\cup_i \{U = \tilde{u}_i\})}$$

Theorem (Issa-Kamath-Wagner)**:** If X and Y are discrete then for any positive integer k,

$$\mathcal{L}_k(X \to Y) = \mathcal{L}_1(X \to Y) = \mathcal{L}(X \to Y).$$

Conditional Form

Definition: The conditional maximal leakage from X to Y given Z is

$$\mathcal{L}(X \to Y|Z) = \sup_{U \leftrightarrow X \leftrightarrow Y|Z} \log \frac{\sup_{\tilde{u}(\cdot,\cdot)} \Pr(U = \tilde{u}(Y, Z))}{\sup_{\tilde{u}(\cdot)} \Pr(U = \tilde{u}(Z))}$$

Conditional Form

Definition: The conditional maximal leakage from X to Y given Z is

$$\mathcal{L}(X \to Y|Z) = \sup_{U \leftrightarrow X \leftrightarrow Y|Z} \log \frac{\sup_{\tilde{u}(\cdot,\cdot)} \Pr(U = \tilde{u}(Y,Z))}{\sup_{\tilde{u}(\cdot)} \Pr(U = \tilde{u}(Z))}$$

vs. $U \leftrightarrow X \leftrightarrow (Y,Z)$

Conditional Form

Definition: The conditional maximal leakage from X to Y given Z is

$$\mathcal{L}(X \to Y|Z) = \sup_{U \leftrightarrow X \leftrightarrow Y|Z} \log \frac{\sup_{\tilde{u}(\cdot,\cdot)} \Pr(U = \tilde{u}(Y,Z))}{\sup_{\tilde{u}(\cdot)} \Pr(U = \tilde{u}(Z))}$$

Conditional Form

Definition: The conditional maximal leakage from X to Y given Z is

$$\mathcal{L}(X \to Y|Z) = \sup_{U \leftrightarrow X \leftrightarrow Y|Z} \log \frac{\sup_{\tilde{u}(\cdot,\cdot)} \Pr(U = \tilde{u}(Y,Z))}{\sup_{\tilde{u}(\cdot)} \Pr(U = \tilde{u}(Z))}$$

Theorem (Issa-Wagner):

$$\mathcal{L}(X \to Y|Z) = \max_{z} \mathcal{L}(X \to Y|Z = z)$$

Properties of Cond. Max. Leakage

Corollary: For any joint distribution P_{XYZ} on finite alphabets

- Data processing inequality: If $X \leftrightarrow Y \leftrightarrow V|Z$ then

$$\mathcal{L}(X \to V|Z) \leq \min\{\mathcal{L}(X \to Y|Z), \mathcal{L}(Y \to V|Z)\}$$

- Cond. independence: $\mathcal{L}(X \to Y|Z) = 0$ iff

$$X \leftrightarrow Z \leftrightarrow Y$$

- Mutual information:

$$\mathcal{L}(X \to Y|Z) \geq I(X; Y|Z)$$

Properties of Cond. Max. Leakage

Properties of Cond. Max. Leakage

- Conditioning reduces max. leakage: if $Z \longleftrightarrow X \longleftrightarrow Y$ then

$$\mathcal{L}(X \to Y | Z) \leq \mathcal{L}(X \to Y)$$

Properties of Cond. Max. Leakage

- Conditioning reduces max. leakage: if $Z \longleftrightarrow X \longleftrightarrow Y$ then

$$\mathcal{L}(X \to Y|Z) \leq \mathcal{L}(X \to Y)$$

- Chain rule:

$$\mathcal{L}(X \to (Y, Z)) \leq \mathcal{L}(X \to Z) + \mathcal{L}(X \to Y|Z)$$

Properties of Cond. Max. Leakage

- Conditioning reduces max. leakage: if $Z \longleftrightarrow X \longleftrightarrow Y$ then

$$\mathcal{L}(X \to Y | Z) \leq \mathcal{L}(X \to Y)$$

- Chain rule:

$$\mathcal{L}(X \to (Y, Z)) \leq \mathcal{L}(X \to Z) + \mathcal{L}(X \to Y | Z)$$

- Composition theorem: if $Z \longleftrightarrow X \longleftrightarrow Y$ then

$$\mathcal{L}(X \to (Y, Z)) \leq \mathcal{L}(X \to Z) + \mathcal{L}(X \to Y)$$

Guessing X

Def:

$$\mathcal{L}_I(X \to Y) = \sup_{P_X} \log \frac{\max_{\hat{x}(\cdot)} P(X = \hat{x}(Y))}{\max_{\hat{x}} P(X = \hat{x})}$$

Guessing X

Def:

$$\mathcal{L}_I(X \to Y) = \sup_{P_X} \log \frac{\max_{\hat{x}(\cdot)} P(X = \hat{x}(Y))}{\max_{\hat{x}} P(X = \hat{x})}$$

Theorem:

$$\mathcal{L}_I(X \to Y) = I_\infty [= \mathcal{L}(X \to Y)]$$

Guessing X

Def: [Braun *et al.* '09; Kopf and Smith '10]:

$$\mathcal{L}_I(X \to Y) = \sup_{P_X} \log \frac{\max_{\hat{x}(\cdot)} P(X = \hat{x}(Y))}{\max_{\hat{x}} P(X = \hat{x})}$$

Theorem: [Braun *et al.* '09; Kopf and Smith '10]:

$$\mathcal{L}_I(X \to Y) = I_\infty [= \mathcal{L}(X \to Y)]$$

Guessing X

Def: [Braun *et al.* '09; Kopf and Smith '10]:

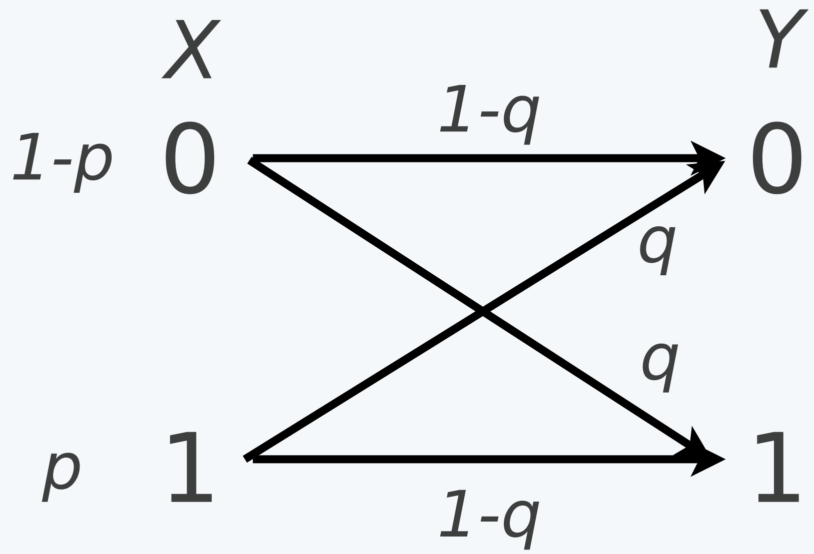
$$\mathcal{L}_I(X \to Y) = \sup_{P_X} \log \frac{\max_{\hat{x}(\cdot)} P(X = \hat{x}(Y))}{\max_{\hat{x}} P(X = \hat{x})}$$

Theorem: [Braun *et al.* '09; Kopf and Smith '10]:

$$\mathcal{L}_I(X \to Y) = I_\infty [= \mathcal{L}(X \to Y)]$$

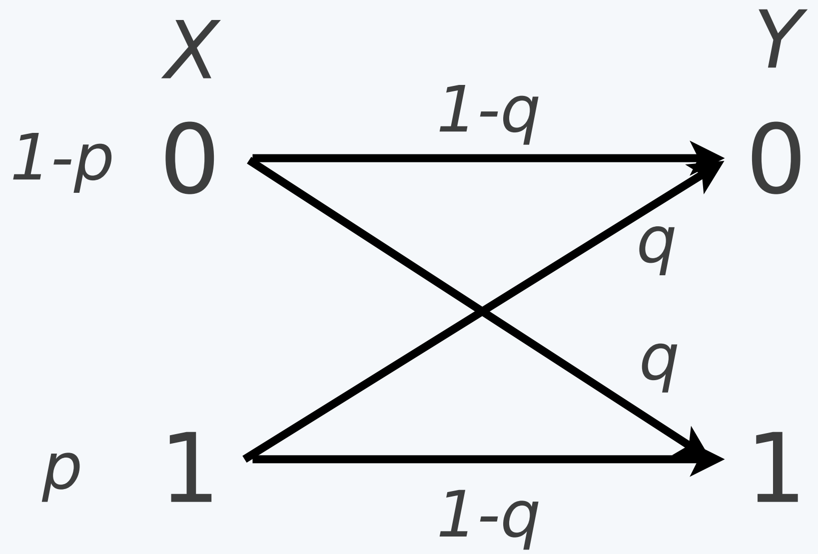
[*maximal leakage*: not in Wagner and Eckhoff ('15)]

Discrete Examples: BSC

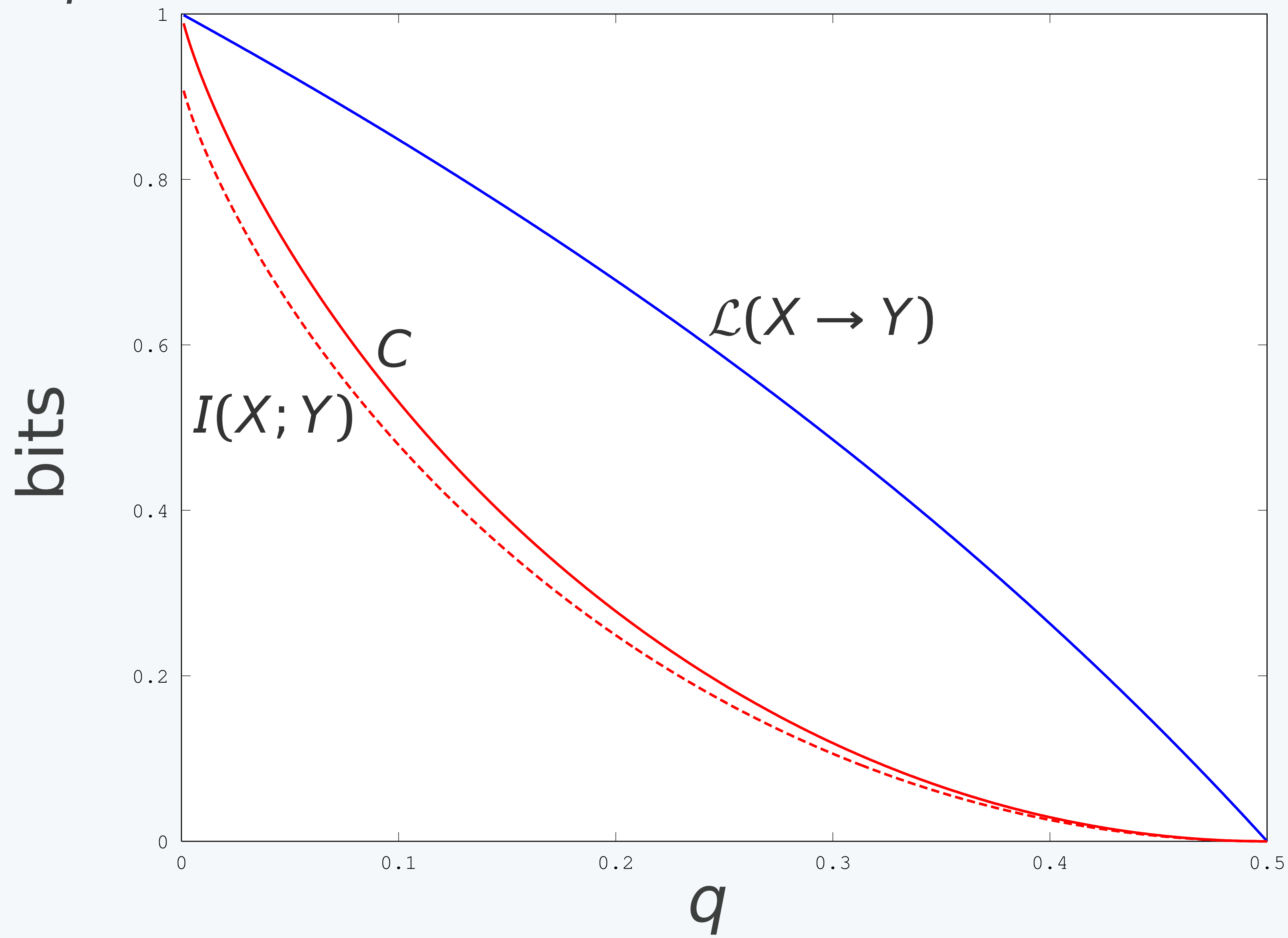


$$\mathcal{L}(X \to Y) = \log(2(1-q))$$

Discrete Examples: BSC



$$\mathcal{L}(X \to Y) = \log(2(1-q))$$



$[p = 1/3]$

Continuous Example

Theorem (Issa-Kamath-Wagner)**:** If $f_X(x)$ *and* $f_{Y|X}(y|x)$ are continuous then:

$$\mathcal{L}(X \to Y) = \log \int \sup_{x : f_X(x) > 0} f_{Y|X}(y|x)\, dy$$

Continuous Example

Theorem (Issa-Kamath-Wagner)**:** If $f_X(x)$ *and* $f_{Y|X}(y|x)$ are continuous then:

$$\mathcal{L}(X \to Y) = \log \int \sup_{x : f_X(x) > 0} f_{Y|X}(y|x)\, dy$$

If X and Y are jointly Gaussian then

$$\mathcal{L}(X \to Y) = \begin{cases} 0 & \text{if } X, Y \text{ indep.} \\ \infty & \text{otherwise} \end{cases}$$

Continuous Example

Theorem (Issa-Kamath-Wagner)**:** If $f_X(x)$ *and* $f_{Y|X}(y|x)$ are continuous then:

$$\mathcal{L}(X \to Y) = \log \int \sup_{x : f_X(x) > 0} f_{Y|X}(y|x)\, dy$$

If X and Y are jointly Gaussian then

$$\mathcal{L}(X \to Y) = \begin{cases} 0 & \text{if } X, Y \text{ indep.} \\ \infty & \text{otherwise} \end{cases}$$

[“adding noise” (as opposed to quantizing) leaks]

Continuous Example

Theorem (Issa-Kamath-Wagner)**:** If $f_X(x)$ *and* $f_{Y|X}(y|x)$ are continuous then:

$$\mathcal{L}(X \to Y) = \log \int \sup_{x : f_X(x) > 0} f_{Y|X}(y|x)\, dy$$

If X and Y are jointly Gaussian then

$$\mathcal{L}(X \to Y) = \begin{cases} 0 & \text{if } X, Y \text{ indep.} \\ \infty & \text{otherwise} \end{cases}$$

Other Metrics

- Mutual information (or equivocation)
- Expected distortion at eavesdropper
- Probability of (approximately) guessing X
- Expected number of guesses to guess X correctly
- Maximal correlation
- k-correlation
- Cryptographic advantage
- Entropic security
- (Local) differential privacy
- ...

Other Metrics

- Mutual information (or equivocation)
- Expected distortion at eavesdropper
- Probability of (approximately) guessing X
- Expected number of guesses to guess X correctly
- Maximal correlation
- k-correlation
- Cryptographic advantage
- Entropic security
- (Local) differential privacy
- ...

Mutual Information

$$I(X;Y) = \sum_{x,y} P_{XY}(x,y) \log \frac{P_{XY}(x,y)}{P_X(x) \cdot P_Y(y)}$$

Mutual Information

$$I(X;Y) = \sum_{x,y} P_{XY}(x,y) \log \frac{P_{XY}(x,y)}{P_X(x) \cdot P_Y(y)}$$

$\Longleftrightarrow H(X|Y)$, first used by Shannon ('49)

Mutual Information

$$I(X;Y) = \sum_{x,y} P_{XY}(x,y) \log \frac{P_{XY}(x,y)}{P_X(x) \cdot P_Y(y)}$$

$\Longleftrightarrow H(X|Y)$, first used by Shannon ('49)

solution concept vs. problem formulation

Mutual Information

$$I(X;Y) = \sum_{x,y} P_{XY}(x,y) \log \frac{P_{XY}(x,y)}{P_X(x) \cdot P_Y(y)}$$

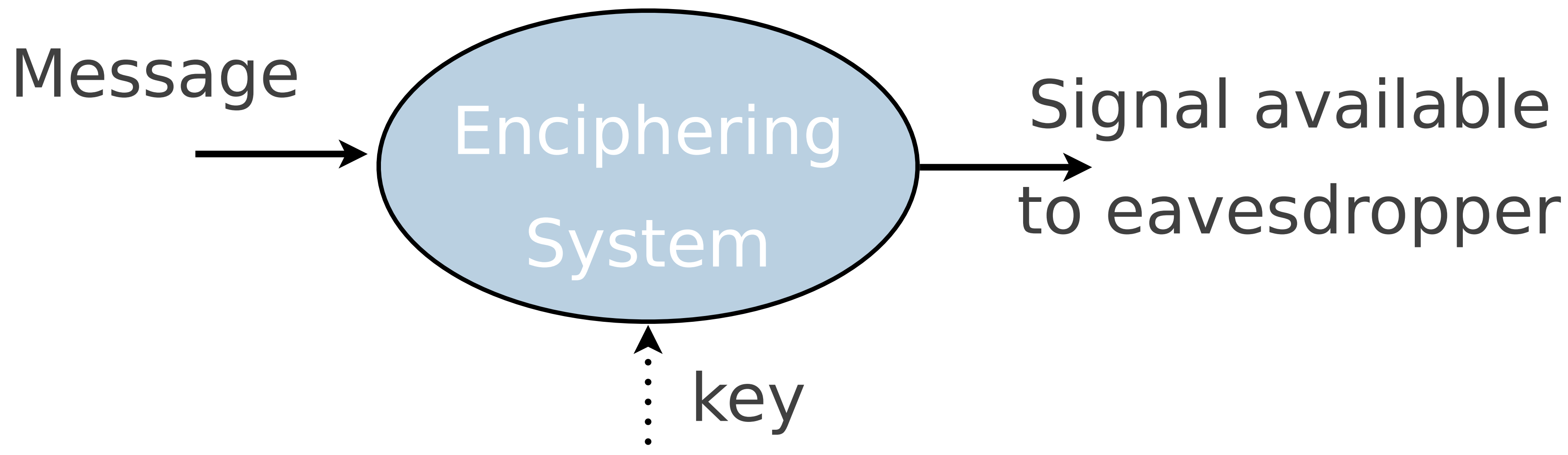
Shannon ('49):

From the point of view of the cryptanalyst, a secrecy system is almost identical with a noisy communication system. The message (transmitted signal) is operated on by a statistical element, the enciphering system, with its statistically chosen key. The result of this operation is the cryptogram (analogous to the perturbed signal) which is available for analysis. The chief differences in the two cases are: first, that the operation of the enciphering transformation is generally of a more complex nature than the perturbing noise in a channel; and, second, the key for a secrecy system is usually chosen from a finite set of possibilities while the noise in a channel is more often continually introduced, in effect chosen from an infinite set.

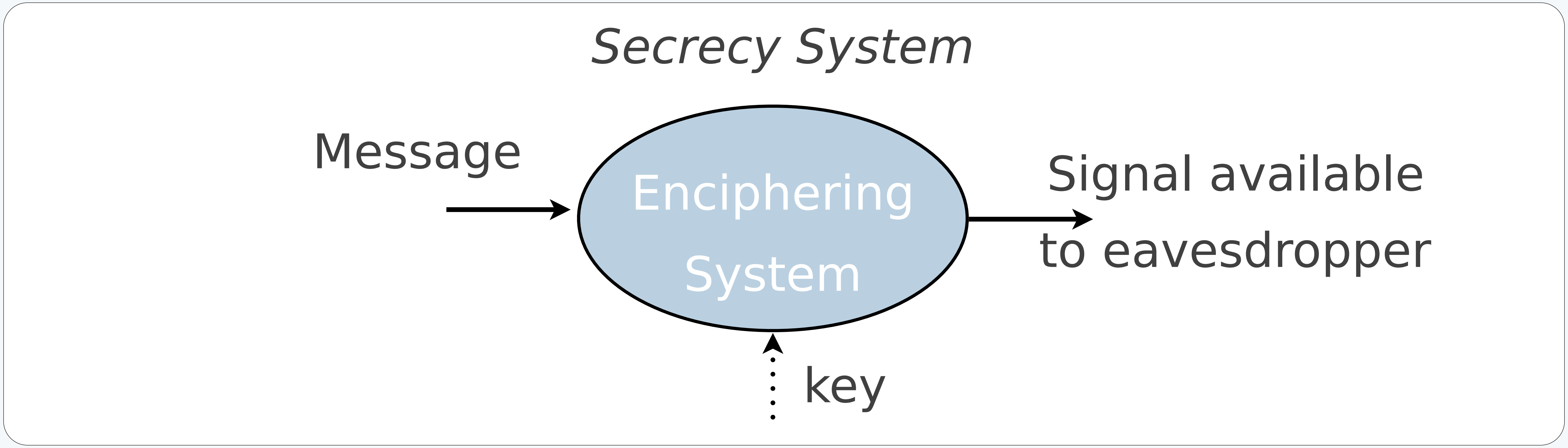
With these considerations in mind it is natural to use the equivocation as a theoretical secrecy index. It may be noted that there are two significant equivocations, that of the key and that of the message. These will be

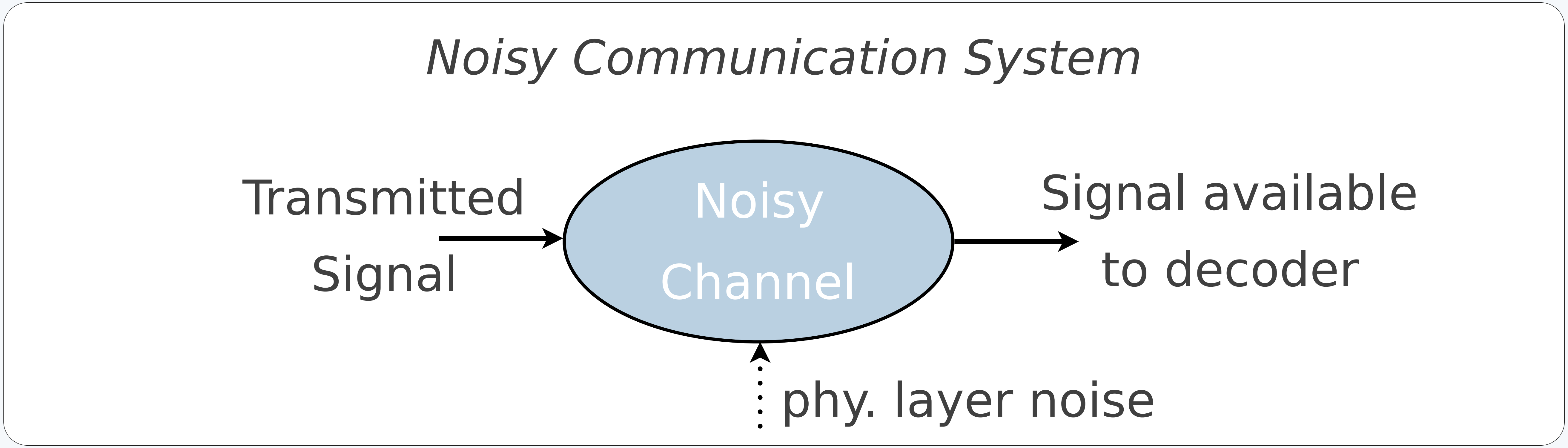
Shannon ('49)

Secrecy System

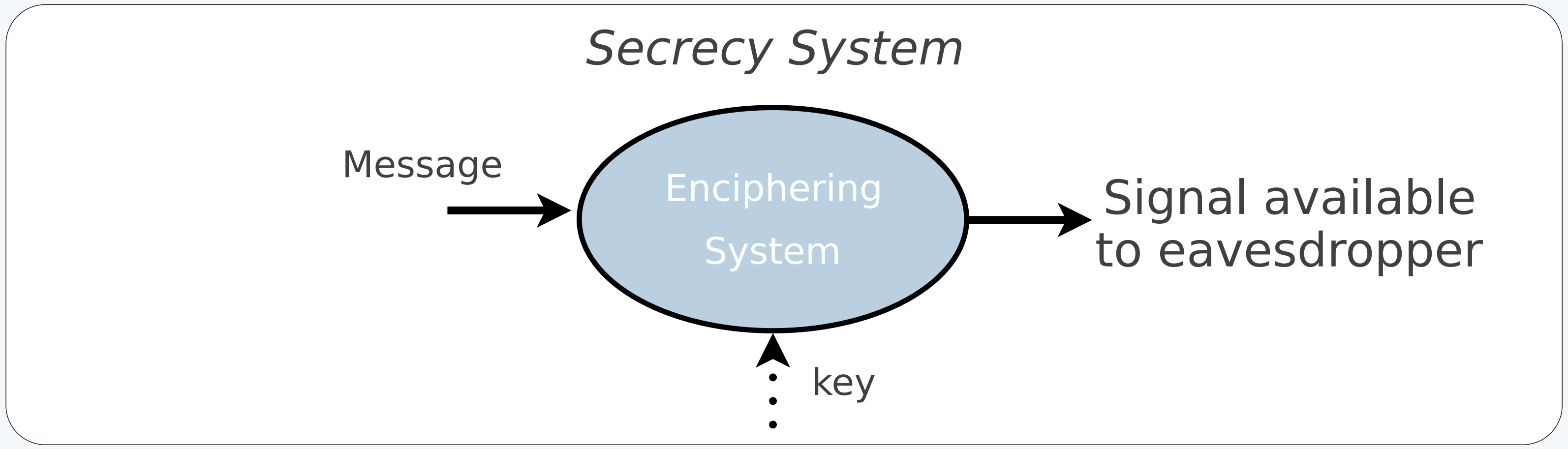


Shannon ('49)

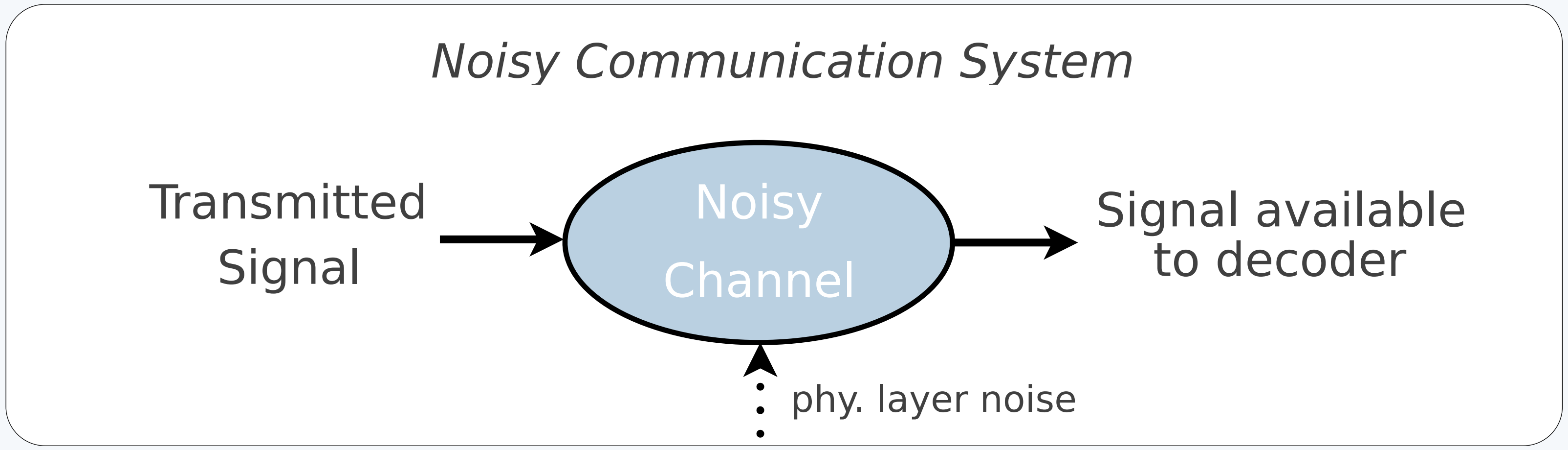

Secrecy System
Message
Enciphering
System
Signal available
to eavesdropper
key


Noisy Communication System
Transmitted
Signal
Noisy
Channel
Signal available
to decoder
phy. layer noise

Shannon ('49)

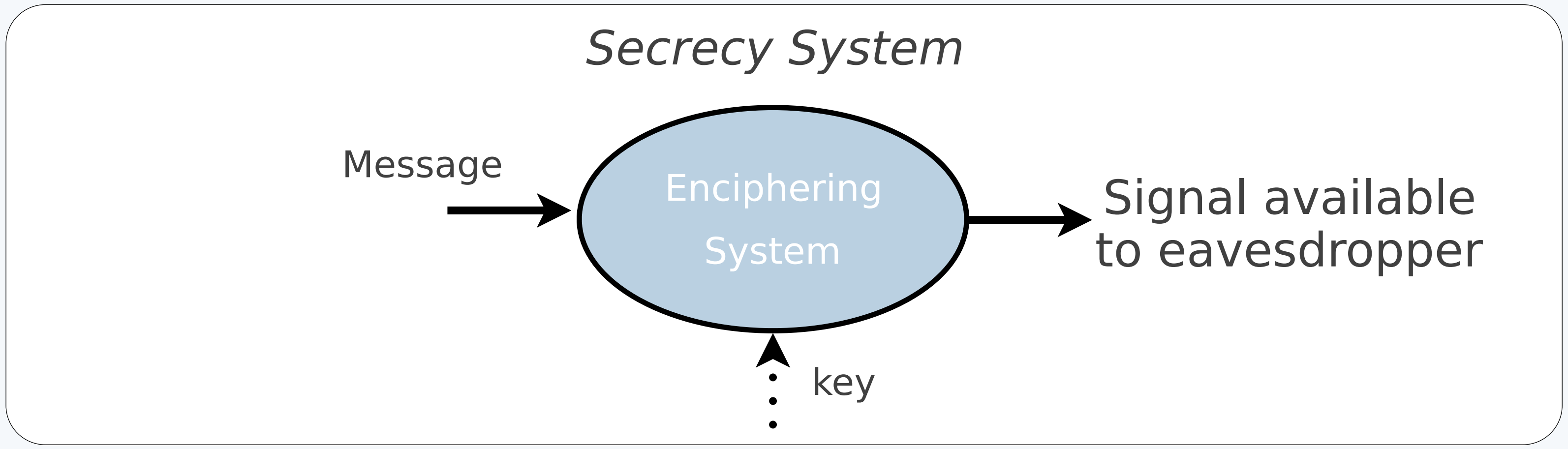


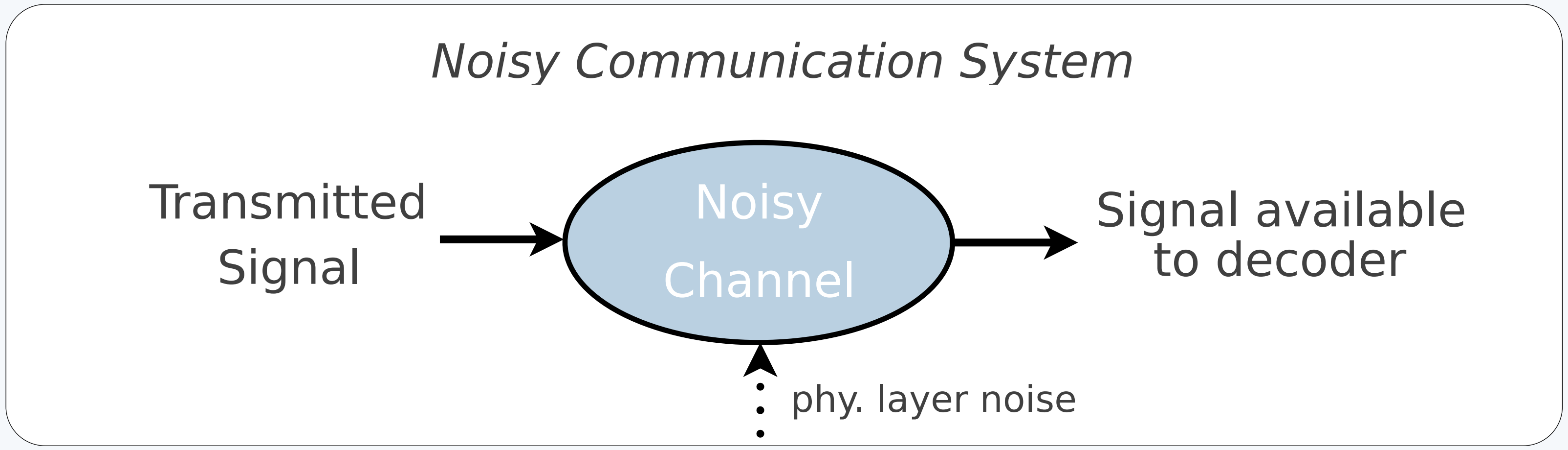
Secrecy System
Message
Enciphering
System
Signal available
to eavesdropper
key



Noisy Communication System
Transmitted
Signal
Noisy
Channel
Signal available
to decoder
phy. layer noise

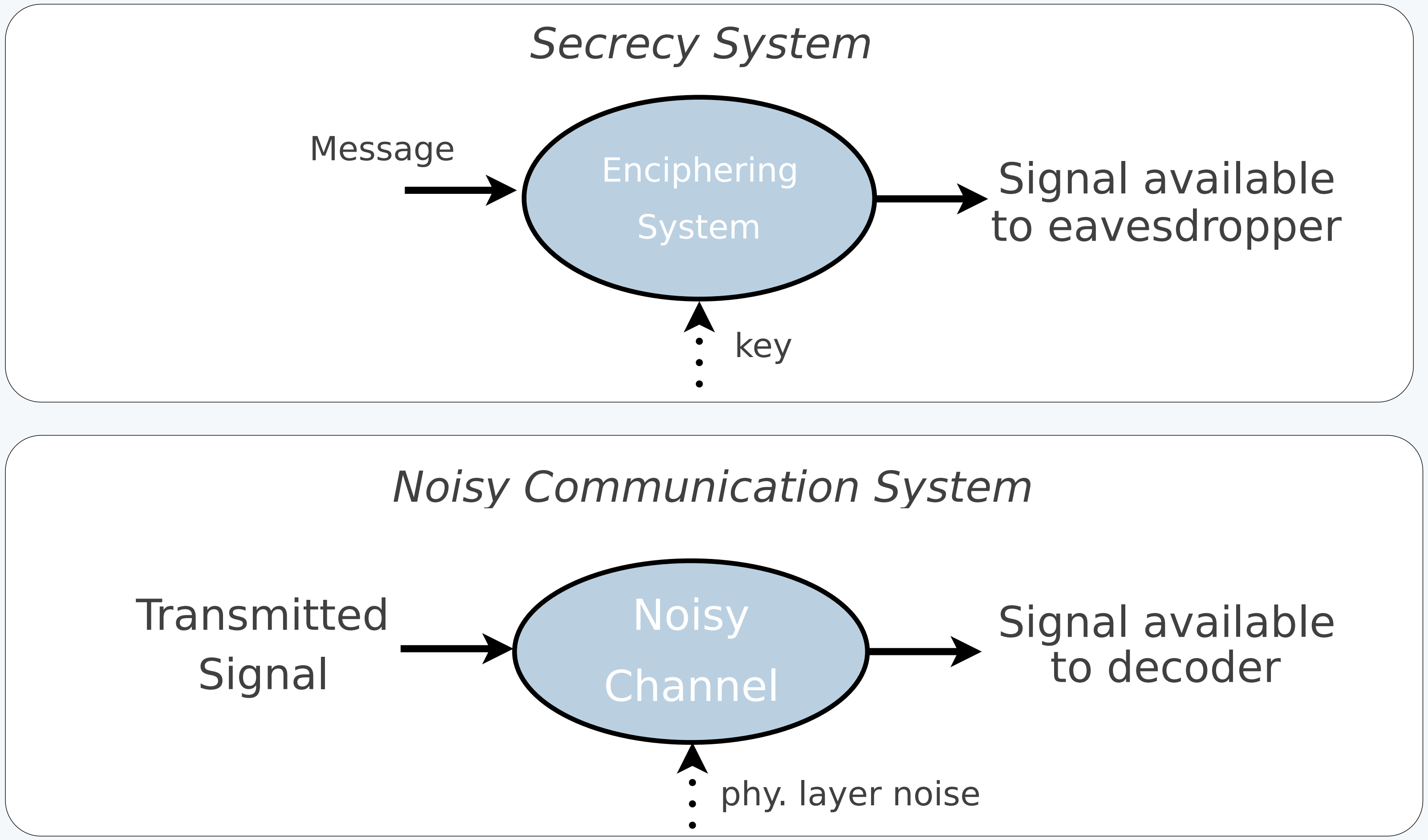
Shannon ('49)





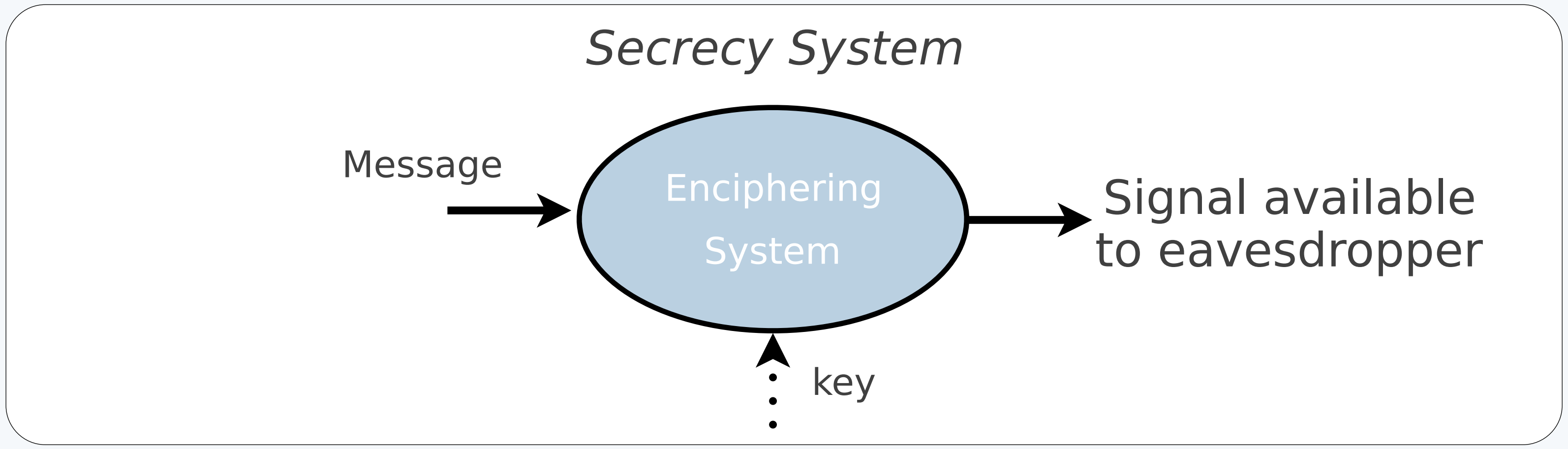
- "Chief" differences: in secrecy system:

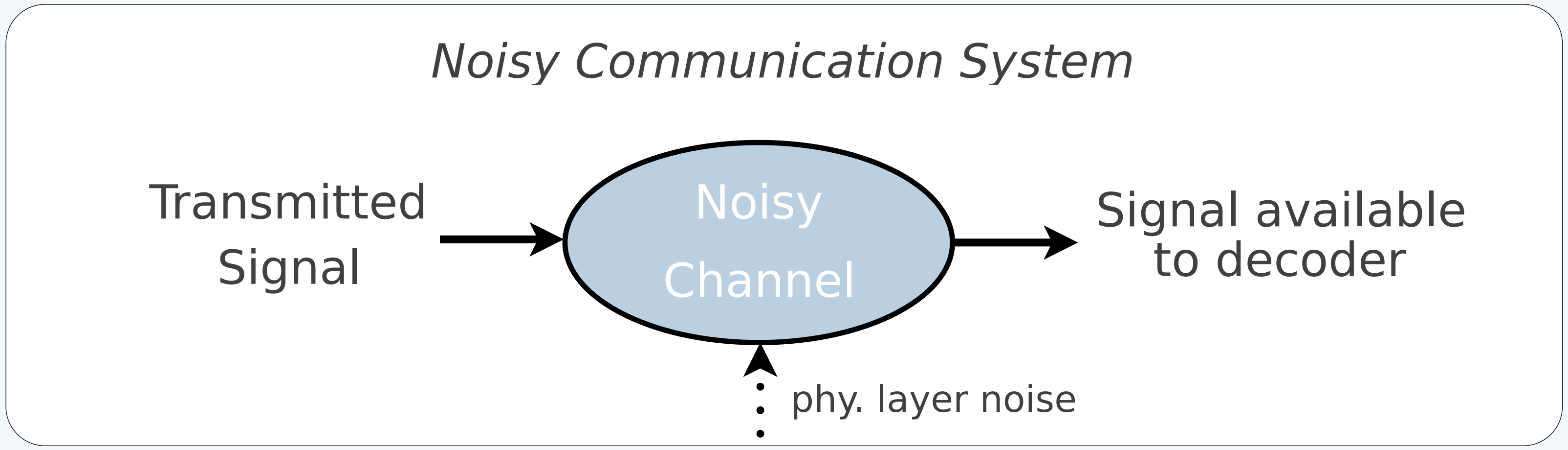
Shannon ('49)



- "Chief" differences: in secrecy system:
  - Injected randomness is of "more complex nature"

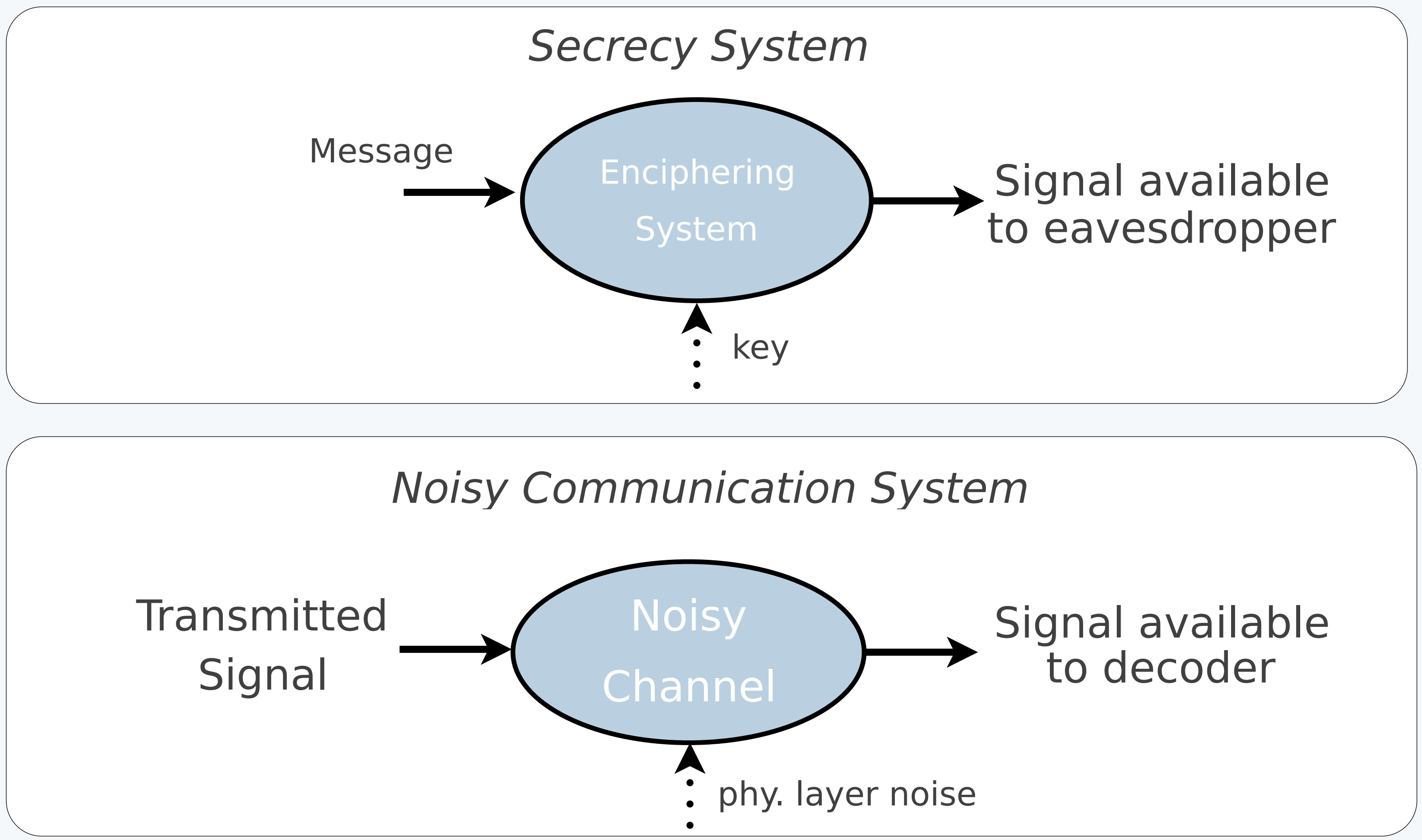
Shannon ('49)





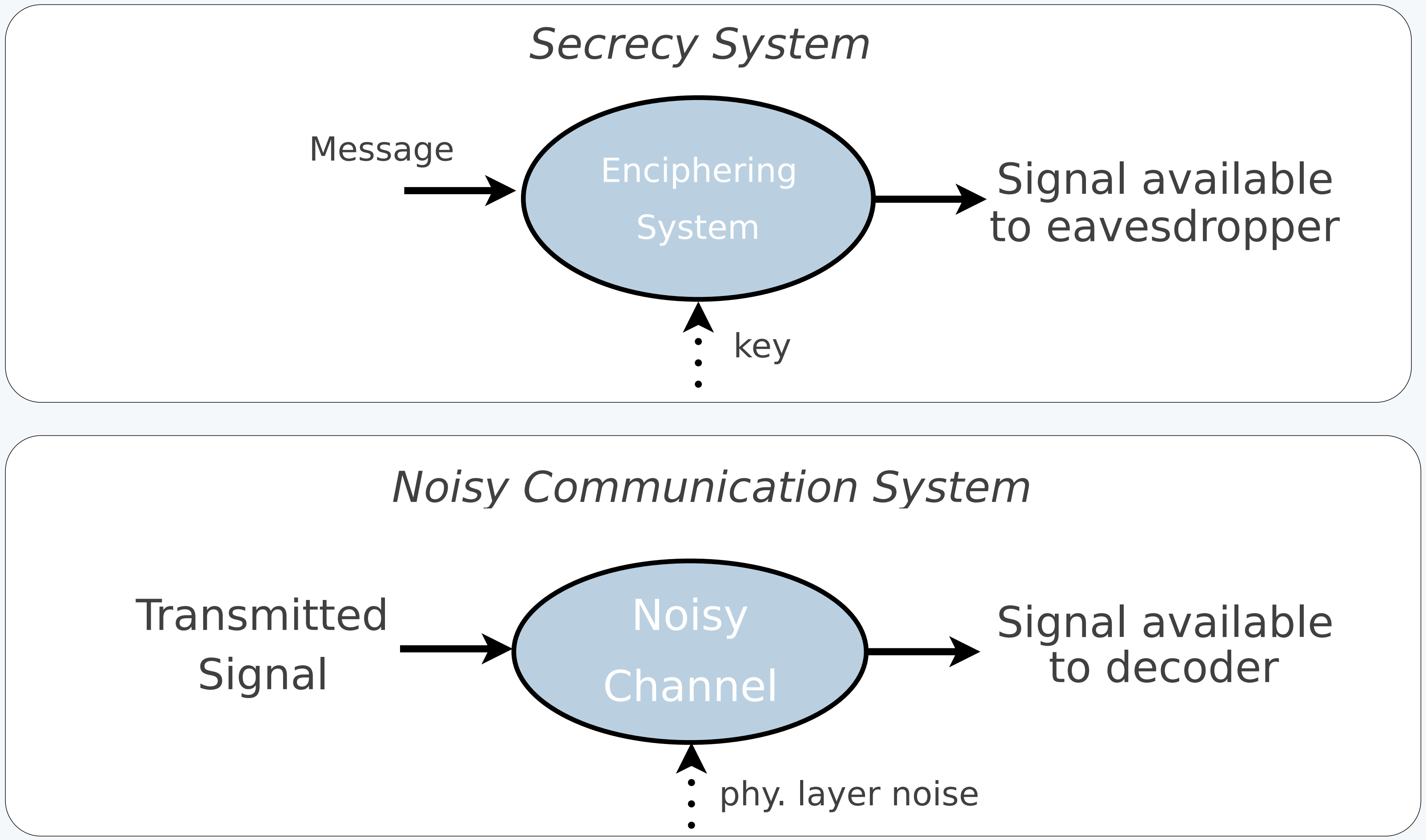
- "Chief" differences: in secrecy system:
  - Injected randomness is of "more complex nature"
  - Injected randomness is discrete

Shannon ('49)



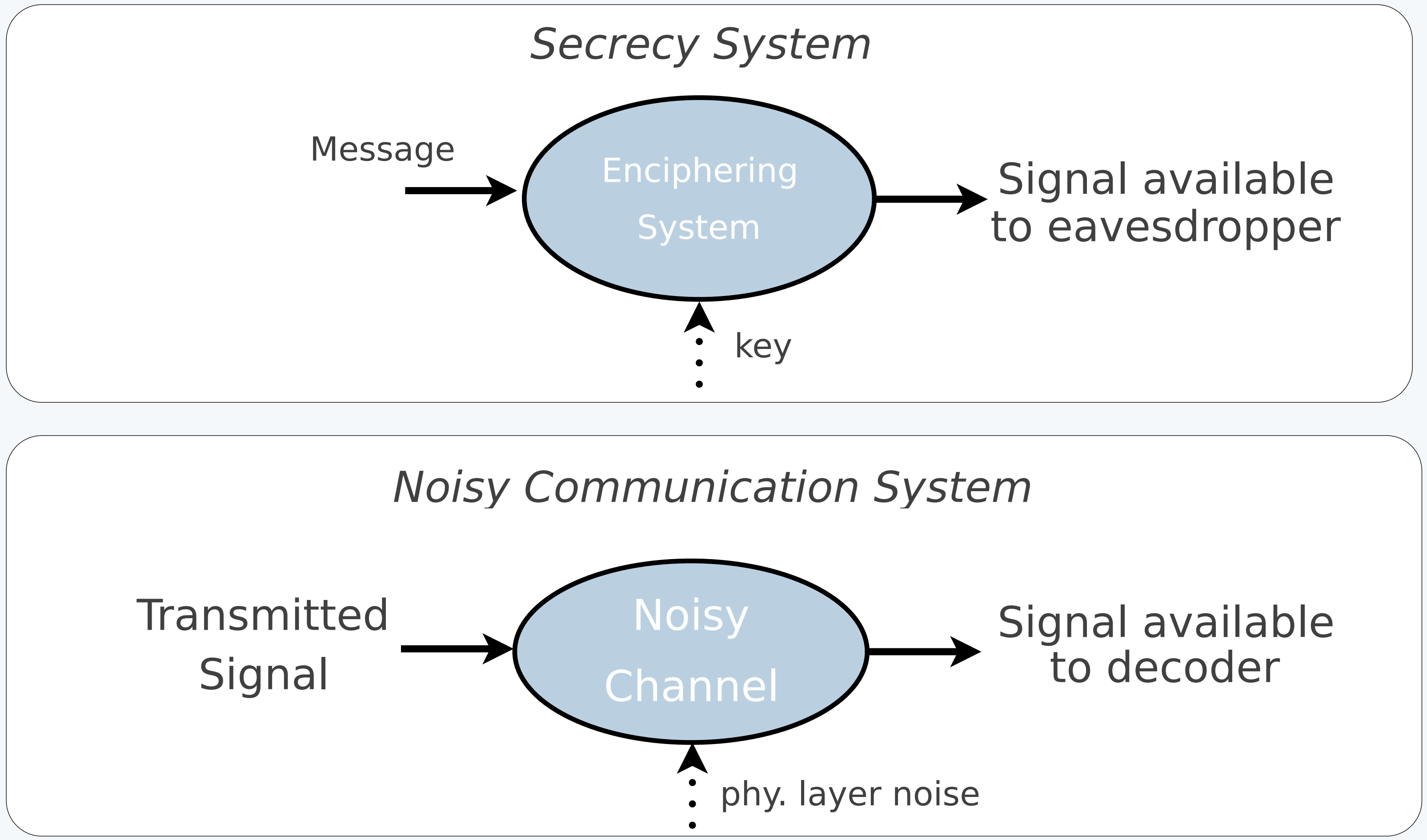
- Other differences: in conventional comm.,

Shannon ('49)



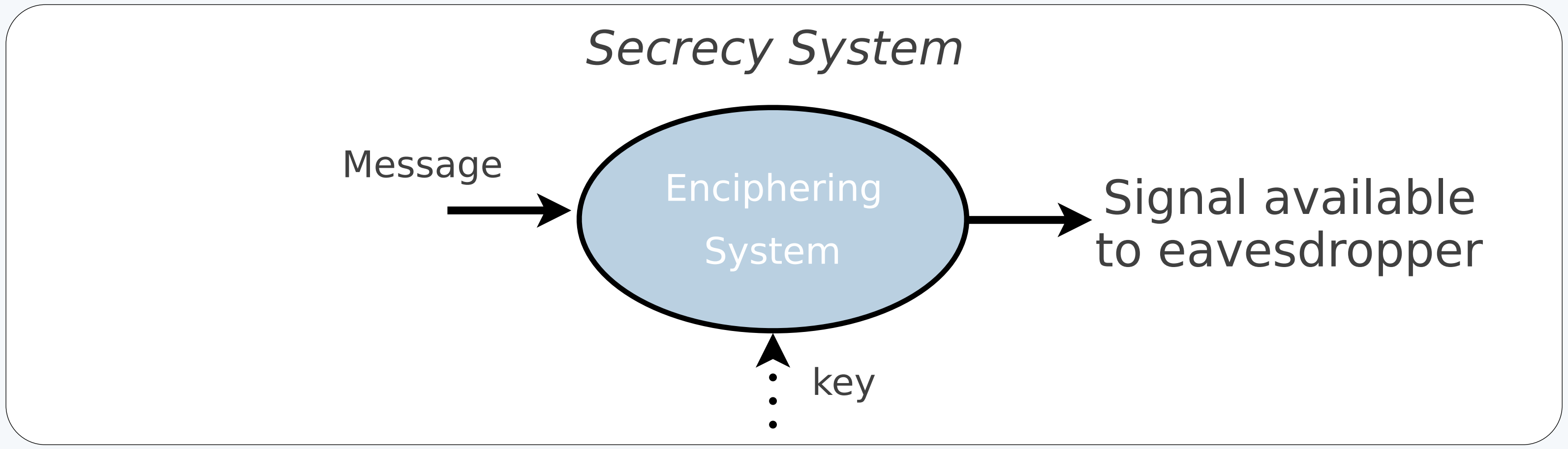
- Other differences: in conventional comm.,
  - Encoder is a willing participant (coding)

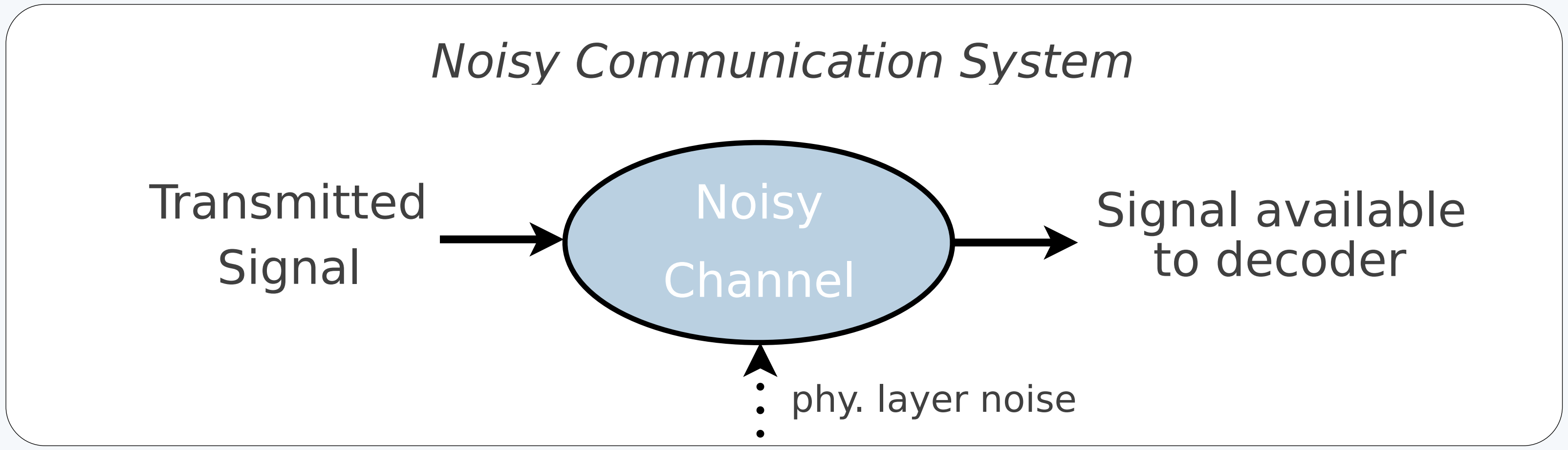
Shannon ('49)



- Other differences: in conventional comm.,
  - Encoder is a willing participant (coding)
  - Communication must be reliable

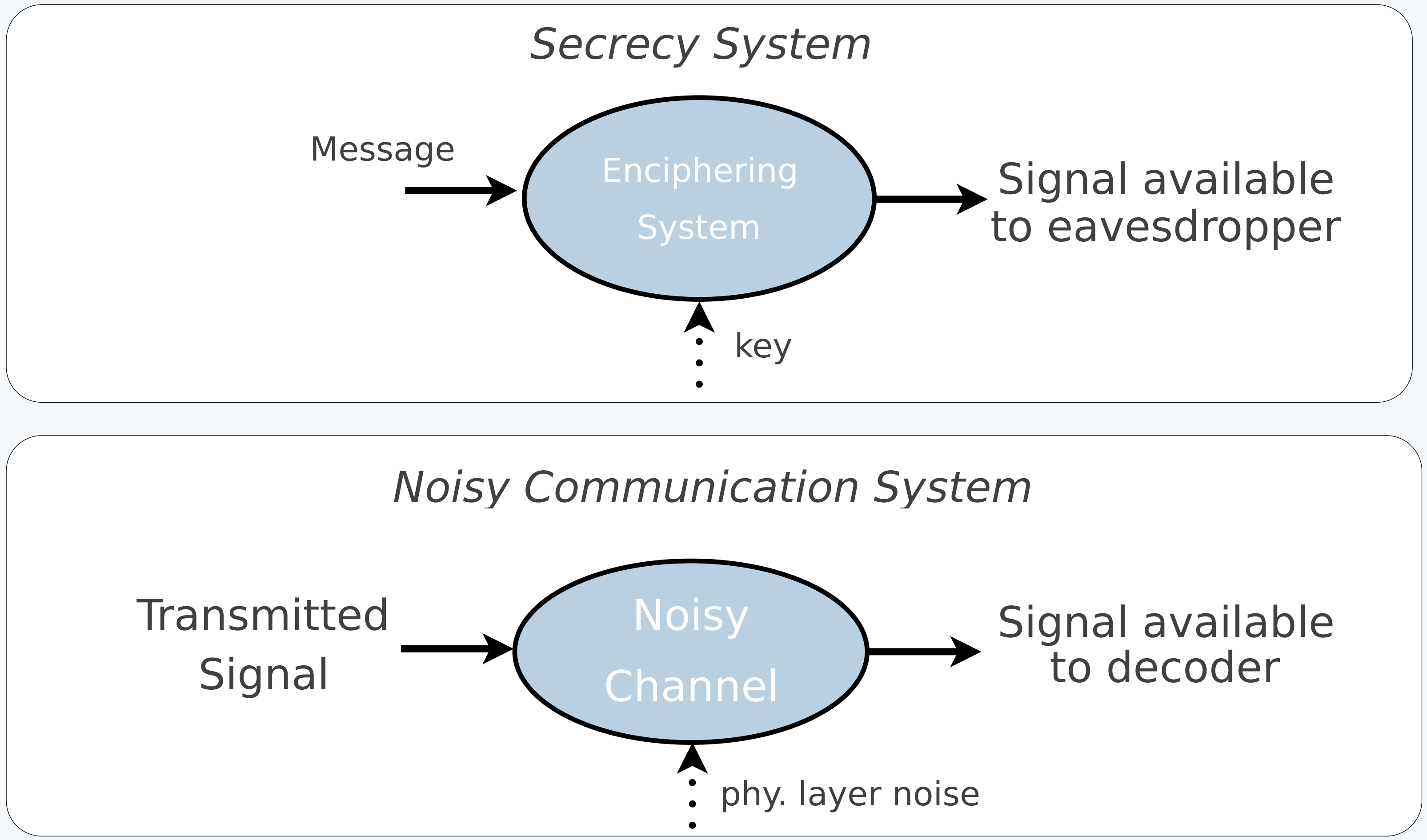
Shannon ('49)





- Other differences: in conventional comm.,
  - Encoder is a willing participant (coding)
  - Communication must be reliable
- Unclear motivation for using MI in secrecy applications

Shannon ('49)



- Other differences: in conventional comm.,
  - Encoder is a willing participant (coding)
  - Communication must be reliable
- Unclear motivation for using MI in secrecy applications
- But isn't capacity an upper bound?

But is Capacity an Upper Bound?

Folk Theorem: Any reasonable measure of "leakage" from X to Y should be upper bounded by the Shannon capacity of the channel $P_{Y|X}$:

$$\text{"}\mathcal{L}(X \to Y)\text{"} \leq C = \max_{p(x)} I(X;Y).$$

But is Capacity an Upper Bound?

Folk Theorem: Any reasonable measure of "leakage" from X to Y should be upper bounded by the Shannon capacity of the channel $P_{Y|X}$:

$$\text{"}\mathcal{L}(X \to Y)\text{"} \leq C = \max_{p(x)} I(X;Y).$$

"Proof:"

But is Capacity an Upper Bound?

Folk Theorem: Any reasonable measure of "leakage" from X to Y should be upper bounded by the Shannon capacity of the channel $P_{Y|X}$:

$$\text{"}\mathcal{L}(X \to Y)\text{"} \le C = \max_{p(x)} I(X;Y).$$

"<u>Proof:</u>" $$\text{"}\mathcal{L}(X \to Y)\text{"} \le \max_{P_X} \text{"}\mathcal{L}(X \to Y)\text{"}$$

But is Capacity an Upper Bound?

Folk Theorem: Any reasonable measure of "leakage" from X to Y should be upper bounded by the Shannon capacity of the channel $P_{Y|X}$:

$$\text{"}\mathcal{L}(X \to Y)\text{"} \leq C = \max_{p(x)} I(X;Y).$$

"<u>Proof:</u>"

$$\begin{aligned} \text{"}\mathcal{L}(X \to Y)\text{"} &\leq \max_{P_X} \text{"}\mathcal{L}(X \to Y)\text{"} \\ &\leq \lim_{n\to\infty} \max_{P_{X^n}} \frac{1}{n} \text{"}\mathcal{L}(X^n \to Y^n)\text{"} \end{aligned}$$

But is Capacity an Upper Bound?

Folk Theorem: Any reasonable measure of "leakage" from X to Y should be upper bounded by the Shannon capacity of the channel $P_{Y|X}$:

$$\text{"}\mathcal{L}(X \to Y)\text{"} \leq C = \max_{p(x)} I(X;Y).$$

"<u>Proof:</u>"

$$\begin{aligned}\text{"}\mathcal{L}(X \to Y)\text{"} &\leq \max_{P_X} \text{"}\mathcal{L}(X \to Y)\text{"} \\ &\leq \lim_{n\to\infty} \max_{P_{X^n}} \frac{1}{n} \text{"}\mathcal{L}(X^n \to Y^n)\text{"} \\ &\leq C = \max_{p(x)} I(X;Y).\end{aligned}$$

But is Capacity an Upper Bound?

Folk Theorem: Any reasonable measure of "leakage" from X to Y should be upper bounded by the Shannon capacity of the channel $P_{Y|X}$:

$$\text{"}\mathcal{L}(X \to Y)\text{"} \leq C = \max_{p(x)} I(X;Y).$$

"Proof:"

$$\begin{aligned}\text{"}\mathcal{L}(X \to Y)\text{"} &\leq \max_{P_X} \text{"}\mathcal{L}(X \to Y)\text{"} \\ &\leq \lim_{n\to\infty} \max_{P_{X^n}} \frac{1}{n} \text{"}\mathcal{L}(X^n \to Y^n)\text{"} \\ &\leq C = \max_{p(x)} I(X;Y).\end{aligned}$$

C is the maximum amortized rate of information transfer over a channel.

But is Capacity an Upper Bound?

Folk Theorem: Any reasonable measure of "leakage" from X to Y should be upper bounded by the Shannon capacity of the channel $P_{Y|X}$:

$$\text{"}\mathcal{L}(X \to Y)\text{"} \leq C = \max_{p(x)} I(X;Y).$$

"Proof:"

$$\begin{aligned}\text{"}\mathcal{L}(X \to Y)\text{"} &\leq \max_{P_X} \text{"}\mathcal{L}(X \to Y)\text{"} \\ &\leq \lim_{n\to\infty} \max_{P_{X^n}} \frac{1}{n} \text{"}\mathcal{L}(X^n \to Y^n)\text{"} \\ &\leq C = \max_{p(x)} I(X;Y).\end{aligned}$$

C is the maximum amortized rate of information transfer over a channel.

[Yet $\mathcal{L}(X \to Y) > C$]

But is Capacity an Upper Bound?

Folk Theorem: Any reasonable measure of "leakage" from X to Y should be upper bounded by the Shannon capacity of the channel $P_{Y|X}$:

$$\text{"}\mathcal{L}(X \to Y)\text{"} \leq C = \max_{p(x)} I(X;Y).$$

"<u>Proof:</u>"

$$\text{"}\mathcal{L}(X \to Y)\text{"} \leq \max_{P_X} \text{"}\mathcal{L}(X \to Y)\text{"}$$

$$\leq \lim_{n\to\infty} \max_{P_{X^n}} \frac{1}{n} \text{"}\mathcal{L}(X^n \to Y^n)\text{"}$$

$$\leq C = \max_{p(x)} I(X;Y).$$

C is the maximum amortized rate of **reliable** information transfer over a channel.

[Yet $\mathcal{L}(X \to Y) > C$]

Leakage vs. Capacity

If X has full support:

$$\mathcal{L}(X \to Y) = \sup_{U \leftrightarrow X \leftrightarrow Y} \log \frac{\sup_{\tilde{u}(\cdot)} \Pr(U = \tilde{u}(Y))}{\sup_{\tilde{u}} \Pr(U = \tilde{u})}$$

Leakage vs. Capacity

If X has full support:

$$\mathcal{L}(X \to Y) = \sup_{U \leftrightarrow X \leftrightarrow Y} \log \frac{\sup_{\tilde{u}(\cdot)} \Pr(U = \tilde{u}(Y))}{\sup_{\tilde{u}} \Pr(U = \tilde{u})}$$

$$= \sup_{U \leftrightarrow X \leftrightarrow Y \leftrightarrow \tilde{U}} \log \frac{\Pr(U = \tilde{U})}{\sup_{\tilde{u}} \Pr(U = \tilde{u})}$$

Leakage vs. Capacity

If X has full support:

$$\mathcal{L}(X \to Y) = \sup_{U \leftrightarrow X \leftrightarrow Y} \log \frac{\sup_{\tilde{u}(\cdot)} \Pr(U = \tilde{u}(Y))}{\sup_{\tilde{u}} \Pr(U = \tilde{u})}$$

$$= \sup_{U \leftrightarrow X \leftrightarrow Y \leftrightarrow \tilde{U}} \log \frac{\Pr(U = \tilde{U})}{\sup_{\tilde{u}} \Pr(U = \tilde{u})}$$

$$= \lim_{n \to \infty} \sup_{U \leftrightarrow X^n \leftrightarrow Y^n \leftrightarrow \tilde{U}} \frac{1}{n} \log \frac{\Pr(U = \tilde{U})}{\sup_{\tilde{u}} \Pr(U = \tilde{u})}$$

Leakage vs. Capacity

If X has full support:

$$\mathcal{L}(X \to Y) = \sup_{U \leftrightarrow X \leftrightarrow Y} \log \frac{\sup_{\tilde{u}(\cdot)} \Pr(U = \tilde{u}(Y))}{\sup_{\tilde{u}} \Pr(U = \tilde{u})}$$

$$= \sup_{U \leftrightarrow X \leftrightarrow Y \leftrightarrow \tilde{U}} \log \frac{\Pr(U = \tilde{U})}{\sup_{\tilde{u}} \Pr(U = \tilde{u})}$$

$$= \lim_{n \to \infty} \sup_{U \leftrightarrow X^n \leftrightarrow Y^n \leftrightarrow \tilde{U}} \frac{1}{n} \log \frac{\Pr(U = \tilde{U})}{\sup_{\tilde{u}} \Pr(U = \tilde{u})}$$

$$= \lim_{n \to \infty} \sup_{P_{X^n}} \sup_{U \leftrightarrow X^n \leftrightarrow Y^n \leftrightarrow \tilde{U}} \frac{1}{n} \log \frac{\Pr(U = \tilde{U})}{\sup_{\tilde{u}} \Pr(U = \tilde{u})}$$

Leakage vs. Capacity

If X has full support:

$$\mathcal{L}(X \to Y) = \lim_{n\to\infty} \sup_{P_{X^n}} \sup_{U \leftrightarrow X^n \leftrightarrow Y^n \leftrightarrow \tilde{U}} \frac{1}{n} \log \frac{\Pr(U = \tilde{U})}{\sup_{\tilde{u}} \Pr(U = \tilde{u})}$$

Leakage vs. Capacity

If X has full support:

$$\mathcal{L}(X \to Y) = \lim_{n\to\infty} \sup_{P_{X^n}} \sup_{U\leftrightarrow X^n \leftrightarrow Y^n \leftrightarrow \tilde{U}} \frac{1}{n} \log \frac{\Pr(U = \tilde{U})}{\sup_{\tilde{u}} \Pr(U = \tilde{u})}$$

Theorem (Issa-Wagner):

$$C = \lim_{\epsilon\to 0} \lim_{n\to\infty} \sup_{P_{X^n}} \sup_{\substack{U\leftrightarrow X^n \leftrightarrow Y^n \to \tilde{U}: \\ P(U=\tilde{U})\geq 1-\epsilon}} \frac{1}{n} \log \frac{\Pr(U = \tilde{U})}{\sup_{\tilde{u}} \Pr(U = \tilde{u})}$$

(Local) Differential Privacy

$$LDP(X \to Y) := \sup_{x,x',y} \log \frac{P_{Y|X}(y|x)}{P_{Y|X}(y|x')}$$

[Warner '65;
Evfimievski *et al.* '03]

(Local) Differential Privacy

$$LDP(X \to Y) := \sup_{x,x',y} \log \frac{P_{Y|X}(y|x)}{P_{Y|X}(y|x')}$$

[Warner '65; Evfimievski *et al.* '03]

Operational interpretation?

(Local) Differential Privacy

$$LDP(X \to Y) := \sup_{x,x',y} \log \frac{P_{Y|X}(y|x)}{P_{Y|X}(y|x')}$$

[Warner ’65; Evfimievski *et al.* ’03]

Operational interpretation?

Theorem (cf. Dwork *et al.* ’06):

$$LDP(X \to Y) = \sup_{f,P_X,y} \left| \log \left(\frac{P(f(X) = 1 | Y = y)}{P(f(X) = 1)} \right) \right|$$

(Local) Differential Privacy

Theorem (cf. Dwork *et al.* '06):

$$LDP(X \to Y) = \sup_{f, P_X, y} \left| \log \left(\frac{P(f(X) = 1 | Y = y)}{P(f(X) = 1)} \right) \right|$$

(Local) Differential Privacy

Theorem (cf. Dwork *et al.* ’06):

$$LDP(X \to Y) = \sup_{f, P_X, y} \left| \log \left(\frac{P(f(X) = 1 | Y = y)}{P(f(X) = 1)} \right) \right|$$

Theorem (Issa-Wagner):

$$LDP(X \to Y) = \sup_{P_X} \sup_{U \leftrightarrow X \leftrightarrow Y} \log \frac{\sup_y \sup_{\tilde{u}} \Pr(U = \tilde{u} | Y = y)}{\sup_{\tilde{u}} \Pr(U = \tilde{u})}$$

Optimal Mechanisms

- Given $p(x)$ and $c(x,y)$, solve

$$\min_{p(y|x)} \sum_y \max_x p(y|x)$$

$$\text{subject to} \quad \sum_{x,y} p(x)p(y|x)c(x,y) \leq C$$

$$\sum_y p(y|x) = 1 \;\; \forall\, x$$

$$p(y|x) \geq 0 \;\; \forall\, x, y$$

Optimal Mechanisms

- Given $p(x)$ and $c(x,y)$, sol

“exp-leakage”

$$\min_{p(y|x)} \sum_y \max_x p(y|x)$$

$$\text{subject to} \quad \sum_{x,y} p(x)p(y|x)c(x,y) \leq C$$

$$\sum_y p(y|x) = 1 \ \ \forall\, x$$

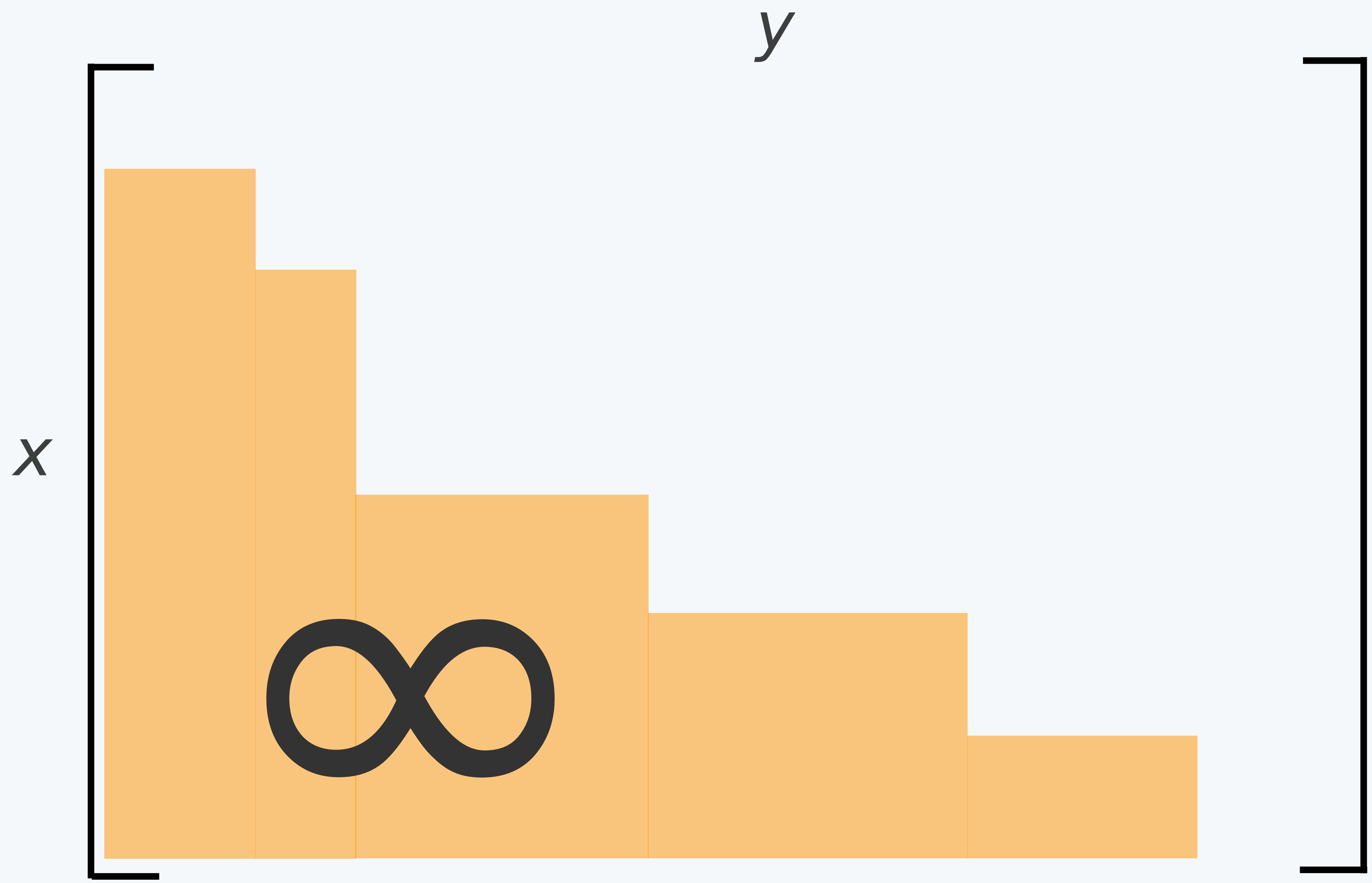
$$p(y|x) \geq 0 \ \ \forall\, x, y$$

Formulation as an LP

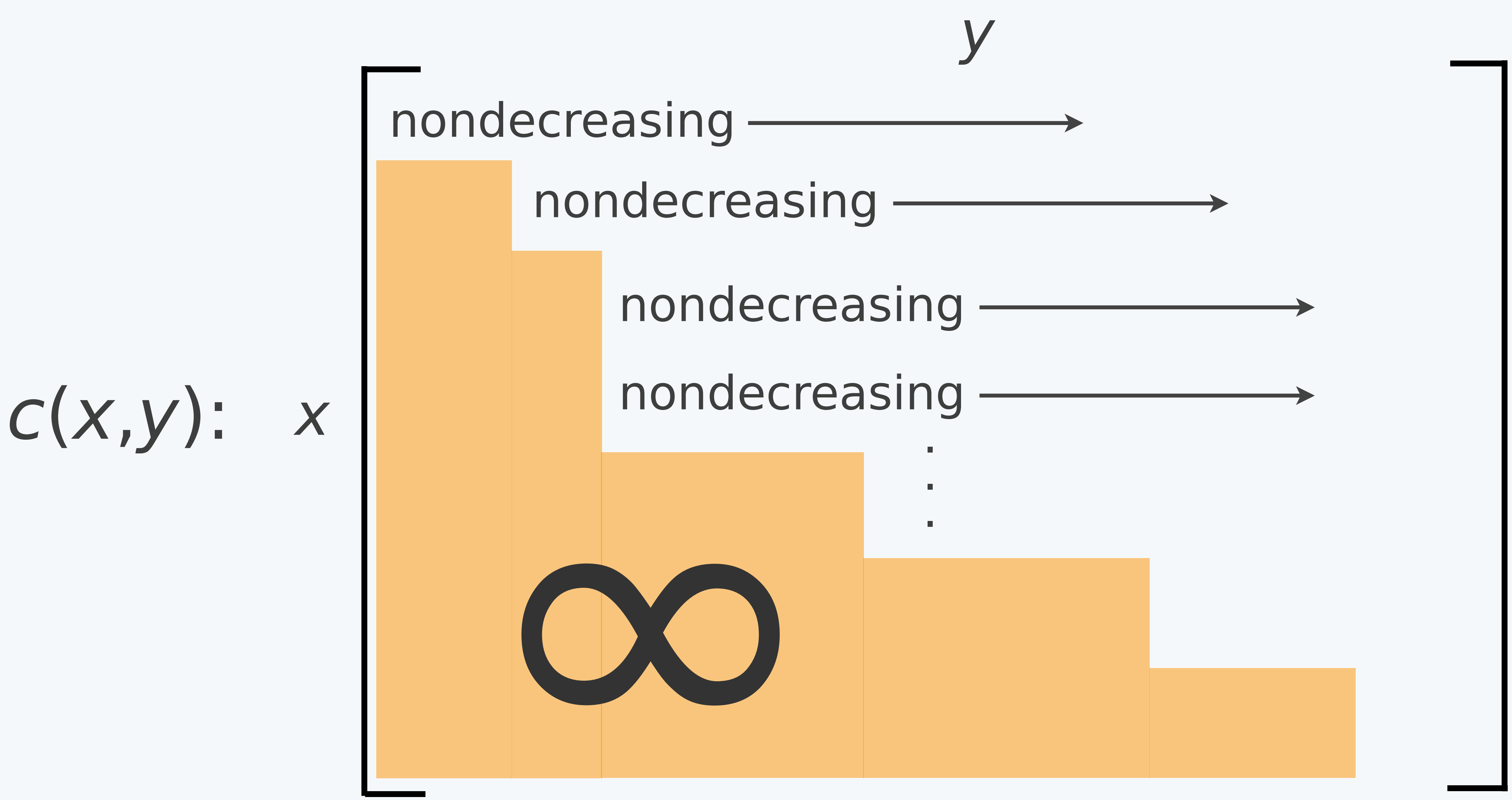
$$
\begin{aligned}
\min_{p(y|x),q_y} \quad & \sum_y q_y \\
\text{subject to} \quad & \sum_{x,y} p(x)p(y|x)c(x,y) \le C \\
& \sum_y p(y|x) = 1 \ \ \forall\, x \\
& p(y|x) \ge 0 \ \ \forall\, x, y \\
& p(y|x) \le q_y \ \ \forall\, x, y
\end{aligned}
$$

A Structural Assumption

$c(x,y)$:

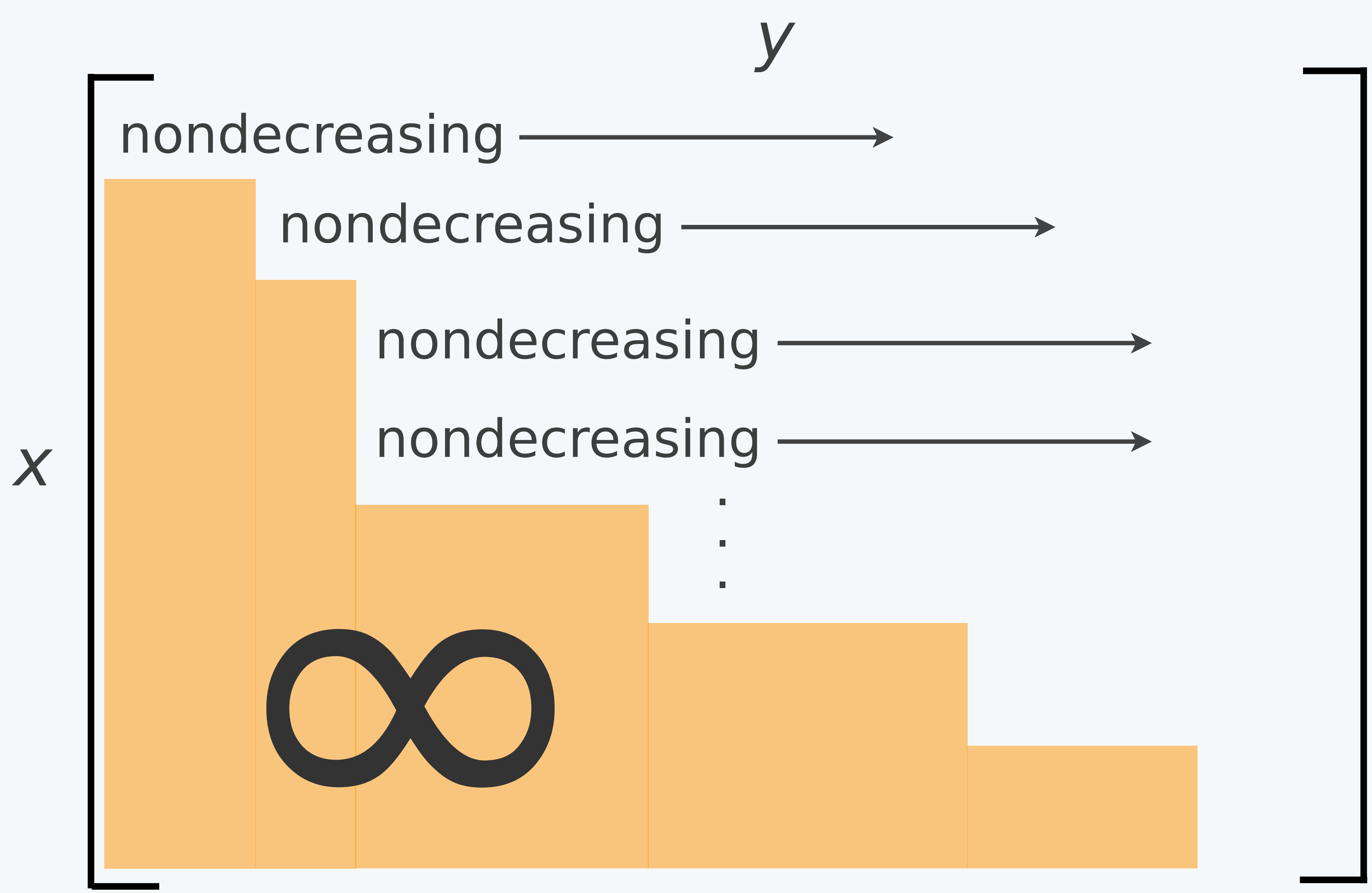


A Structural Assumption



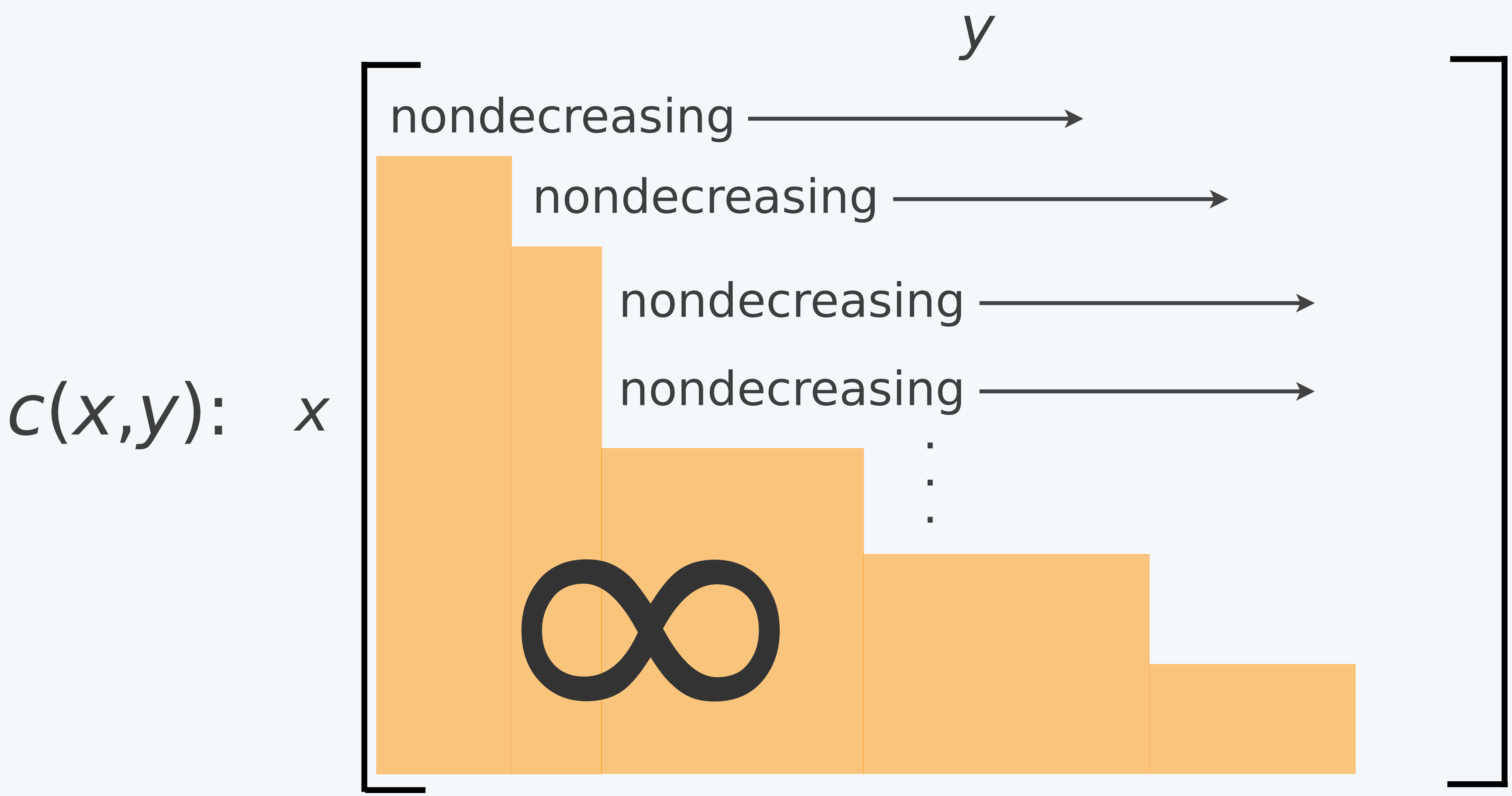
A Structural Assumption

$c(x,y)$:



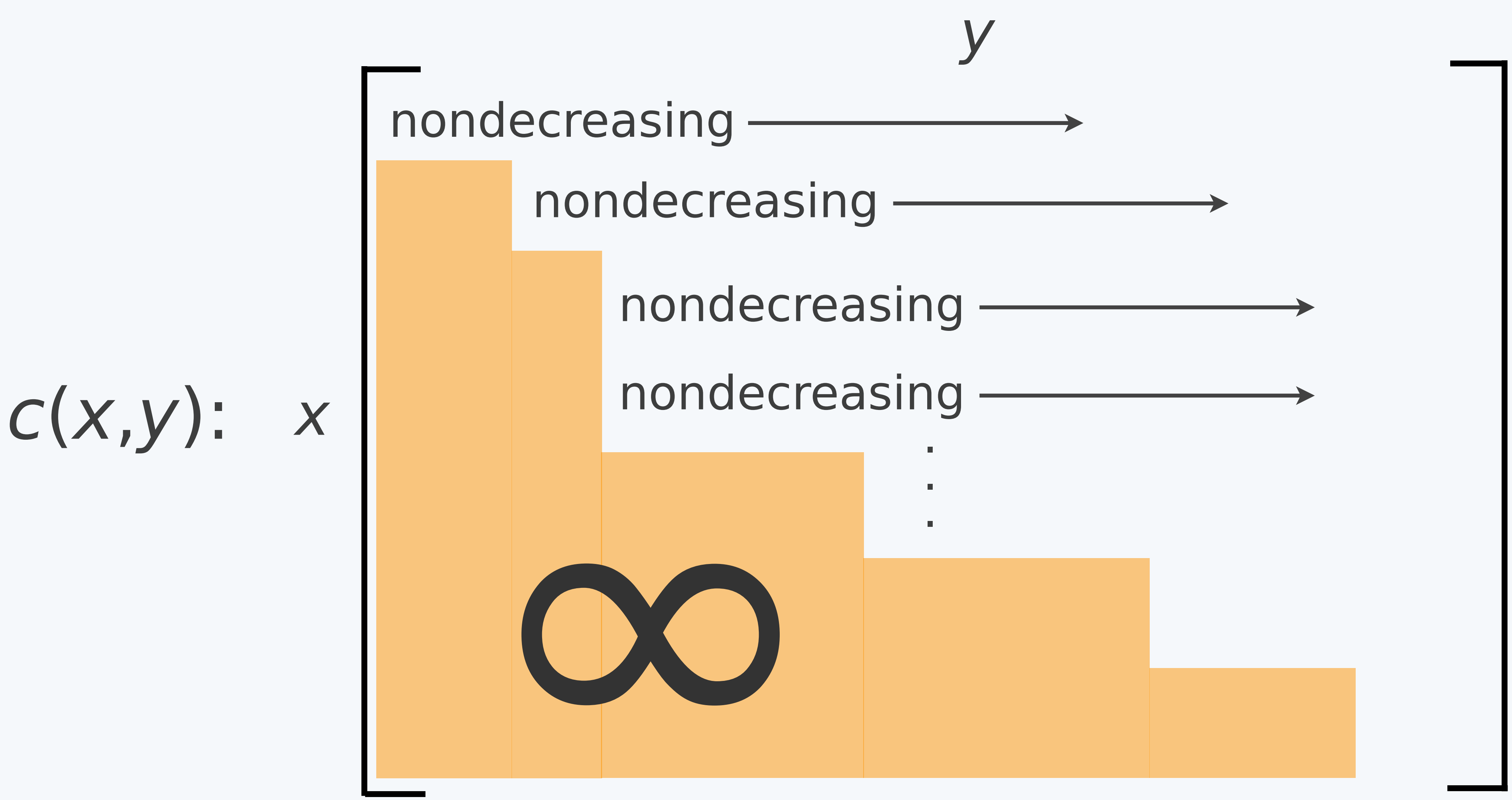
- Examples:

A Structural Assumption



- Examples:
  - Execution time [RSA], power consumption

A Structural Assumption



- Examples:
  - Execution time [RSA], power consumption
  - “Staircase increasing”

Deterministic Mechanisms Are Optimal

Deterministic Mechanisms Are Optimal

Theorem (Wu, Wagner, Suh):

If $c(\cdot,\cdot)$ is staircase increasing, then for any α and P_X,

$$\sum_y \max_x P_{Y|X}(y|x) + \alpha \cdot \sum_x \sum_y P_X(x) P_{Y|X}(y|x) c(x,y)$$

is minimized by a deterministic (0-1) $P_{Y|X}$.

Deterministic Mechanisms Are Optimal

Theorem (Wu, Wagner, Suh):

If $c(\cdot,\cdot)$ is staircase increasing, then for any α and P_X,

$$\sum_y \max_x P_{Y|X}(y|x) + \alpha \cdot \sum_x \sum_y P_X(x) P_{Y|X}(y|x) c(x,y)$$

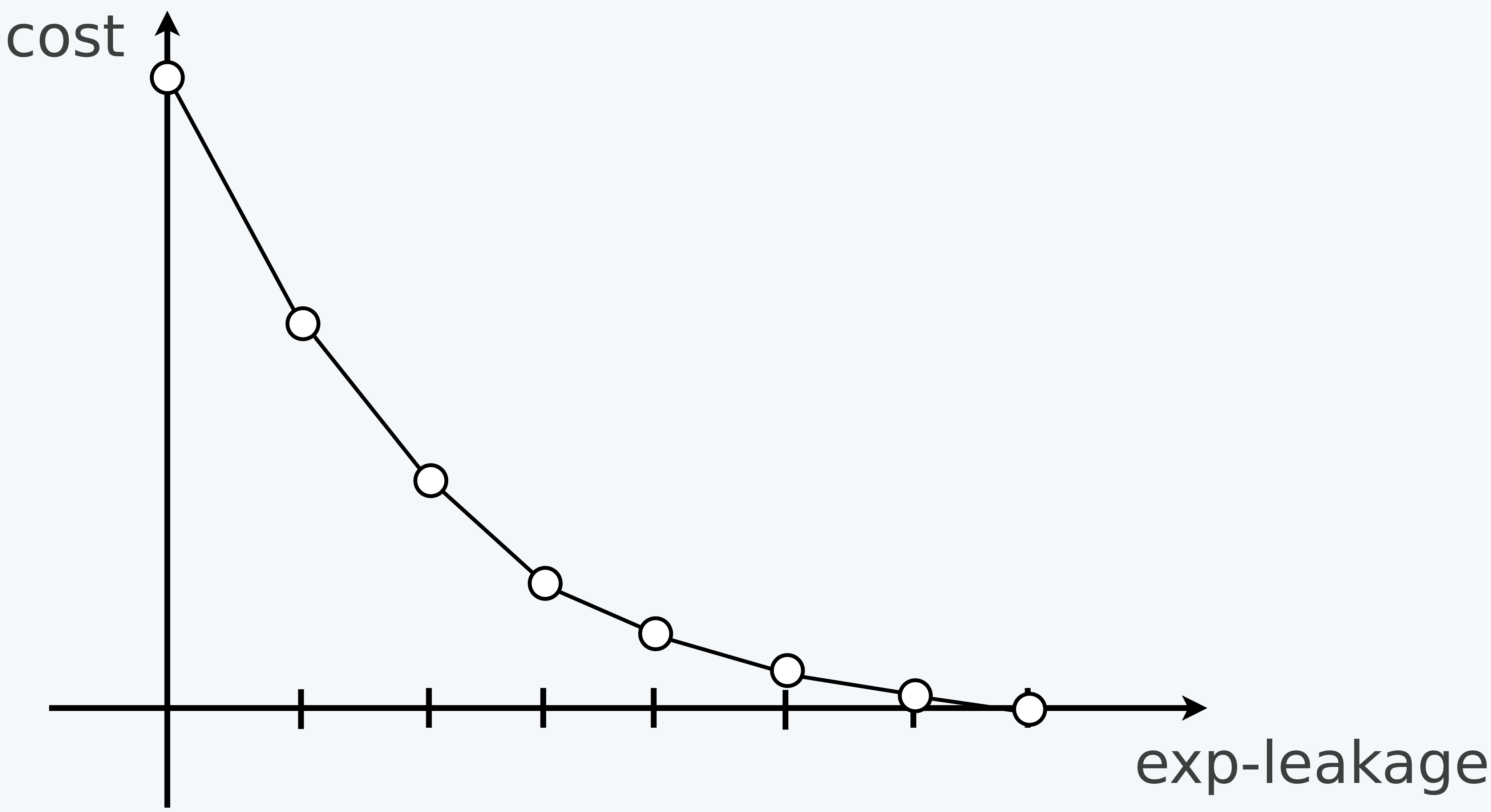
is minimized by a deterministic (0-1) $P_{Y|X}$.

Fails for the cost matrix: $\begin{bmatrix} 0 & 1 & 2 \\ 2 & 0 & 1 \\ 1 & 2 & 0 \end{bmatrix}$

Corollary

Corollary (Wu, Wagner, Suh):

The optimal cost/exp-leakage curve is piecewise linear with kink points only at integer exp-leakage values.



Advantages of Deterministic Mechanisms

Advantages of Deterministic Mechanisms

- Do not require randomness (obviously)

Advantages of Deterministic Mechanisms

- Do not require randomness (obviously)
- Easier to describe and store

Advantages of Deterministic Mechanisms

- Do not require randomness (obviously)
- Easier to describe and store
- Immune to averaging attacks

cf. Other Metrics

$$c(x,y) = \begin{bmatrix} 1 & 2 & 3 & 4 \\ \infty & 1 & 2 & 3 \\ \infty & \infty & 1 & 2 \\ \infty & \infty & \infty & 1 \end{bmatrix}$$

$p(x)$: uniform

minimize $\{E[c(X,Y])] : \text{leakage} \leq 1\}$

cf. Other Metrics

$$c(x,y) = \begin{bmatrix} 1 & 2 & 3 & 4 \\ \infty & 1 & 2 & 3 \\ \infty & \infty & 1 & 2 \\ \infty & \infty & \infty & 1 \end{bmatrix}$$

$p(x)$: uniform

minimize $\{E[c(X,Y])] : \text{leakage} \leq 1\}$

Maximal Leakage:

$$p(y|x) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

cf. Other Metrics

$$c(x,y) = \begin{bmatrix} 1 & 2 & 3 & 4 \\ \infty & 1 & 2 & 3 \\ \infty & \infty & 1 & 2 \\ \infty & \infty & \infty & 1 \end{bmatrix}$$

$p(x)$: uniform

minimize $\{E[c(X,Y])] : \text{leakage} \leq 1\}$

cf. Other Metrics

$$c(x,y) = \begin{bmatrix} 1 & 2 & 3 & 4 \\ \infty & 1 & 2 & 3 \\ \infty & \infty & 1 & 2 \\ \infty & \infty & \infty & 1 \end{bmatrix}$$

$p(x)$: uniform

$$\text{minimize } \{E[c(X,Y])] : \text{leakage} \leq 1\}$$

Mutual Information:

$$p(y|x) = \begin{bmatrix} 0.52 & 0.27 & 0.14 & 0.07 \\ 0 & 0.56 & 0.29 & 0.15 \\ 0 & 0 & 0.69 & 0.34 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

cf. Other Metrics

$$c(x,y) = \begin{bmatrix} 1 & 2 & 3 & 4 \\ \infty & 1 & 2 & 3 \\ \infty & \infty & 1 & 2 \\ \infty & \infty & \infty & 1 \end{bmatrix}$$

$p(x)$: uniform

minimize $\{E[c(X,Y])] : \text{leakage} \leq 1\}$

Mutual Information:

$$p(y|x) = \begin{bmatrix} 0.52 & 0.27 & 0.14 & 0.07 \\ 0 & 0.56 & 0.29 & 0.15 \\ 0 & 0 & 0.69 & 0.34 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Local Diff. Privacy:

$$p(y|x) = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Maximal Leakage: Other Results

Maximal Leakage: Other Results

- Shannon cipher system (Issa, Kamath, Wagner ’16)

Maximal Leakage: Other Results

- Shannon cipher system (Issa, Kamath, Wagner ’16)
- Privacy-utility tradeoffs (Liao, Sankar, Calmon, Tan, ’17)

Maximal Leakage: Other Results

- Shannon cipher system (Issa, Kamath, Wagner ’16)
- Privacy-utility tradeoffs (Liao, Sankar, Calmon, Tan, ’17)
- Sibson MI of other orders (Liao, Kosut, Sankar, Calmon, ’18)

Maximal Leakage: Other Results

- Shannon cipher system (Issa, Kamath, Wagner ’16)
- Privacy-utility tradeoffs (Liao, Sankar, Calmon, Tan, ’17)
- Sibson MI of other orders (Liao, Kosut, Sankar, Calmon, ’18)
- Learning ML from trace data (Issa and Wagner, ’18)

Three Takeaways

Def (Issa-Kamath-Wagner)**:** Given P_{XY}, the *maximal leakage* from X to Y is

$$\mathcal{L}(X \to Y) = \sup_{U : U \leftrightarrow X \leftrightarrow Y} \log \frac{\sup_{\tilde{u}(\cdot)} \Pr(U = \tilde{u}(Y))}{\sup_{\tilde{u}} \Pr(U = \tilde{u})}$$

Three Takeaways

Def (Issa-Kamath-Wagner)**:** Given P_{XY}, the *maximal leakage* from X to Y is

$$\mathcal{L}(X \to Y) = \sup_{U: U \leftrightarrow X \leftrightarrow Y} \log \frac{\sup_{\tilde{u}(\cdot)} \Pr(U = \tilde{u}(Y))}{\sup_{\tilde{u}} \Pr(U = \tilde{u})}$$

Maximal leakage ...

Three Takeaways

Def (Issa-Kamath-Wagner)**:** Given P_{XY}, the *maximal leakage* from X to Y is

$$\mathcal{L}(X \to Y) = \sup_{U: U \leftrightarrow X \leftrightarrow Y} \log \frac{\sup_{\tilde{u}(\cdot)} \Pr(U = \tilde{u}(Y))}{\sup_{\tilde{u}} \Pr(U = \tilde{u})}$$

Maximal leakage ...

1. ... captures the increase in guessing probability of secrets

Three Takeaways

Def (Issa-Kamath-Wagner)**:** Given P_{XY}, the *maximal leakage* from X to Y is

$$\mathcal{L}(X \to Y) = \sup_{U: U \leftrightarrow X \leftrightarrow Y} \log \frac{\sup_{\tilde{u}(\cdot)} \Pr(U = \tilde{u}(Y))}{\sup_{\tilde{u}} \Pr(U = \tilde{u})}$$

Maximal leakage ...

1. ... captures the increase in guessing probability of secrets

   ... is well suited for side channels with keys, passwords.

Three Takeaways

Def (Issa-Kamath-Wagner)**:** Given P_{XY}, the *maximal leakage* from X to Y is

$$\mathcal{L}(X \to Y) = \sup_{U: U \leftrightarrow X \leftrightarrow Y} \log \frac{\sup_{\tilde{u}(\cdot)} \Pr(U = \tilde{u}(Y))}{\sup_{\tilde{u}} \Pr(U = \tilde{u})}$$

Maximal leakage ...

1. ... captures the increase in guessing probability of secrets

   ... is well suited for side channels with keys, passwords.

2. ... is robust to modeling assumptions

Three Takeaways

Def (Issa-Kamath-Wagner)**:** Given P_{XY}, the *maximal leakage* from X to Y is

$$\mathcal{L}(X \to Y) = \sup_{U:U\longleftrightarrow X\longleftrightarrow Y} \log \frac{\sup_{\tilde{u}(\cdot)} \Pr(U = \tilde{u}(Y))}{\sup_{\tilde{u}} \Pr(U = \tilde{u})}$$

Maximal leakage ...

1. ... captures the increase in guessing probability of secrets

   ... is well suited for side channels with keys, passwords.

2. ... is robust to modeling assumptions
3. ... favors deterministic mechanisms (quantization) over "adding noise" in many contexts.

Extra Slides

A Different Question

A Different Question

How many secrecy measures do we need?

A Different Question

How many secrecy measures do we need?

- Probably more than one …

A Different Question

How many secrecy measures do we need?

- Probably more than one …
    - ML ill-suited for e.g., medical databases

A Different Question

How many secrecy measures do we need?

- Probably more than one ...
  - ML ill-suited for e.g., medical databases
  - DP ill-suited for side channels

A Different Question

How many secrecy measures do we need?

- Probably more than one ...
  - ML ill-suited for e.g., medical databases
  - DP ill-suited for side channels
  - Both ML and DP ill-suited for computationally-bounded eavesdroppers

A Different Question

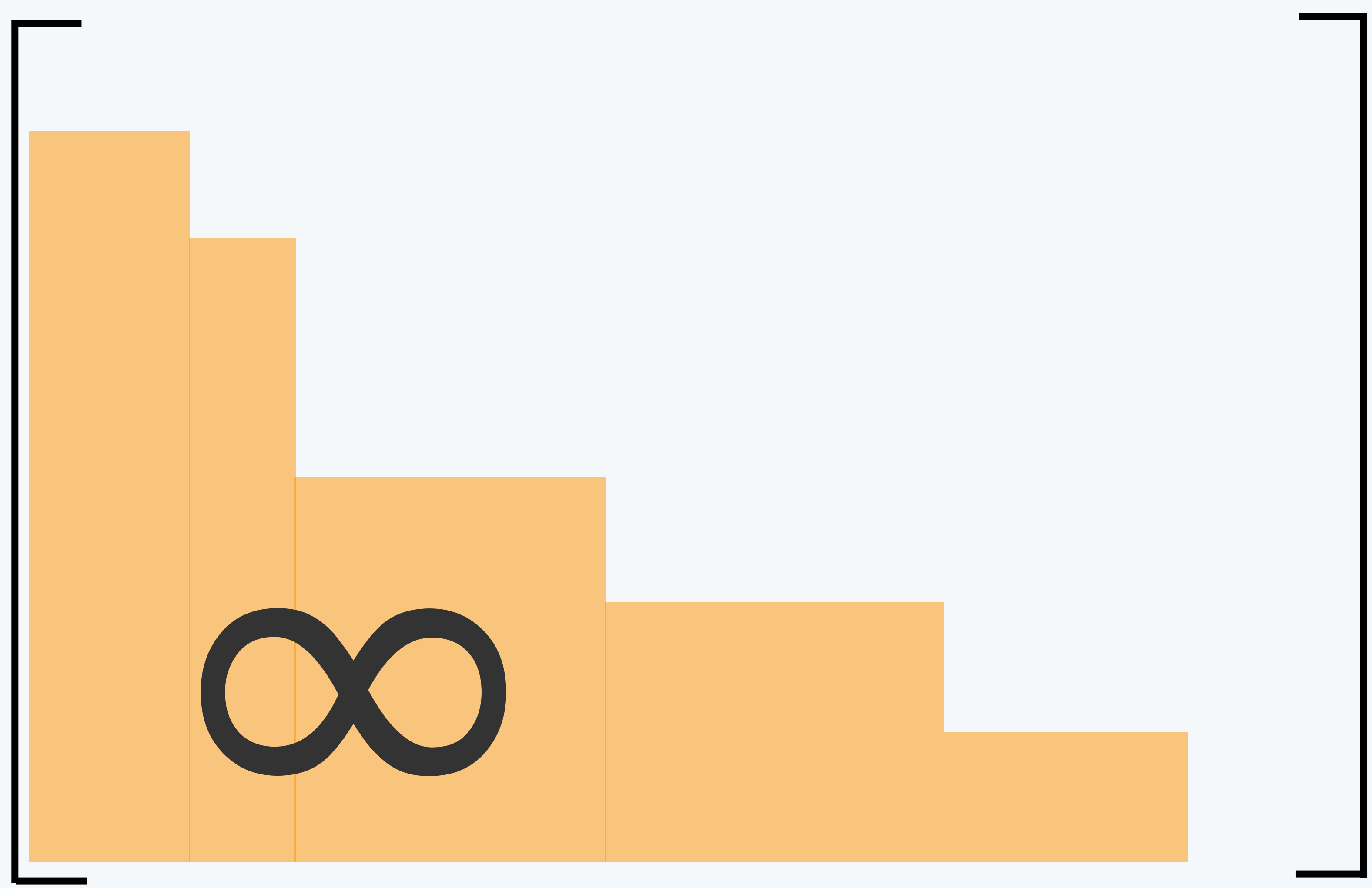
How many secrecy measures do we need?

- Probably more than one …
  - ML ill-suited for e.g., medical databases
  - DP ill-suited for side channels
  - Both ML and DP ill-suited for computationally-bounded eavesdroppers
- … but probably not 80+ either.

A Greedy Algorithm

- Given $A \subseteq \mathcal{Y}$, the *induced deterministic mechanism,* $P_{A,}$ is

$$p(y|x) = 1 \quad \text{if} \quad y = \text{argmin}\{c(x, y') : y' \in A\}$$



A Greedy Algorithm

- Given $A \subseteq \mathcal{Y}$, the *induced deterministic mechanism*, P_{A}, is

$$p(y|x) = 1 \text{ if } y = \operatorname{argmin}\{c(x, y') : y' \in A\}$$

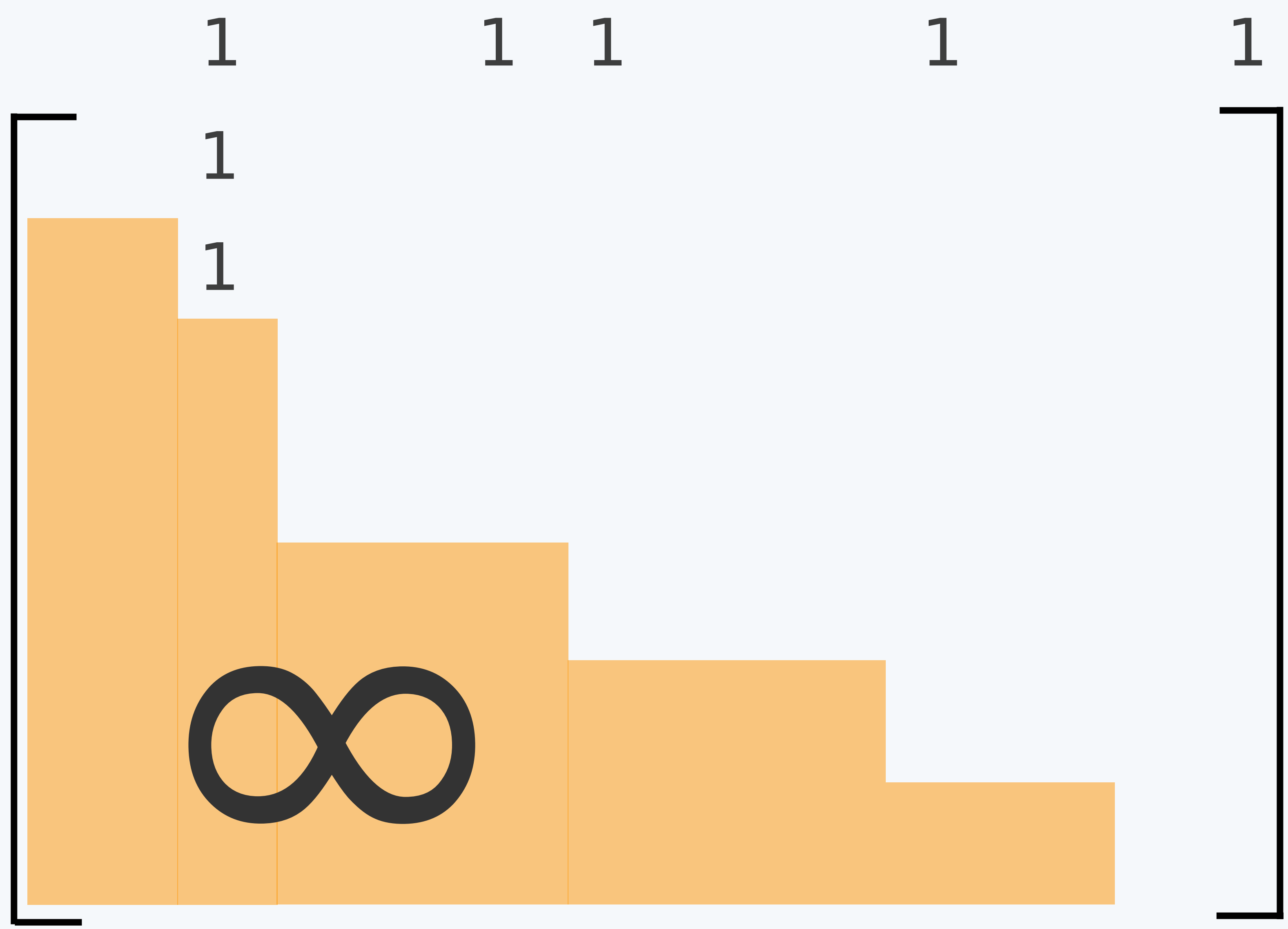
$A = \quad 1 \quad 1 \quad 1 \quad 1 \quad 1$

A Greedy Algorithm

- Given $A \subseteq \mathcal{Y}$, the *induced deterministic mechanism, P_{A},* is

$$p(y|x) = 1 \quad \text{if} \quad y = \operatorname{argmin}\{c(x, y') : y' \in A\}$$

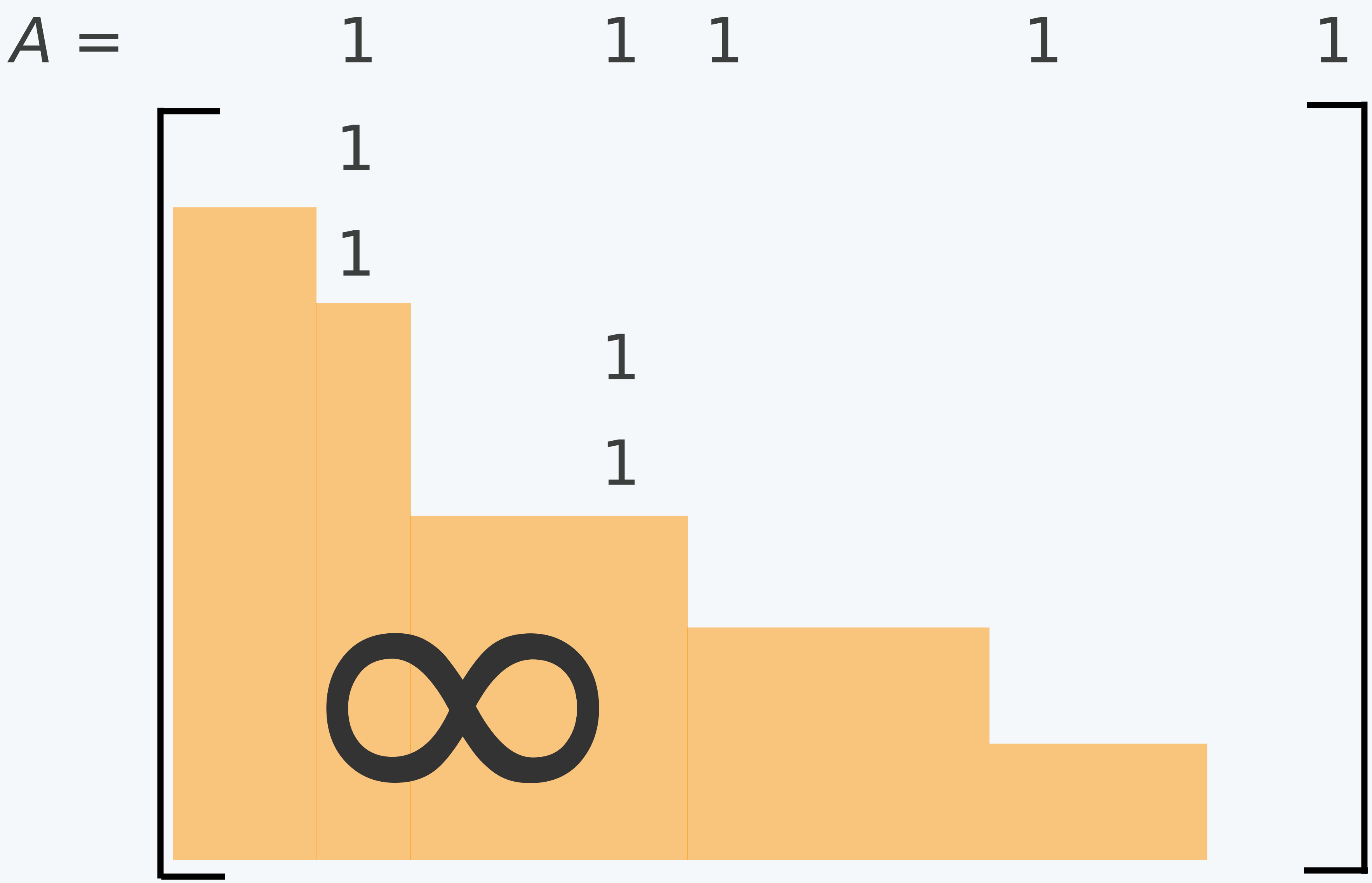
$A =$



A Greedy Algorithm

- Given $A \subseteq \mathcal{Y}$, the *induced deterministic mechanism, P_{A},* is

$$p(y|x) = 1 \quad \text{if} \quad y = \operatorname{argmin}\{c(x, y') : y' \in A\}$$



A Greedy Algorithm

- Given $A \subseteq \mathcal{Y}$, the *induced deterministic mechanism, P_{A},* is

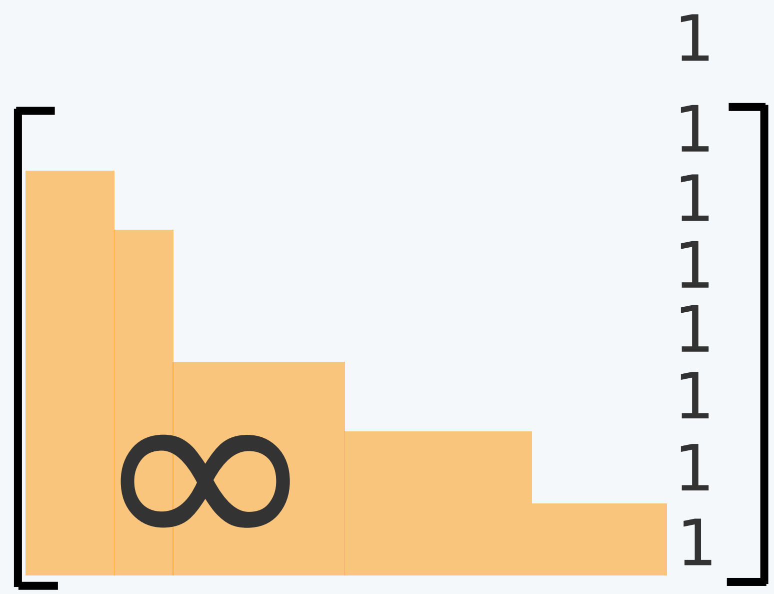
$$p(y|x) = 1 \quad \text{if} \quad y = \operatorname{argmin}\{c(x, y') : y' \in A\}$$

$$A = \begin{array}{c} \begin{array}{cccccc} & 1 & 1 & 1 & 1 & 1 \end{array} \\ \left[\begin{array}{cccccc} & 1 & & & & \\ & 1 & & & & \\ & & 1 & & & \\ & & 1 & & & \\ & & & 1 & & \\ \infty & & & & 1 & \\ & & & & & 1 \end{array} \right] \end{array}$$

A Greedy Algorithm

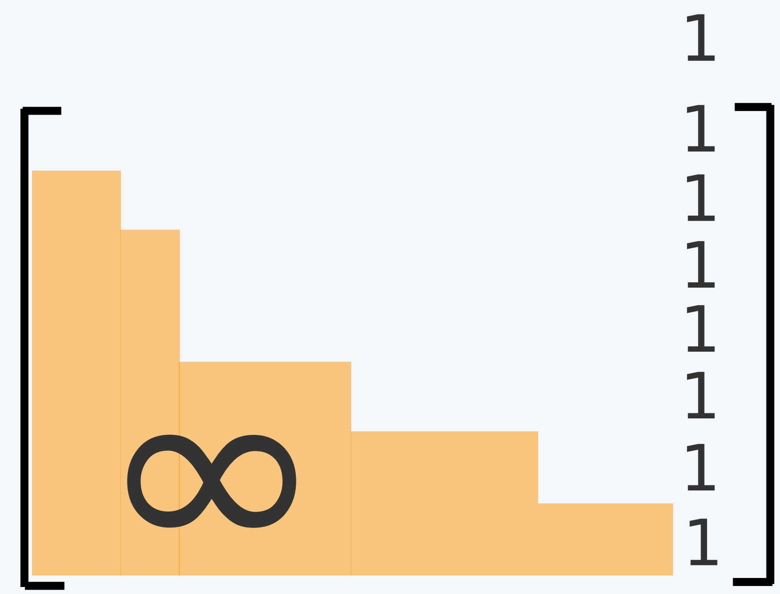
A Greedy Algorithm

- Start with a singleton A that minimizes the cost of P_A.

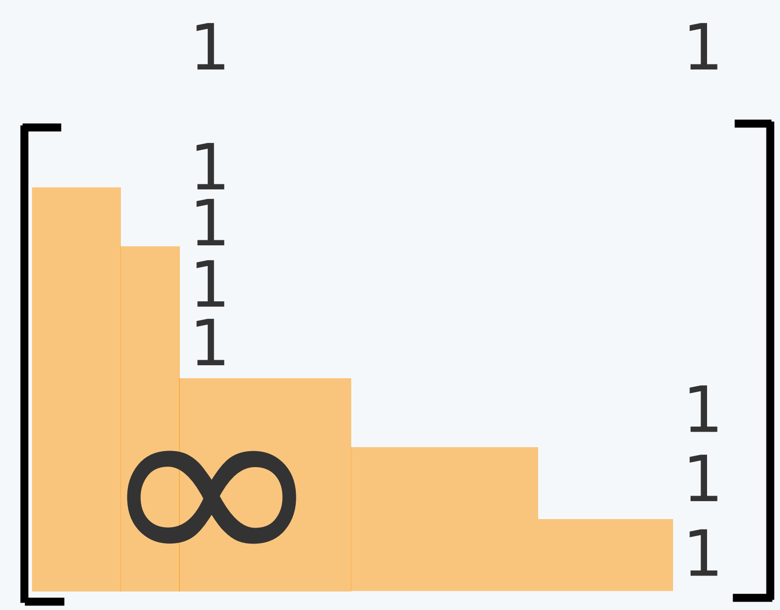


A Greedy Algorithm

- Start with a singleton A that minimizes the cost of P_A.



- Iterate: $A \to A \cup \{j\}$, where $j \notin A$ is chosen to minimize the cost of $P_{A \cup \{j\}}$.



Theorem

Theorem (Wu, Wagner, Suh '19):

For exp-leakage k, let

- $C^*(k)$ denote the optimum cost
- $C_\mathrm{G}(k)$ denote the cost obtained by the greedy algorithm

Then $C^*(1) = C_\mathrm{G}(1)$, $C^*(2) = C_\mathrm{G}(2)$, and

$$
\begin{aligned}
C^*(1) - C_G(k) &\geq \left(1 - \left(\frac{k-2}{k-1}\right)^{k-1}\right)(C^*(1) - C^*(k)) \\
&\geq \left(1 - \frac{1}{e}\right)(C^*(1) - C^*(k)) \\
&\geq 0.63(C^*(1) - C^*(k))
\end{aligned}
$$

Theorem

Theorem (Wu, Wagner, Suh '19):

For exp-leakage k, let

- $C*(k)$ denote the optimum cost
- $C_G(k)$ denote the cost obtained by the greedy algorithm

Then $C*(1) = C_G(1)$, $C*(2) = C_G(2)$, and

$$\begin{aligned} C^*(1) - C_G(k) &\geq \left(1 - \left(\frac{k-2}{k-1}\right)^{k-1}\right)(C^*(1) - C^*(k)) \\ &\geq \left(1 - \frac{1}{e}\right)(C^*(1) - C^*(k)) \\ &\geq 0.63(C^*(1) - C^*(k)) \end{aligned}$$

Proof: submodularity of -cost(P_A).

Theorem

Theorem (Wu, Wagner, Suh '19):

For exp-leakage k, let

- $C^*(k)$ denote the optimum cost
- $C_G(k)$ denote the cost obtained by the greedy algorithm

Then $C^*(1) = C_G(1)$, $C^*(2) = C_G(2)$, and

$$C^*(1) - C_G(k) \geq \left(1 - \left(\frac{k-2}{k-1}\right)^{k-1}\right)(C^*(1) - C^*(k))$$

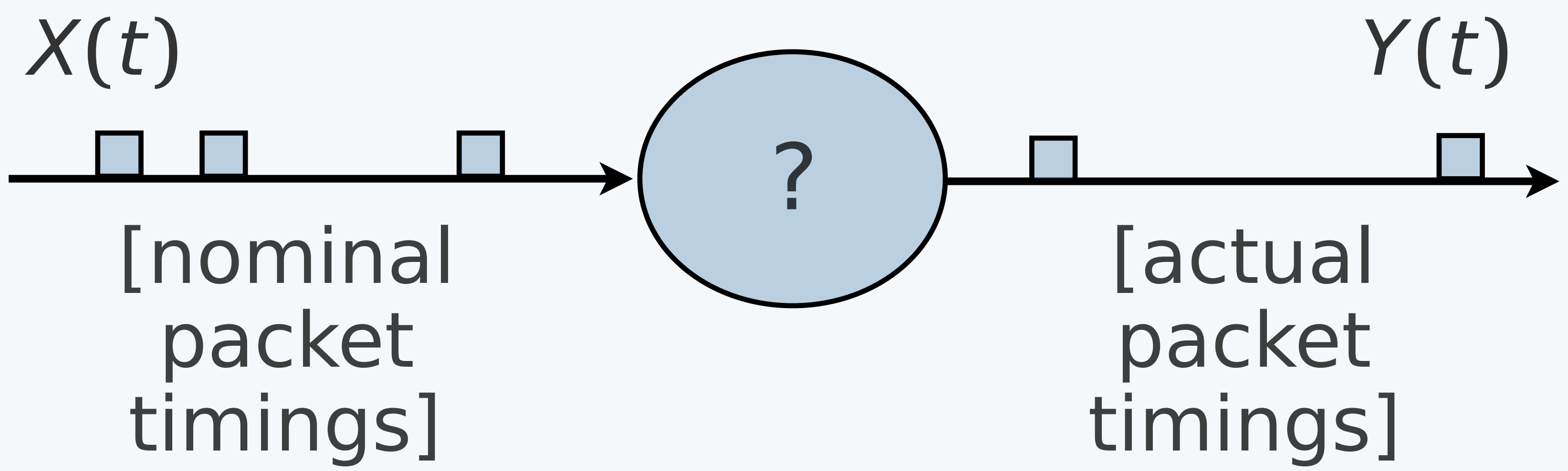
$$\geq \left(1 - \frac{1}{e}\right)(C^*(1) - C^*(k))$$

$$\geq 0.63(C^*(1) - C^*(k))$$

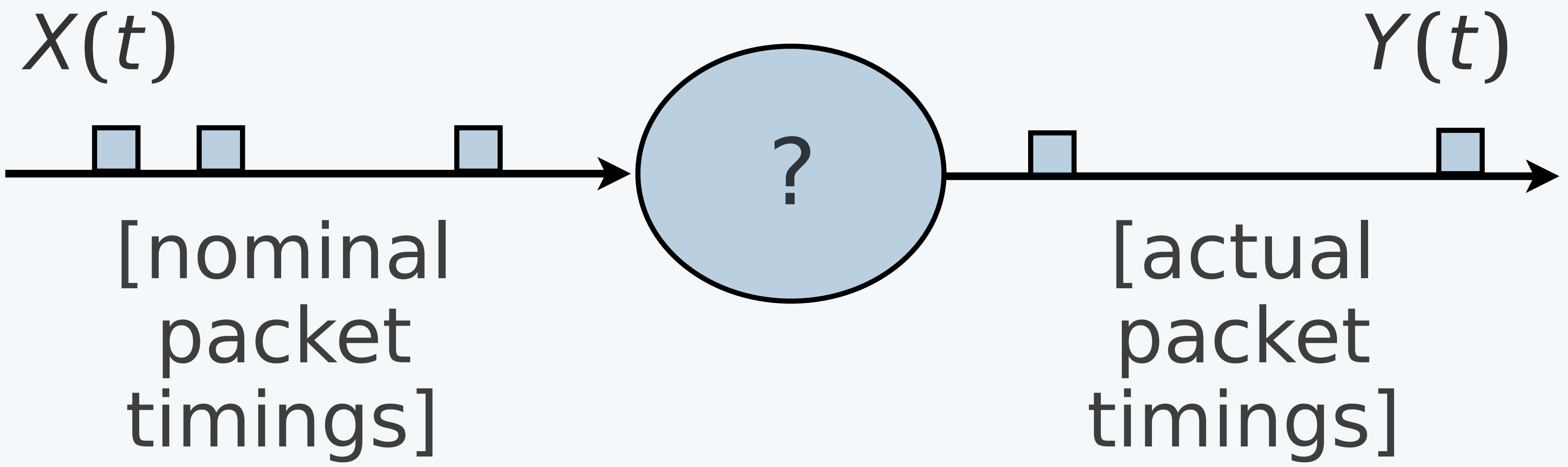
Proof: submodularity of -cost(P_A).

Note: leads to a sequence of approximations.

How to Delay Packets?

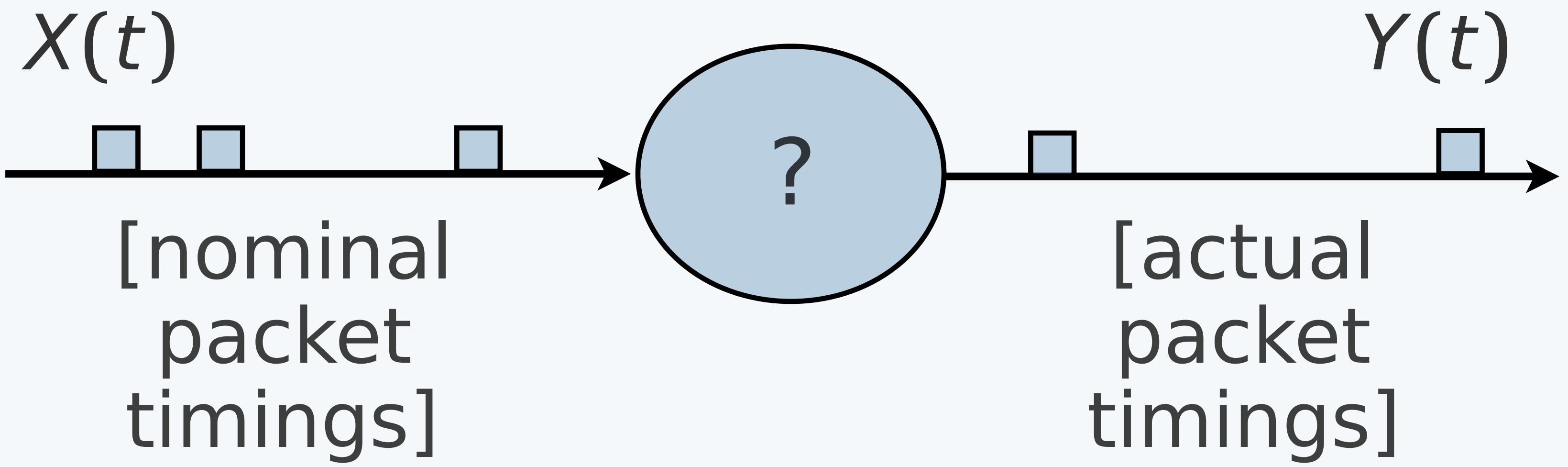


How to Delay Packets?



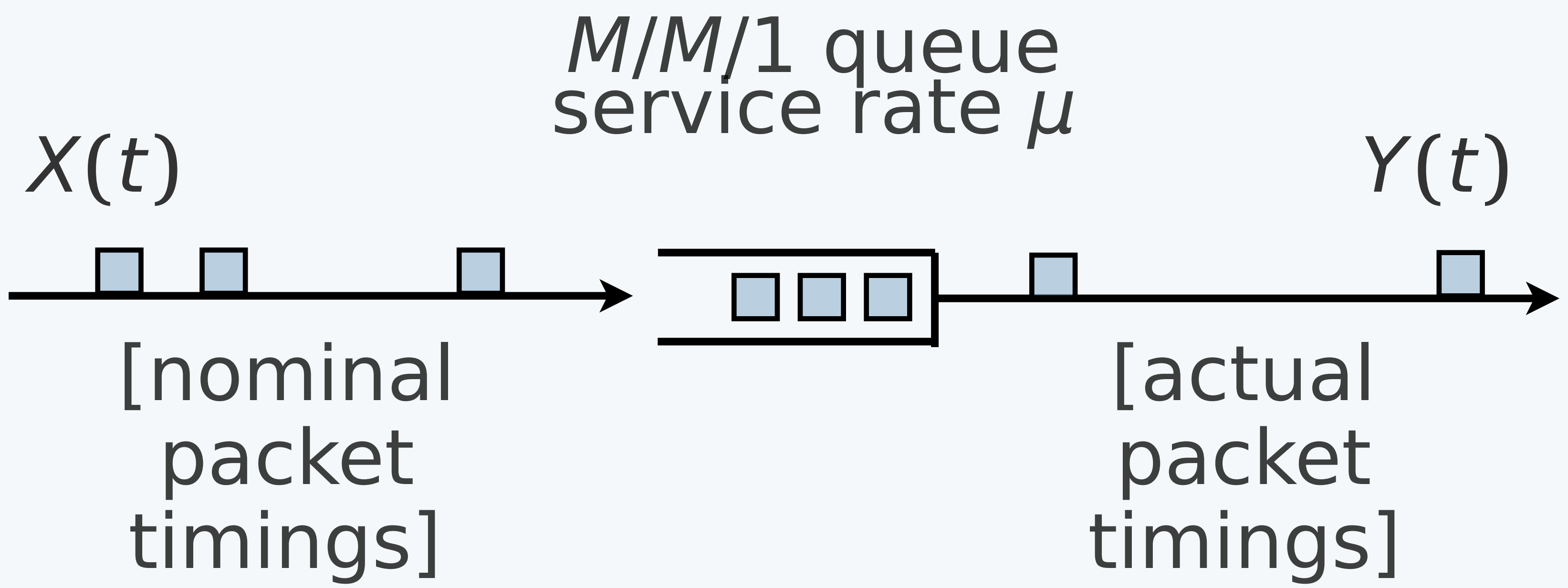
- Suppose $X(t)$ is a Poisson process with rate λ

How to Delay Packets?

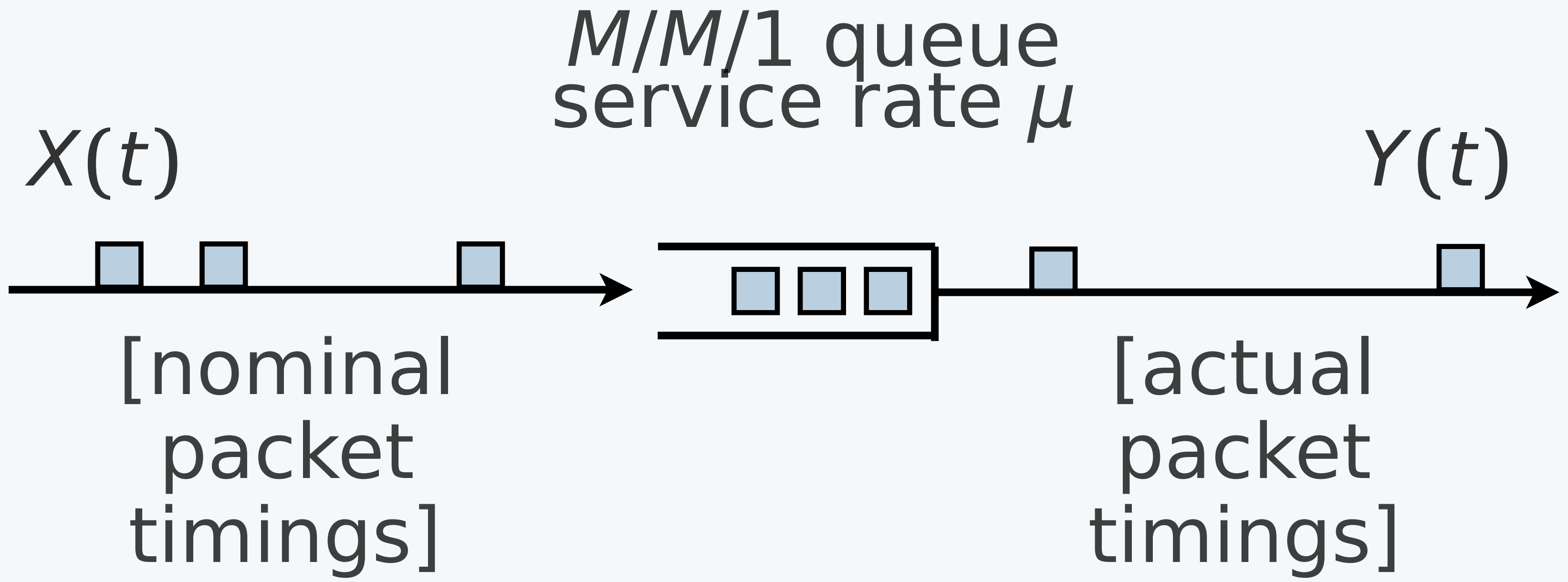


- Suppose $X(t)$ is a Poisson process with rate λ
- How to blur the packet timings to minimize leakage?

Try an $M/M/1$ Queue

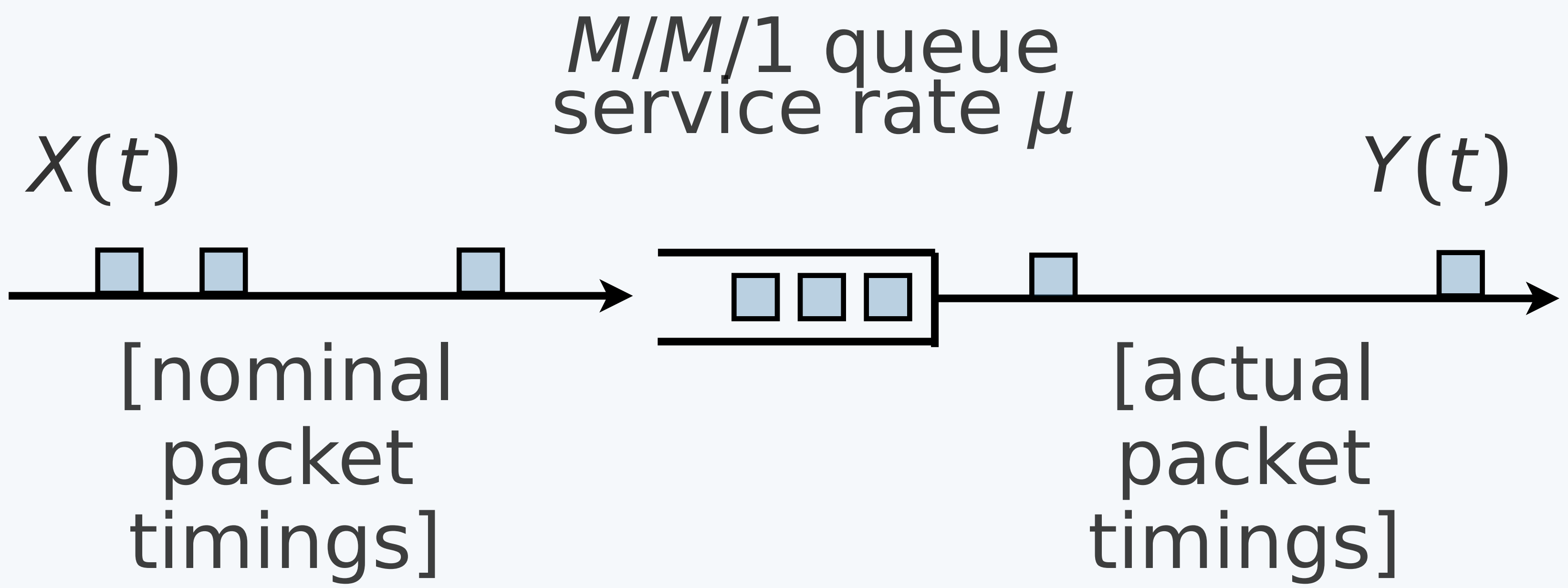


Try an $M/M/1$ Queue



$$\frac{1}{T} \cdot \mathcal{L}\left(\{X(t)\}_{t=0}^{T} \to \{Y(t)\}_{t=0}^{T}\right) = \mu \quad \text{nats}$$

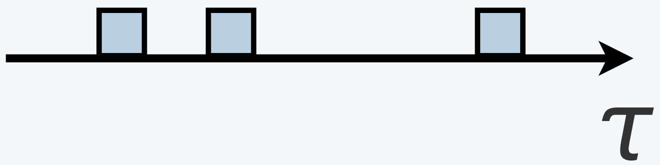
Try an $M/M/1$ Queue



$$\frac{1}{T} \cdot \mathcal{L}\left(\{X(t)\}_{t=0}^{T} \to \{Y(t)\}_{t=0}^{T} \right) = \mu \quad \text{nats}$$

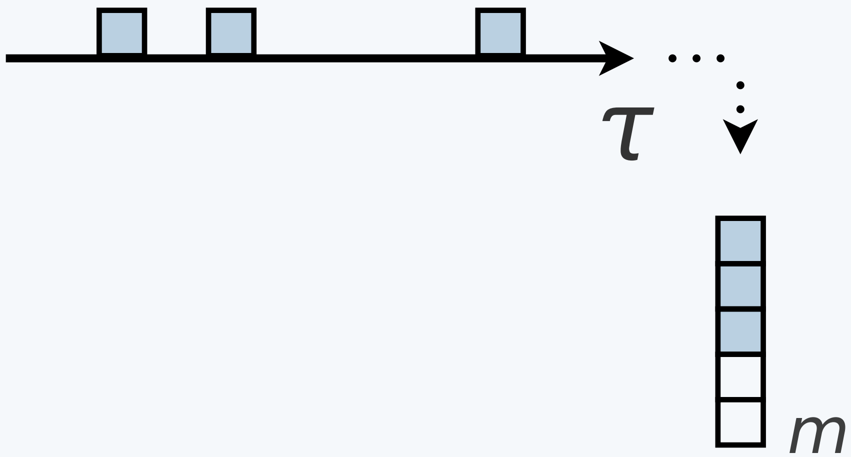
[leakage rate is at least λ]

Accumulate and Dump



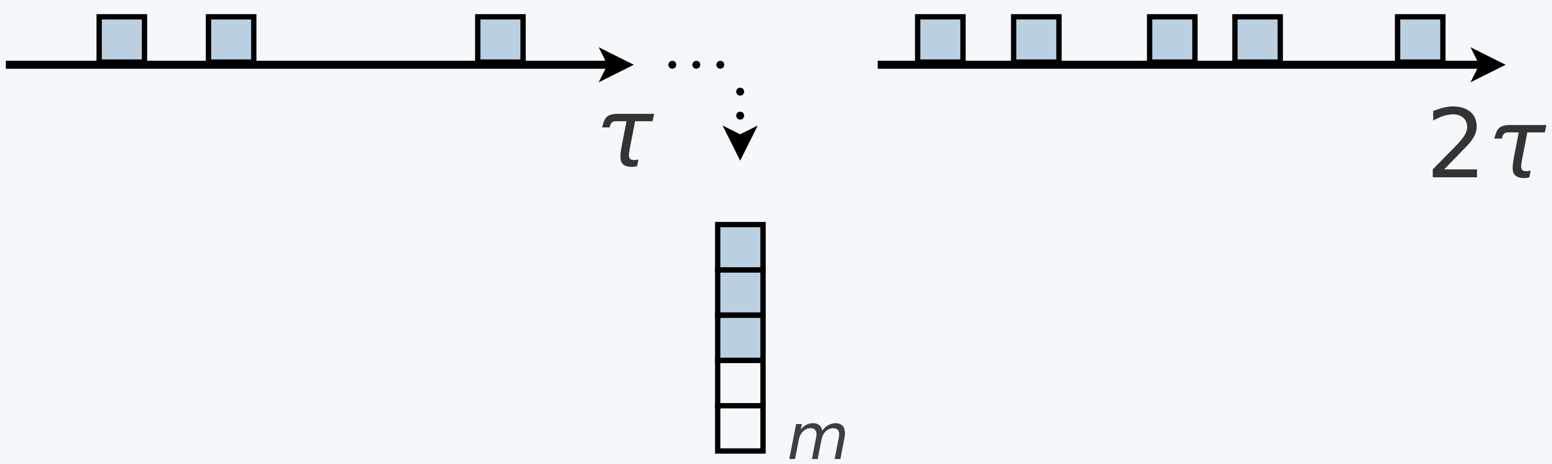
$$\frac{1}{T} \cdot \mathcal{L}\left(\{X(t)\}_{t=0}^{T} \to \{Y(t)\}_{t=0}^{T}\right) \leq \frac{1}{\tau} \log m$$

Accumulate and Dump



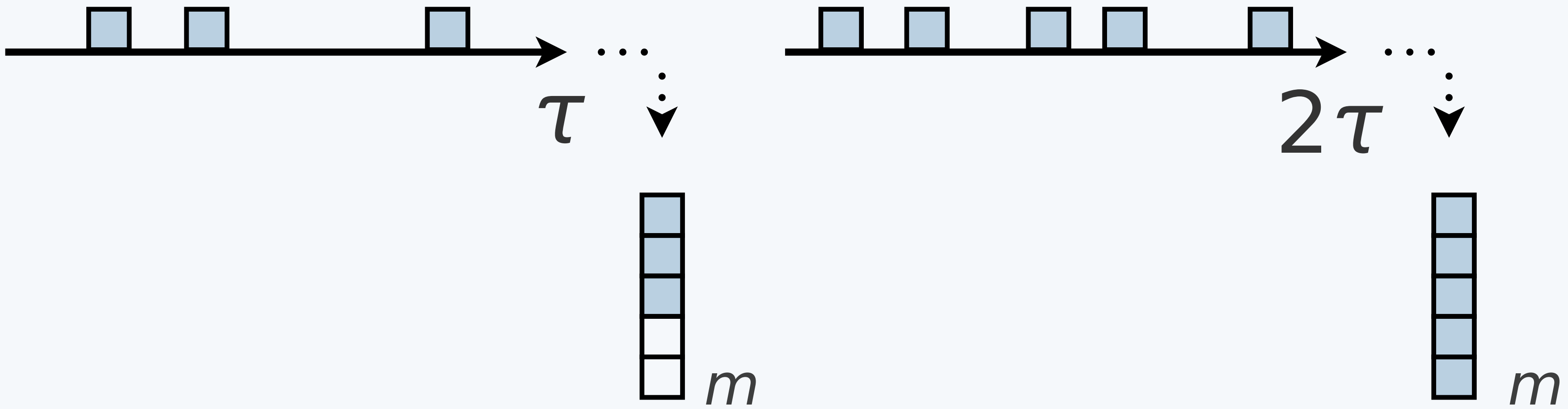
$$\frac{1}{T} \cdot \mathcal{L}\left(\{X(t)\}_{t=0}^{T} \to \{Y(t)\}_{t=0}^{T}\right) \leq \frac{1}{\tau} \log m$$

Accumulate and Dump



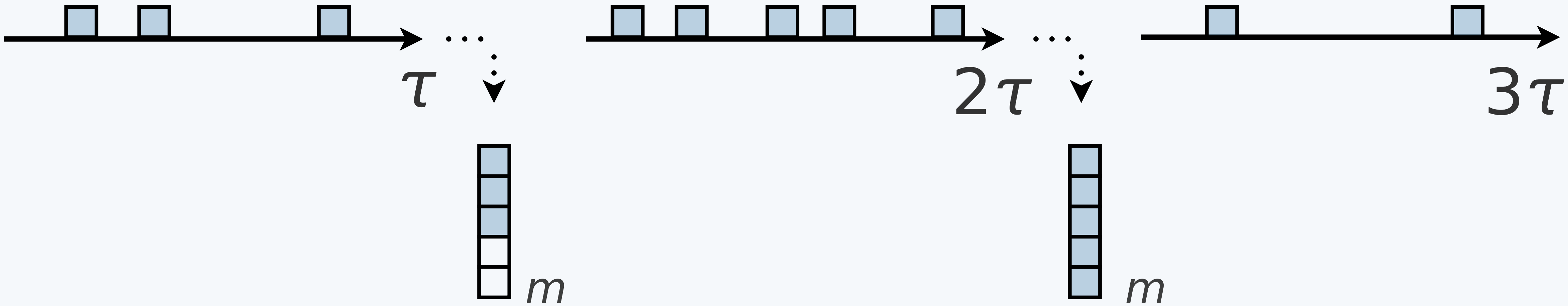
$$\frac{1}{T} \cdot \mathcal{L}\left(\{X(t)\}_{t=0}^{T} \to \{Y(t)\}_{t=0}^{T}\right) \leq \frac{1}{\tau} \log m$$

Accumulate and Dump



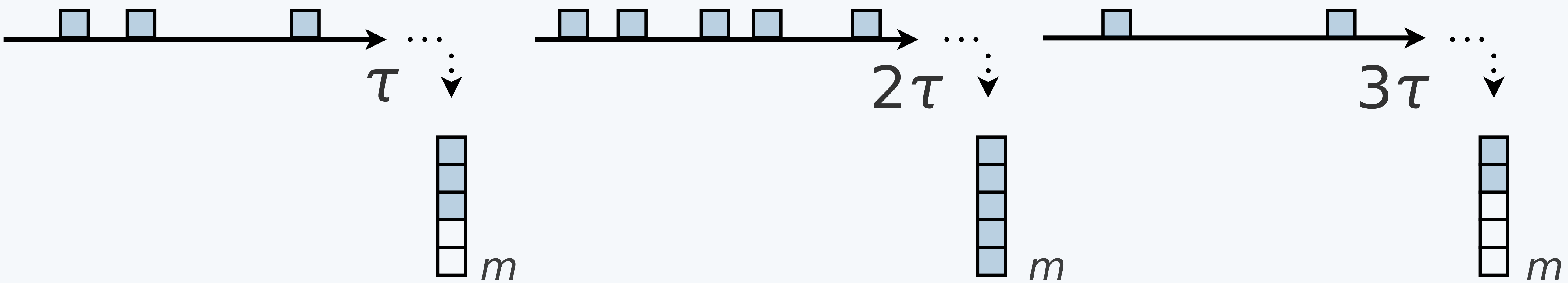
$$\frac{1}{T} \cdot \mathcal{L}\left(\{X(t)\}_{t=0}^{T} \to \{Y(t)\}_{t=0}^{T}\right) \leq \frac{1}{\tau} \log m$$

Accumulate and Dump



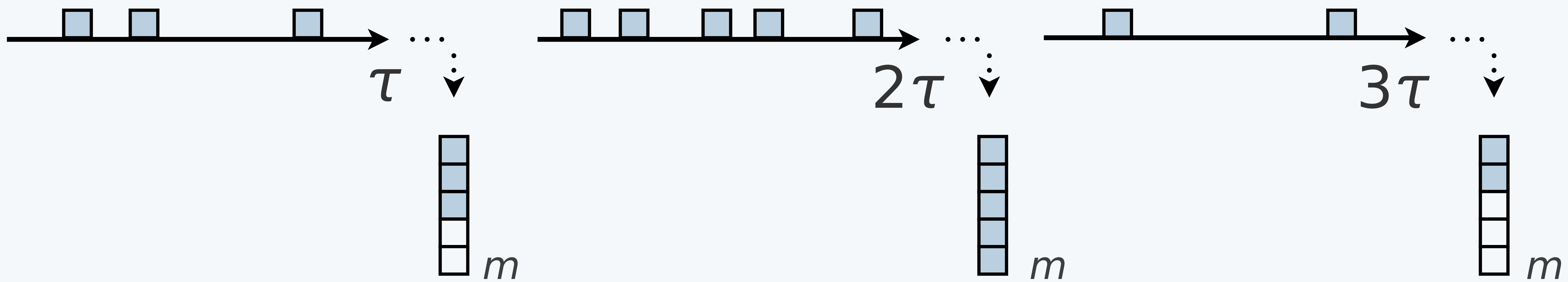
$$\frac{1}{T} \cdot \mathcal{L}\left(\{X(t)\}_{t=0}^{T} \to \{Y(t)\}_{t=0}^{T}\right) \leq \frac{1}{\tau} \log m$$

Accumulate and Dump



$$\frac{1}{T} \cdot \mathcal{L}\left(\{X(t)\}_{t=0}^{T} \to \{Y(t)\}_{t=0}^{T}\right) \leq \frac{1}{\tau} \log m$$

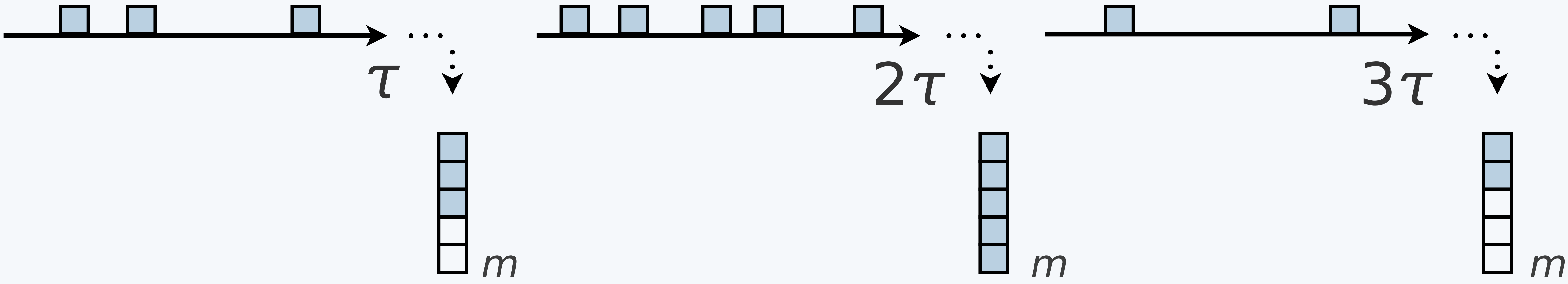
Accumulate and Dump



$$\frac{1}{T} \cdot \mathcal{L}\left(\{X(t)\}_{t=0}^{T} \to \{Y(t)\}_{t=0}^{T}\right) \leq \frac{1}{\tau} \log m$$

[quantization leaks less than “adding noise”]

Accumulate and Dump

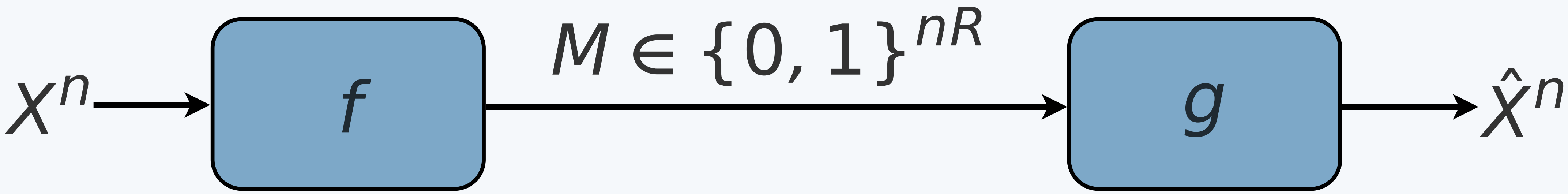


$$\frac{1}{T} \cdot \mathcal{L}\left(\{X(t)\}_{t=0}^{T} \to \{Y(t)\}_{t=0}^{T}\right) \leq \frac{1}{\tau} \log m$$

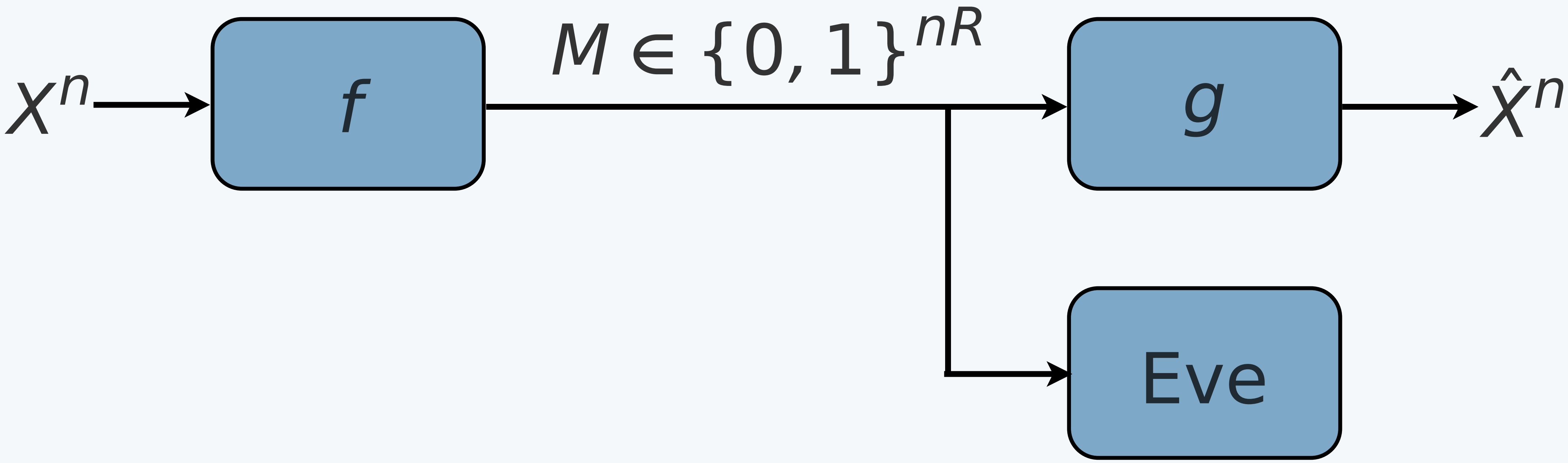
[quantization leaks less than "adding noise"]

[cf. Kadloor, Kiyavash, and Venkitasubramaniam '16]

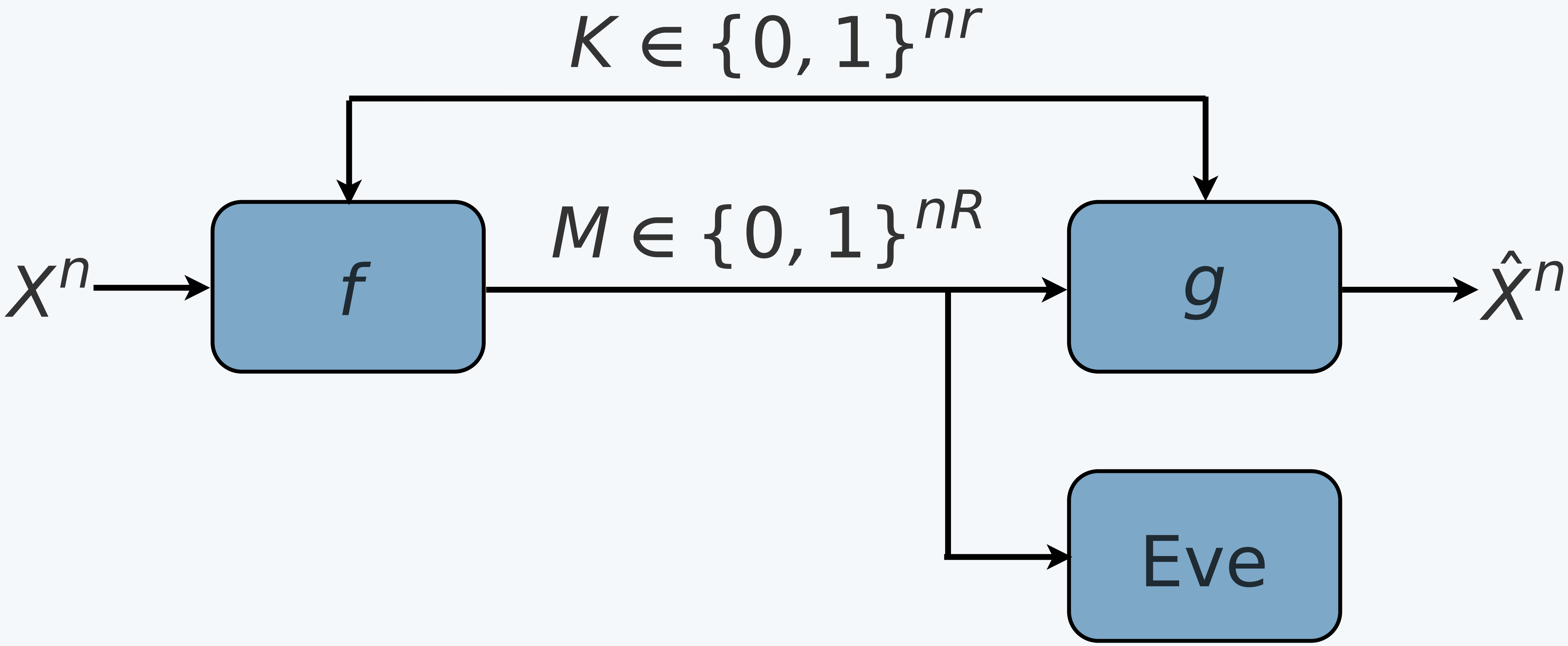
The Shannon Cipher System



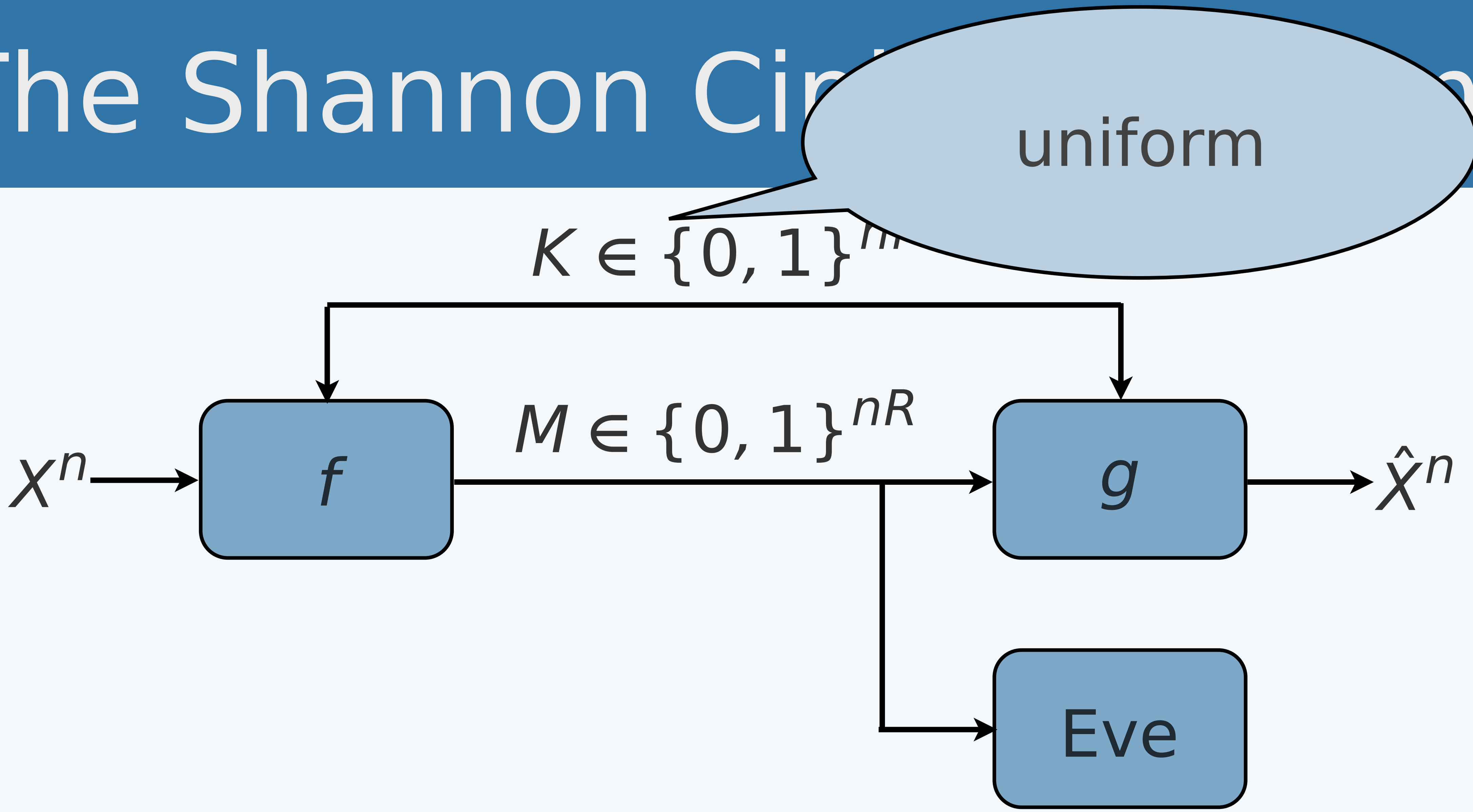
The Shannon Cipher System



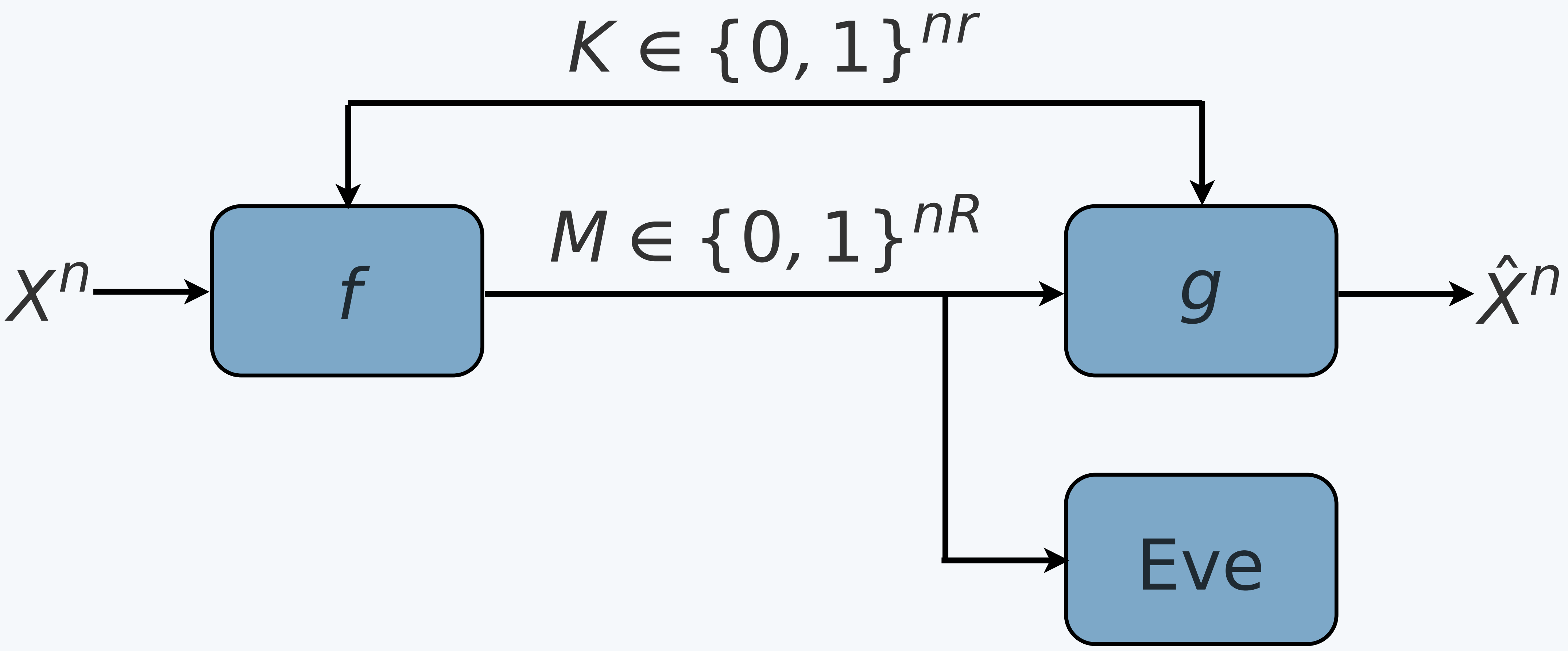
The Shannon Cipher System



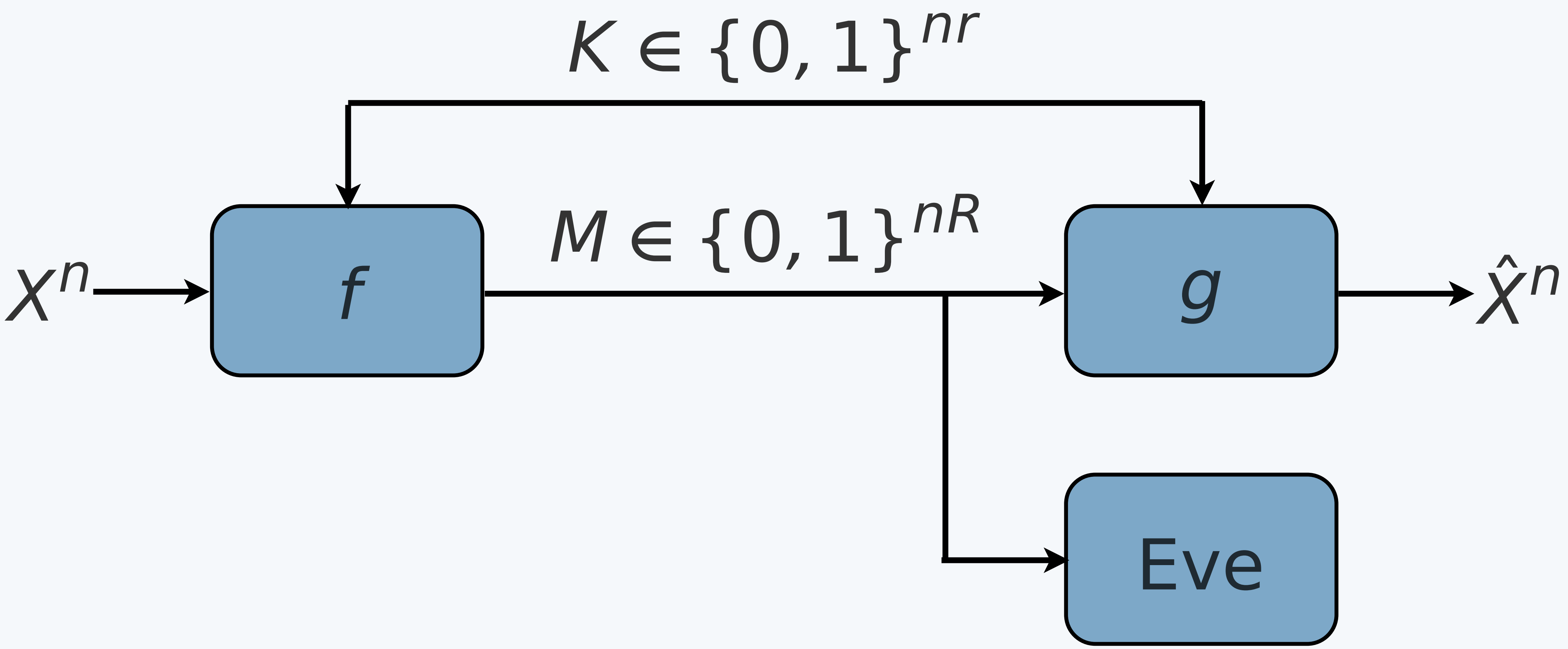
The Shannon Cip[...]



The Shannon Cipher System

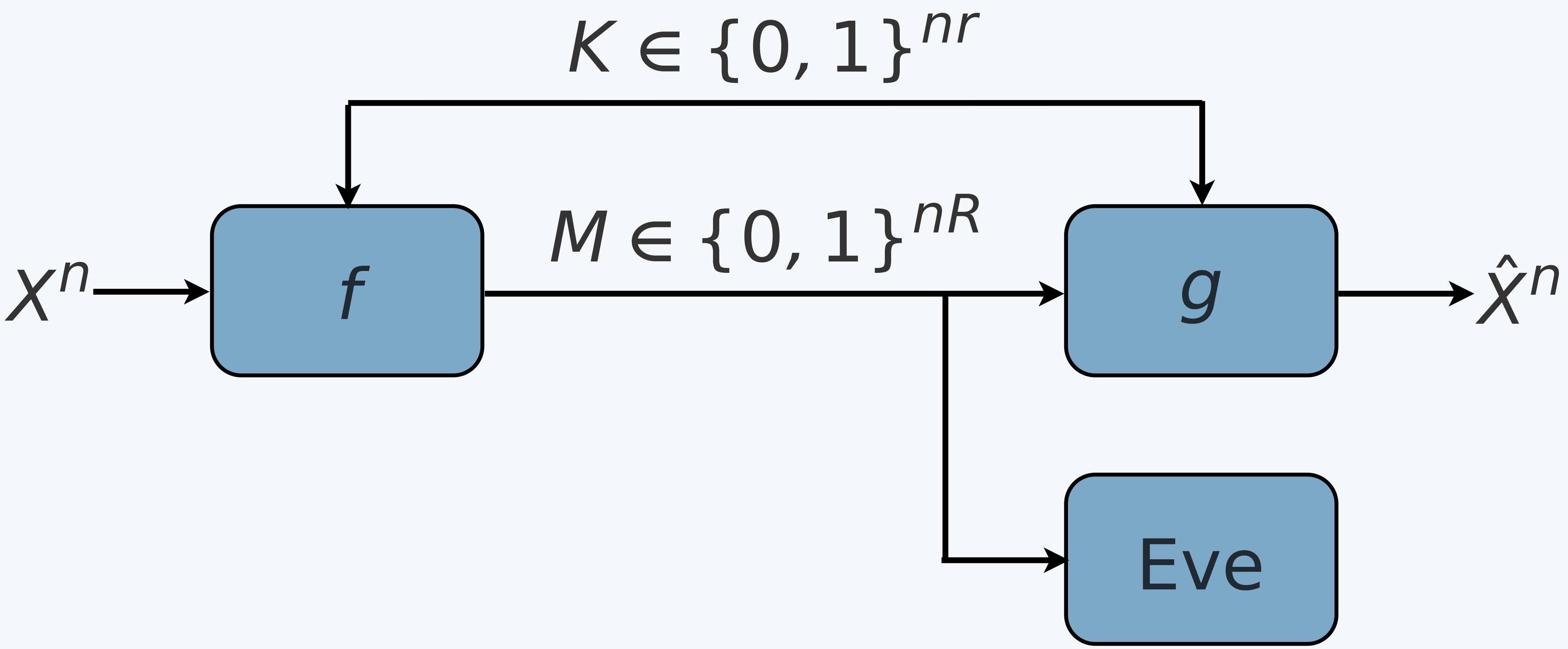


The Shannon Cipher System



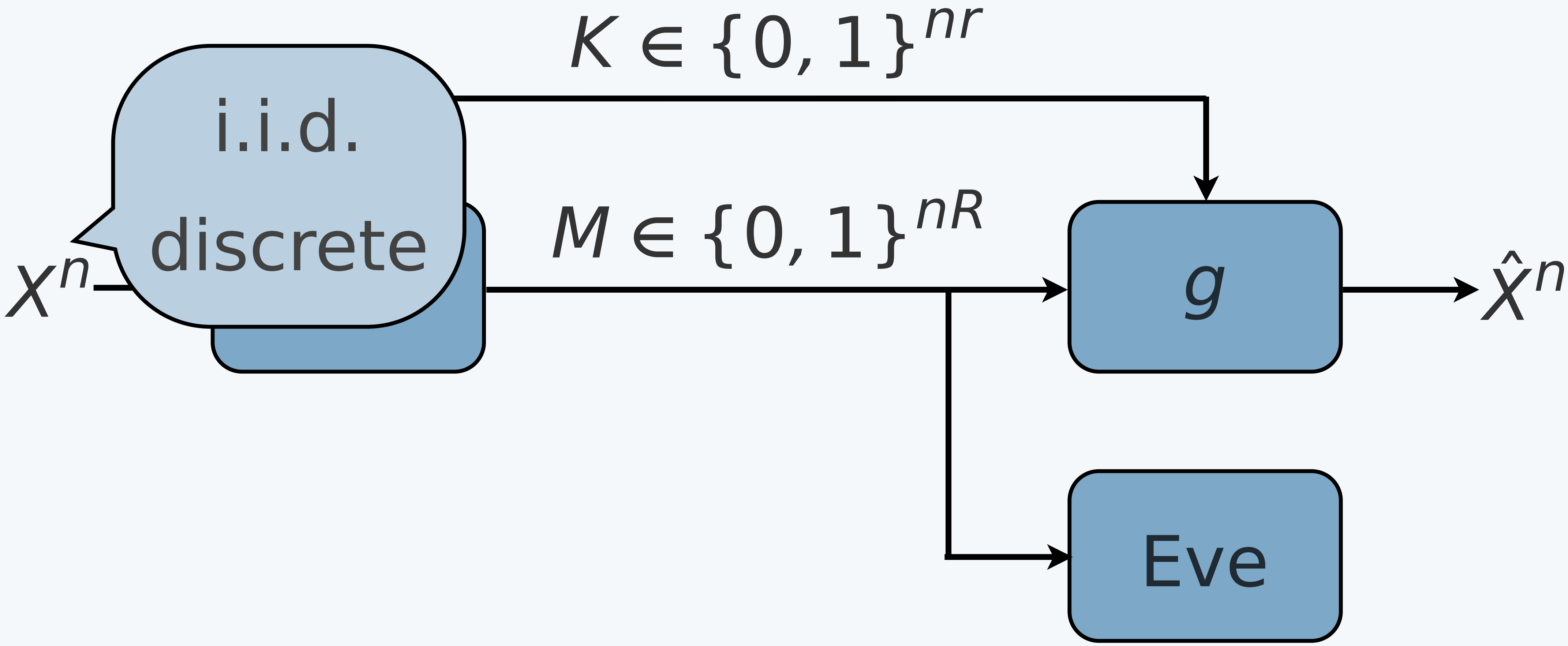
- Shannon ('49): perfect secrecy is possible if(f) the key rate r exceeds the message rate R.

The Shannon Cipher System

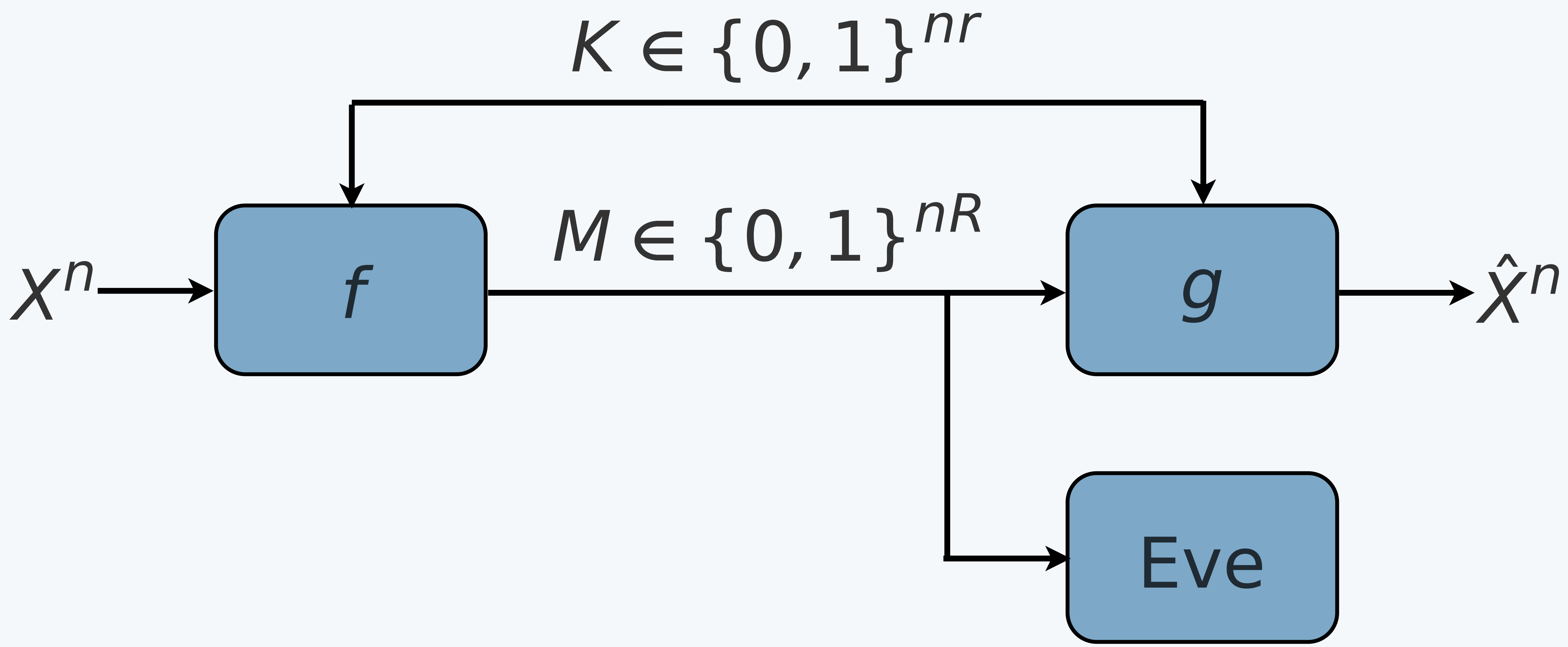


- Shannon ('49): perfect secrecy is possible if(f) the key rate r exceeds the message rate R.
- How to design f and g to minimize leakage when $r < R$?

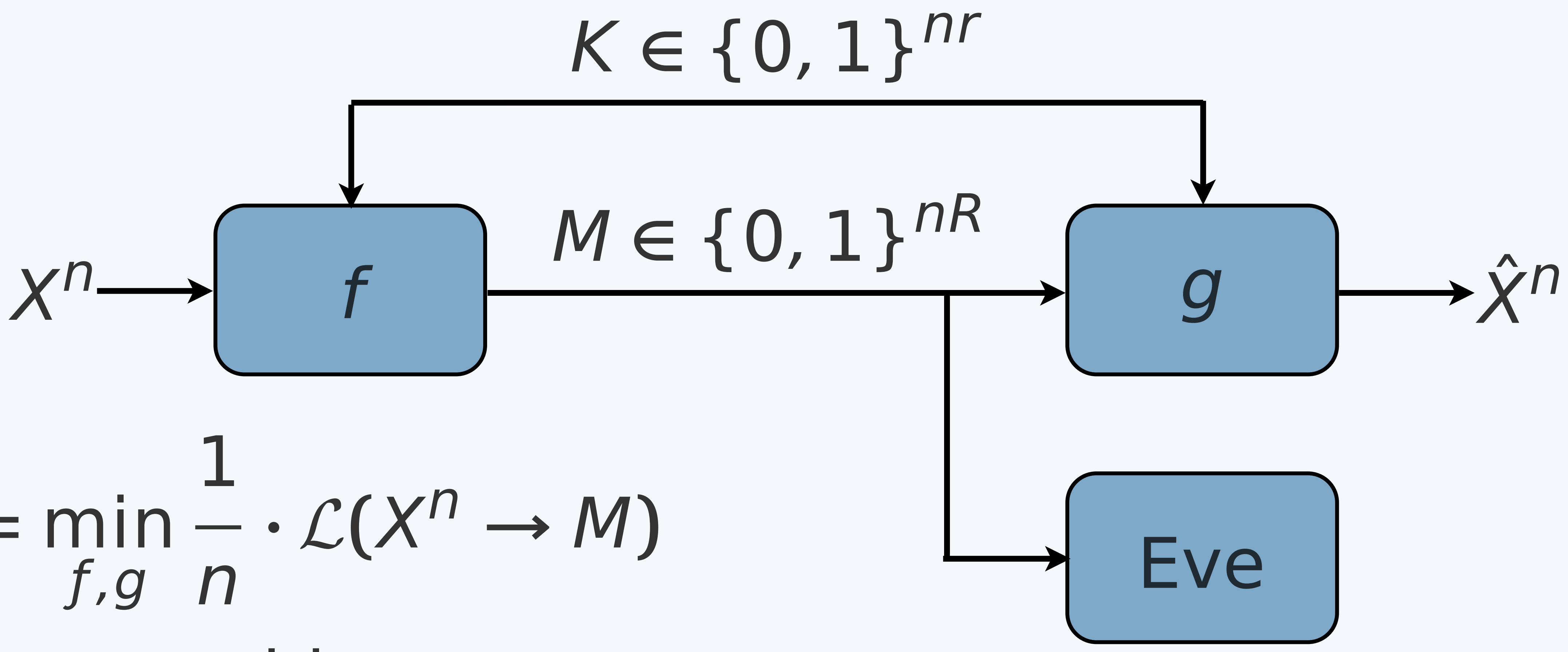
Leakage and Shannon's Cipher



Leakage and Shannon's Cipher



Leakage and Shannon's Cipher



$$L_n = \min_{f,g} \frac{1}{n} \cdot \mathcal{L}(X^n \to M)$$

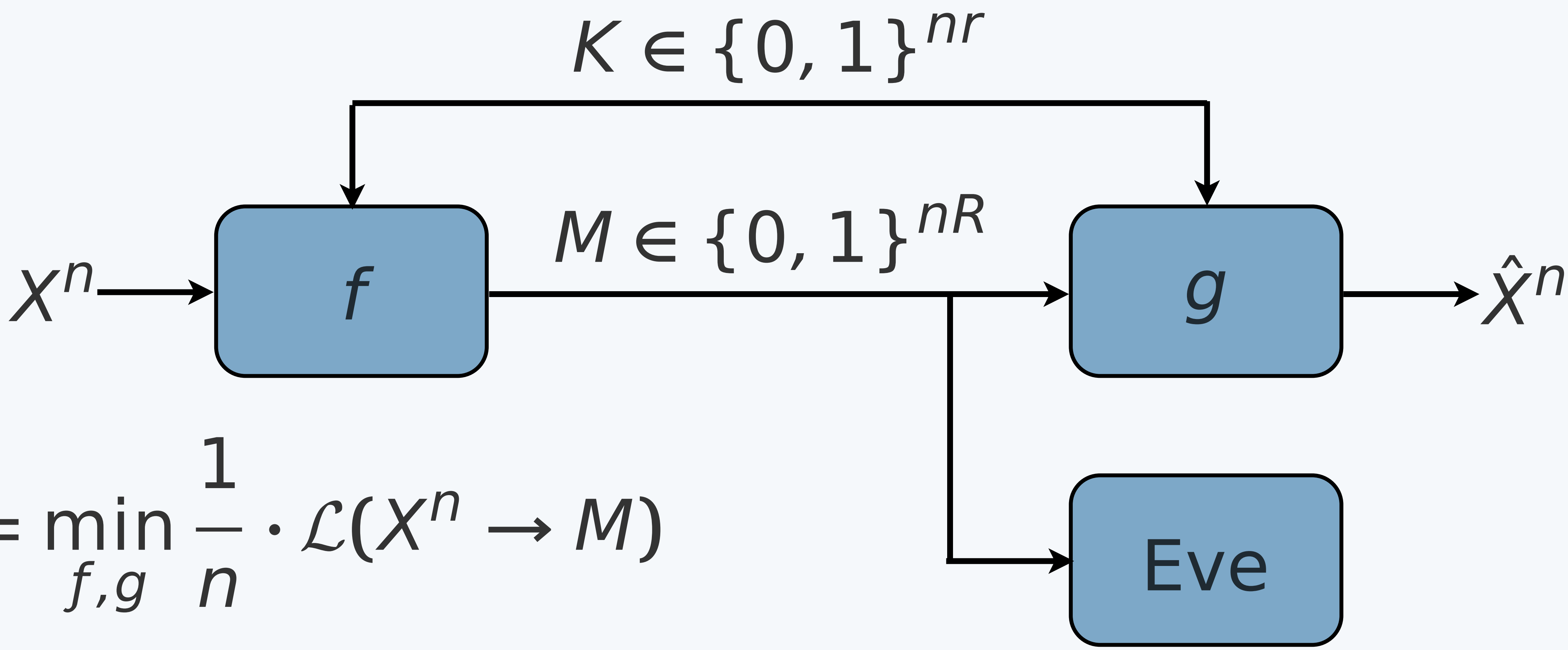
subject to

$$f : \mathcal{X}^n \times \{0,1\}^{nr} \mapsto \{0,1\}^{nR}$$

$$g : \{0,1\}^{nR} \times \{0,1\}^{nr} \mapsto \hat{\mathcal{X}}^n$$

$$\frac{1}{n}\sum_{i=1}^{n} E[d(X_i, \hat{X}_i)] \leq D$$

Leakage and Shannon's Cipher



$$L_n = \min_{f,g} \frac{1}{n} \cdot \mathcal{L}(X^n \to M)$$

subject to

$$f : \mathcal{X}^n \times \{0,1\}^{nr} \mapsto \{0,1\}^{nR}$$

$$g : \{0,1\}^{nR} \times \{0,1\}^{nr} \mapsto \hat{\mathcal{X}}^n$$

$$\frac{1}{n} \sum_{i=1}^{n} E[d(X_i, \hat{X}_i)] \le D$$

$$L = \lim_{n \to \infty} L_n$$

Leakage and Shannon's Cipher

Theorem (Issa-Kamath-Wagner): Let $R(D)$ denote the rate-distortion function for the source. If

$$R < R(D),$$

then the problem is infeasible. Otherwise, the min. max. leakage is

$$L = [R(D) - r]^{+}$$

Leakage and Shannon's Cipher

Theorem (Issa-Kamath-Wagner): Let $R(D)$ denote the rate-distortion function for the source. If

$$R < R(D),$$

then the problem is infeasible. Otherwise, the min. max. leakage is

$$L = [R(D) - r]^+$$

Notes:

Leakage and Shannon’s Cipher

Theorem (Issa-Kamath-Wagner): Let $R(D)$ denote the rate-distortion function for the source. If

$$R < R(D),$$

then the problem is infeasible. Otherwise, the min. max. leakage is

$$L = [R(D) - r]^{+}$$

Notes:

- Using MI instead of leakage gives same result

Leakage and Shannon's Cipher

Theorem (Issa-Kamath-Wagner): Let $R(D)$ denote the rate-distortion function for the source. If

$$R < R(D),$$

then the problem is infeasible. Otherwise, the min. max. leakage is

$$L = [R(D) - r]^{+}$$

Notes:

- Using MI instead of leakage gives same result
  - Though difference in optimal schemes...

Leakage and Shannon's Cipher

Theorem (Issa-Kamath-Wagner): Let $R(D)$ denote the rate-distortion function for the source. If

$$R < R(D),$$

then the problem is infeasible. Otherwise, the min. max. leakage is

$$L = [R(D) - r]^+$$

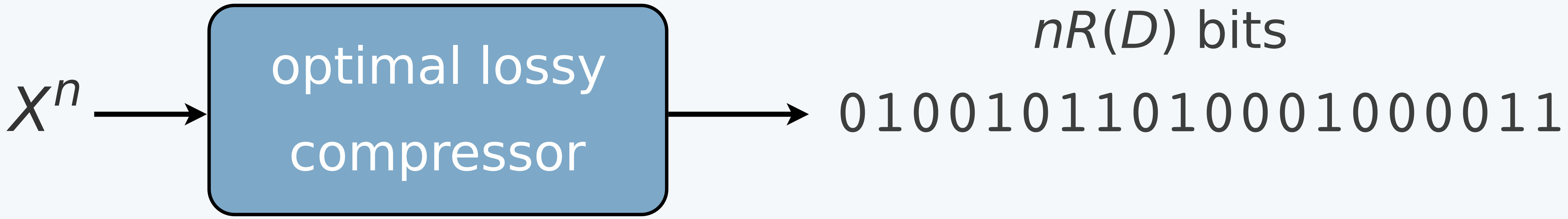
Notes:

- Using MI instead of leakage gives same result
  - Though difference in optimal schemes...
- Large deviations (and a.s.) result

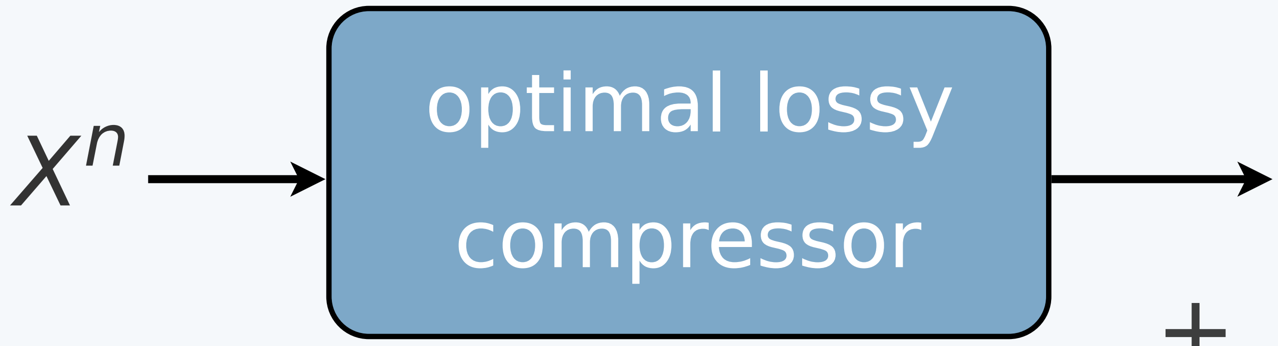
Achievability for Primary User

$X^n \longrightarrow$

Achievability for Primary User



Achievability for Primary User

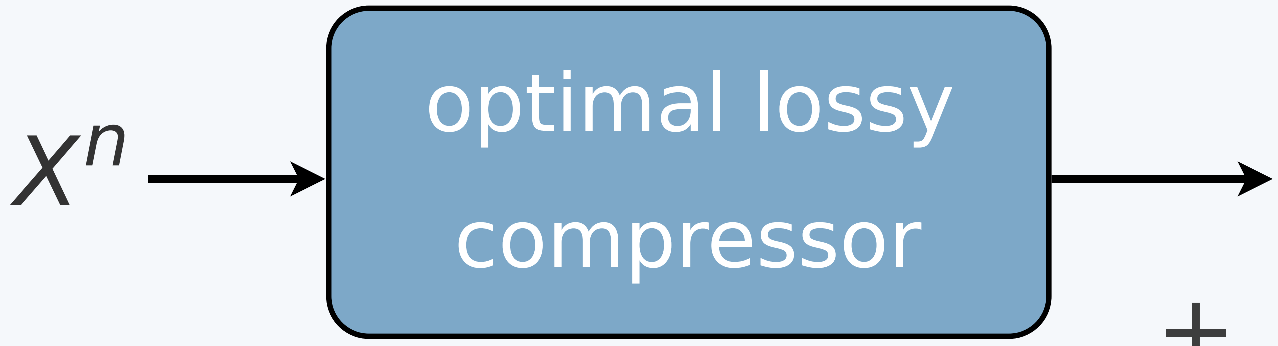


$nR(D)$ bits

0 1 0 0 1 0 1 1 0 1 0 0 0 1 0 0 0 0 1 1

\+ 1 1 0 1 0 1 1 1 0 0 0 (key)

Achievability for Primary User



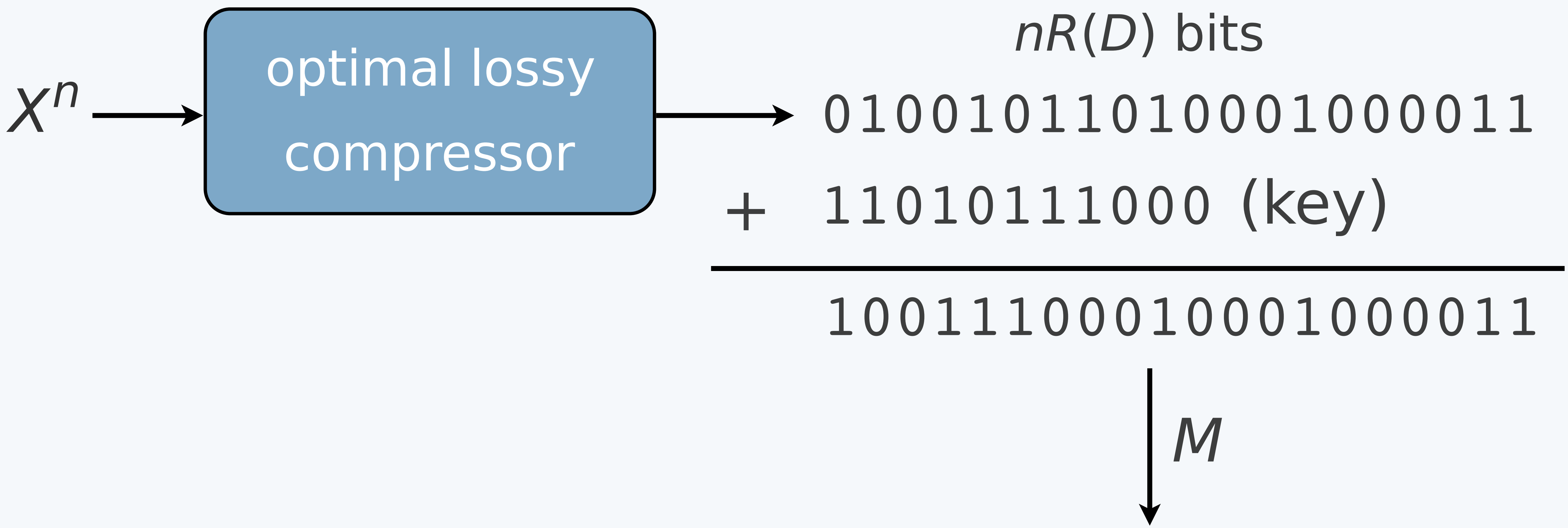
$nR(D)$ bits

010010111010001000011

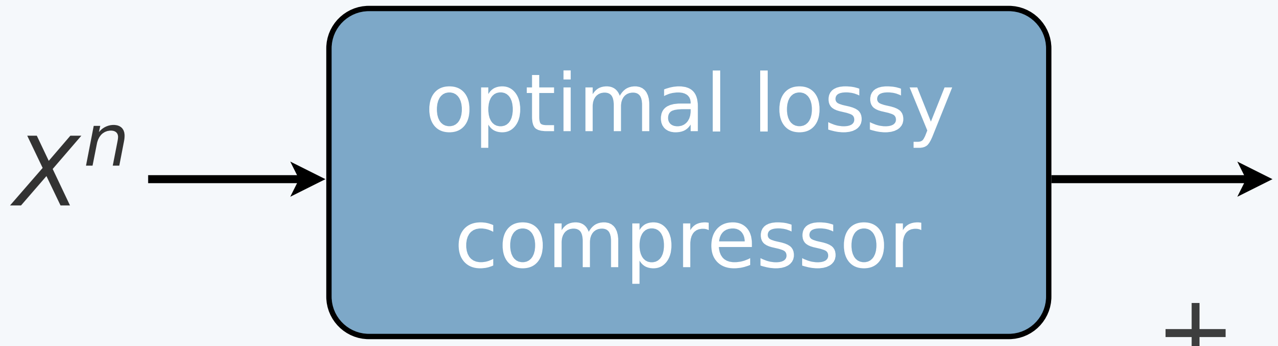
\+ 11010111000 (key)

10011100010001000011

Achievability for Primary User



Achievability for Primary User



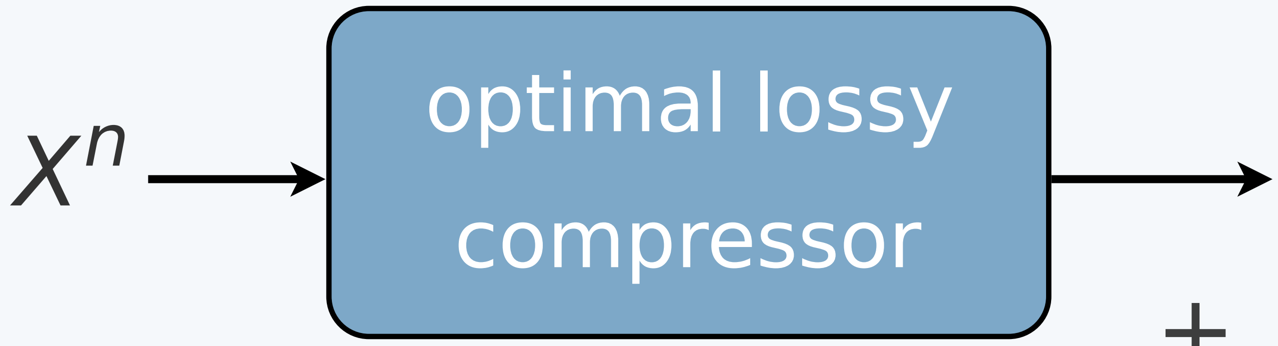
$nR(D)$ bits

$$
\begin{array}{rl}
 & 01001011010001000011 \\
+ & 11010111000 \text{ (key)} \\
\hline
 & 10011100010001000011
\end{array}
$$

$\Big\downarrow M$

10011100010001000011

Achievability for Primary User



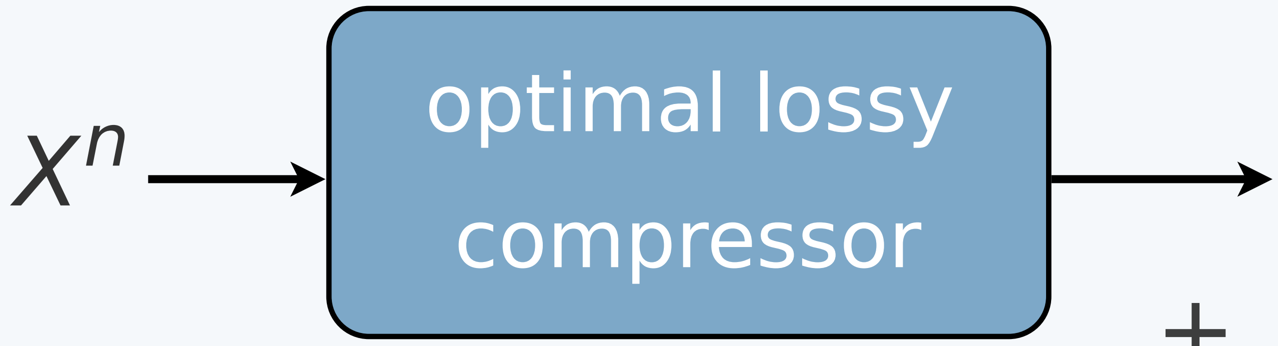
$nR(D)$ bits

```
  0100101101000100011
+ 11010111000 (key)
---------------------
  1001110001000100011
```

$\downarrow M$

```
  1001110001000100011
+ 11010111000 (key)
```

Achievability for Primary User



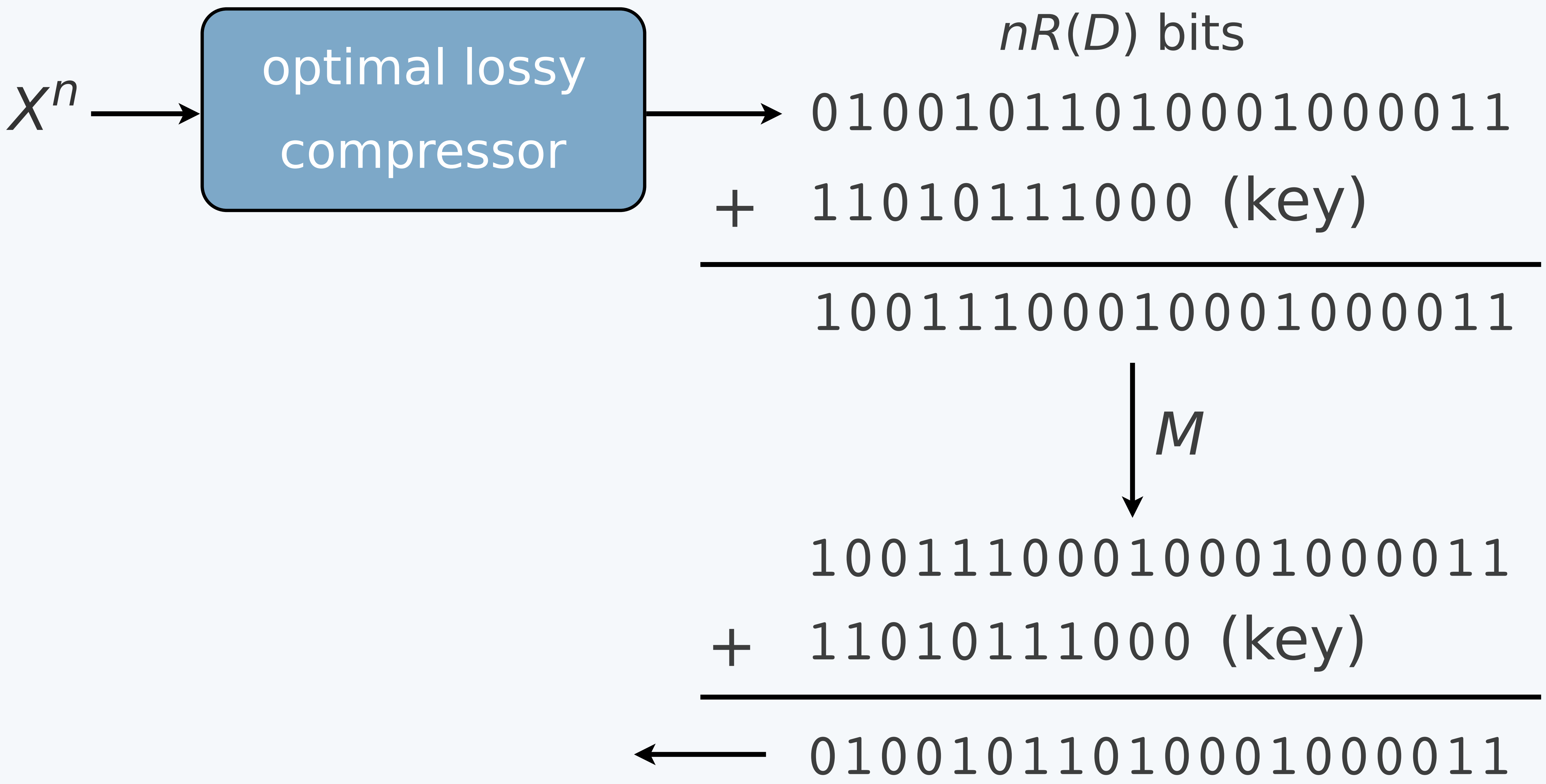
$nR(D)$ bits

$$
\begin{array}{rl}
 & 01001011010001000011 \\
+ & 11010111000 \text{ (key)} \\
\hline
 & 10011100010001000011
\end{array}
$$

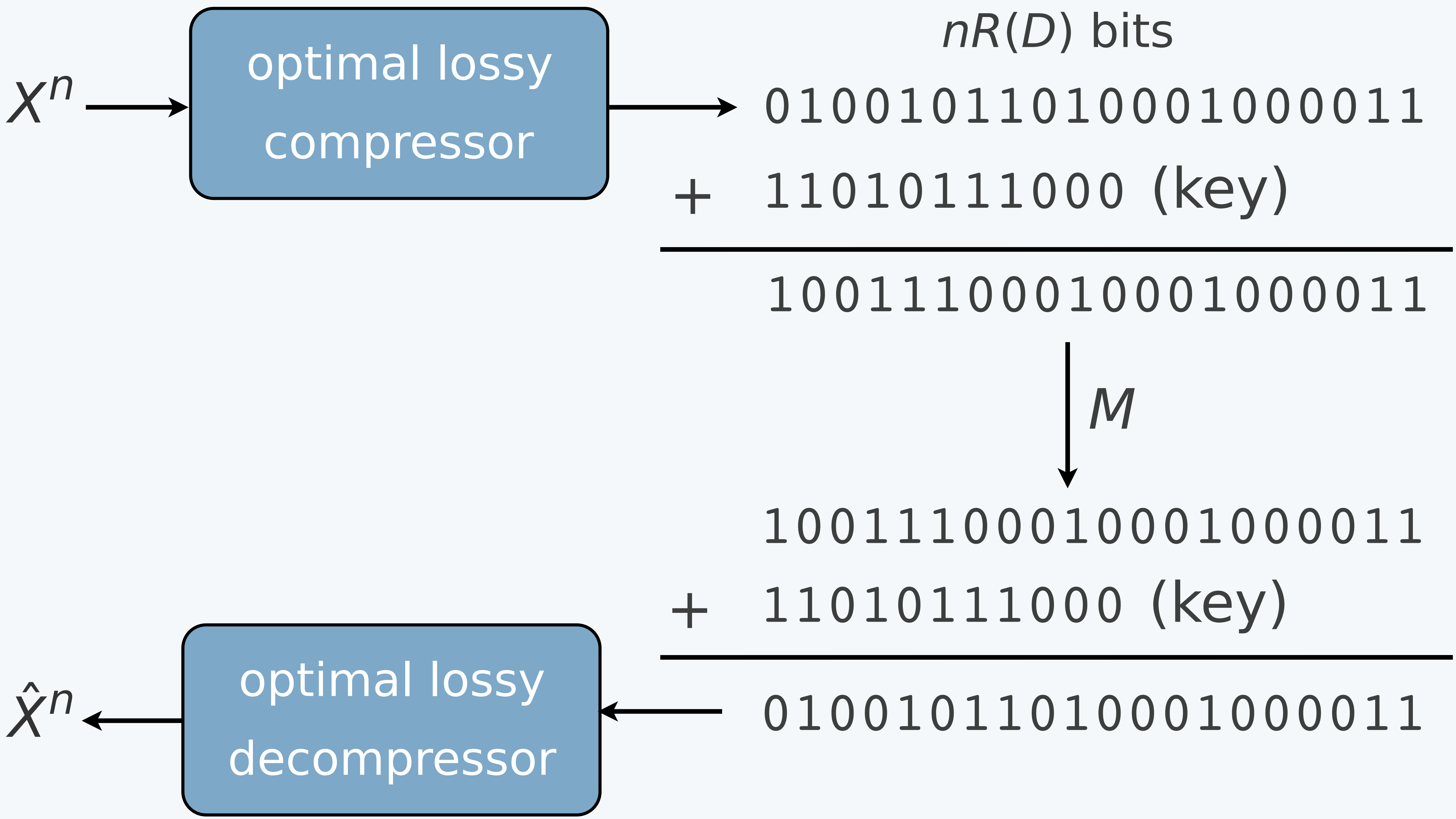
$\Big\downarrow M$

$$
\begin{array}{rl}
 & 10011100010001000011 \\
+ & 11010111000 \text{ (key)} \\
\hline
 & 01001011010001000011
\end{array}
$$

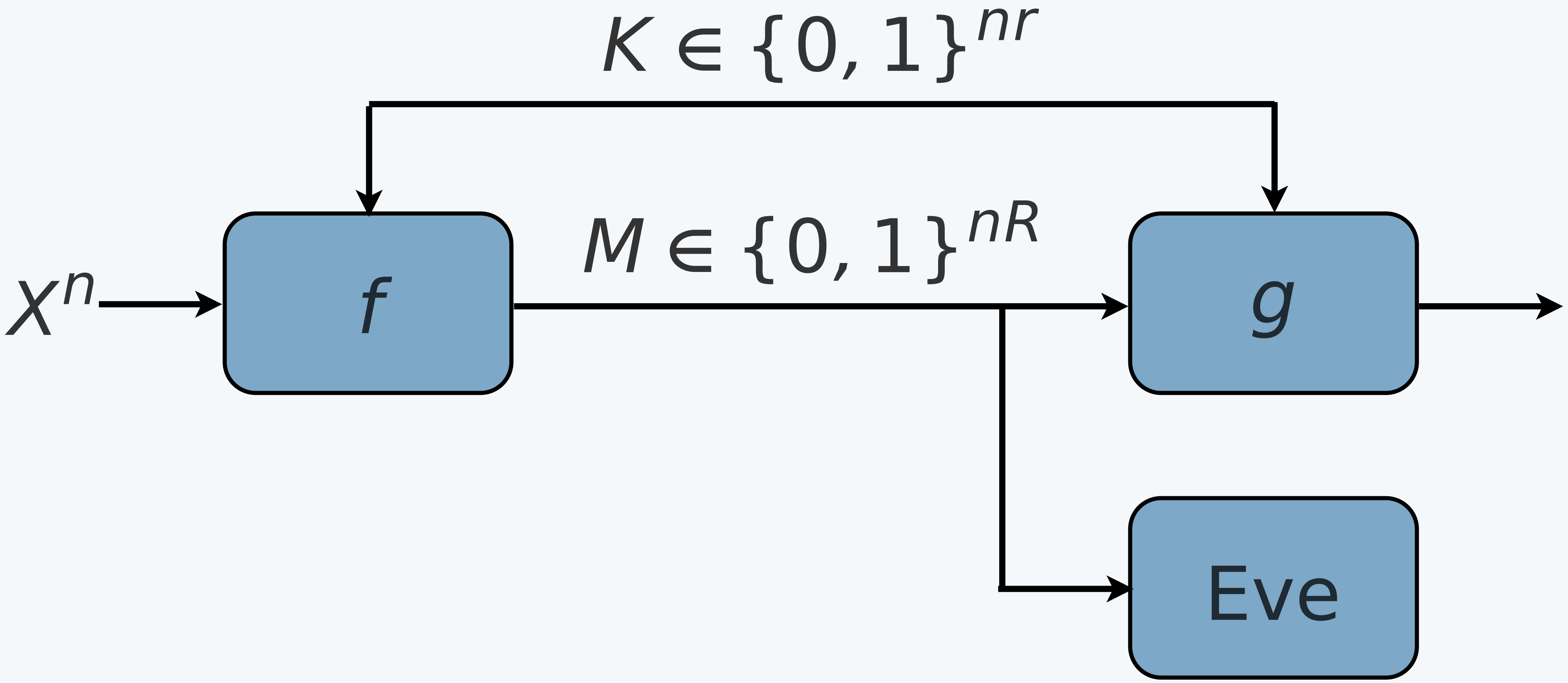
Achievability for Primary User



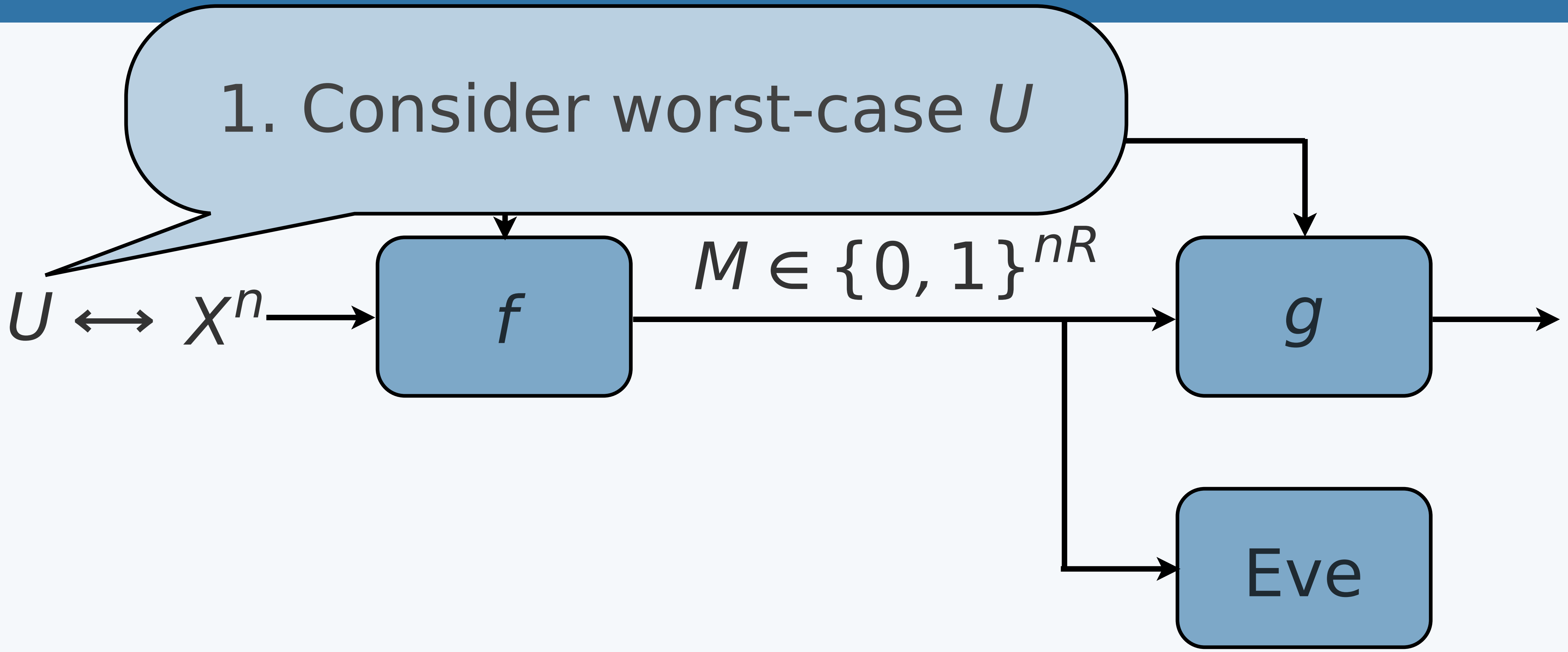
Achievability for Primary User



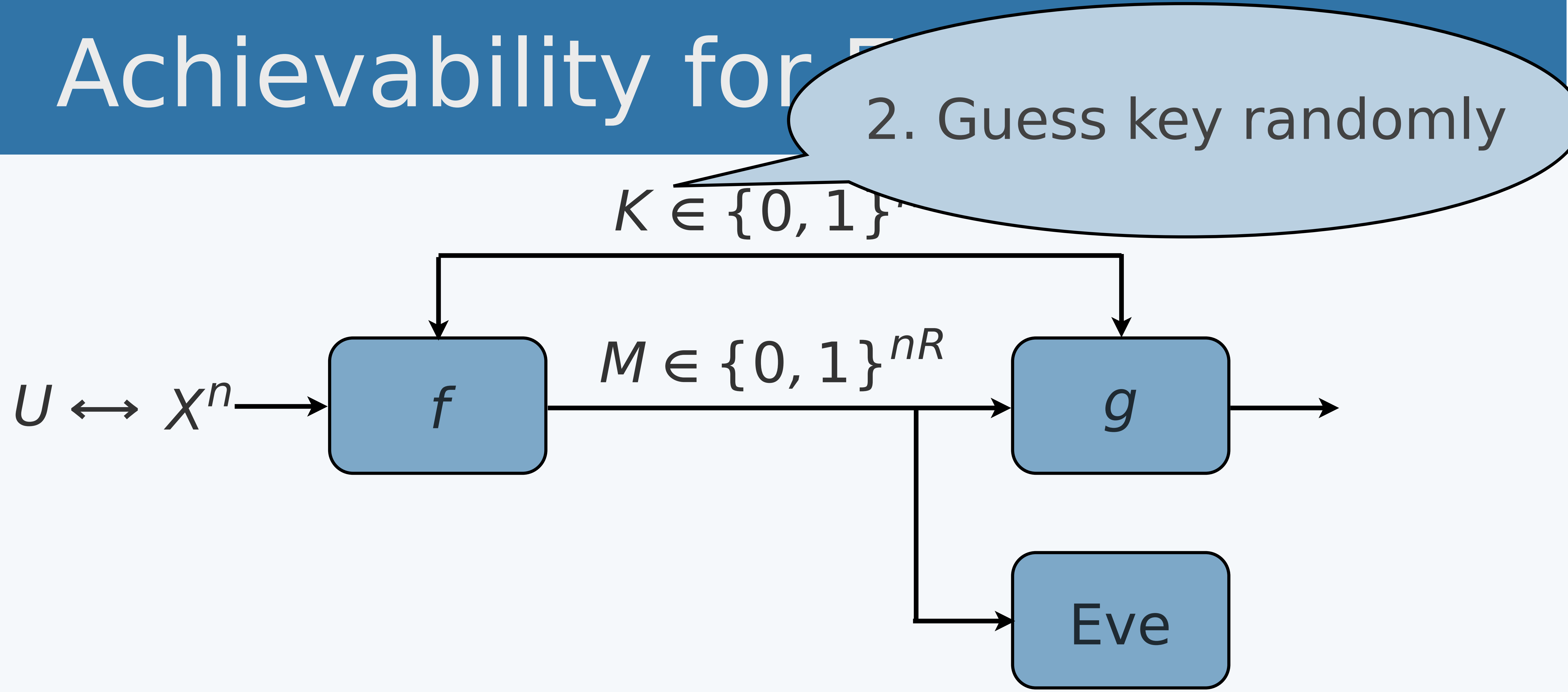
Achievability for Eavesdropper



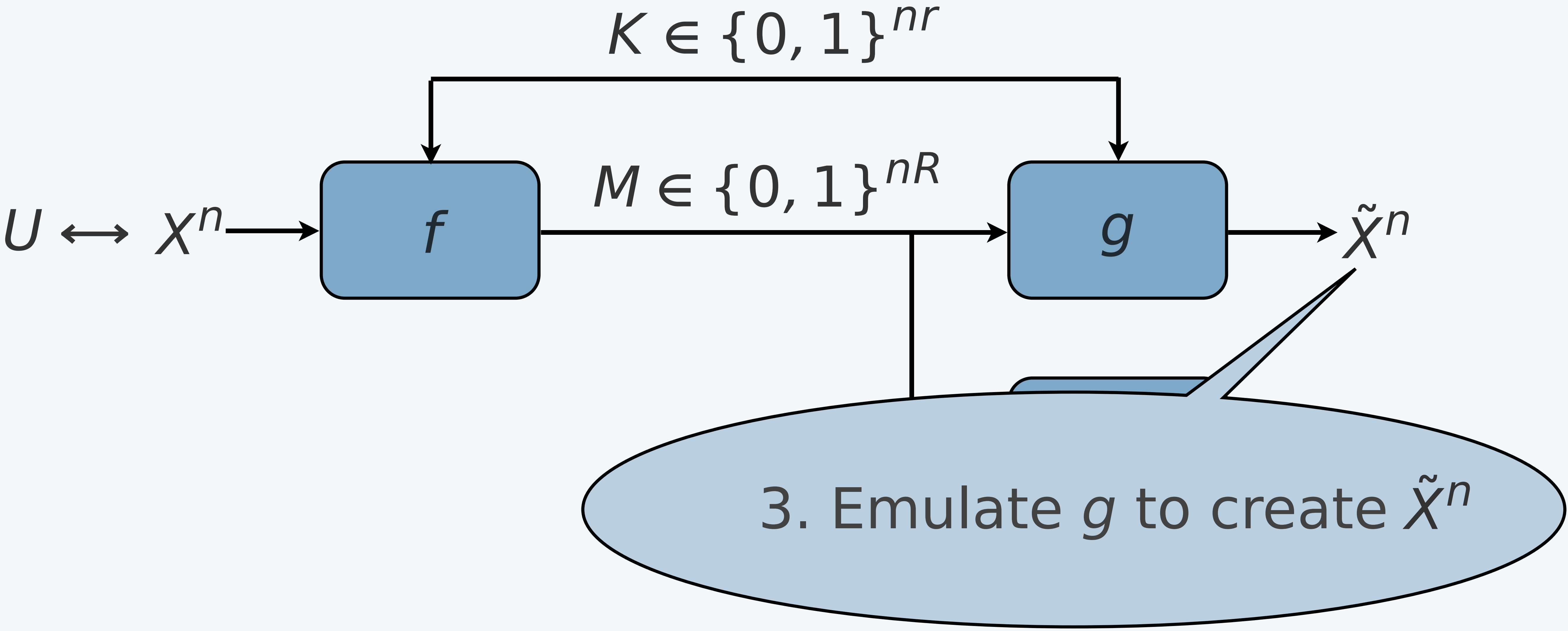
Achievability for Eavesdropper



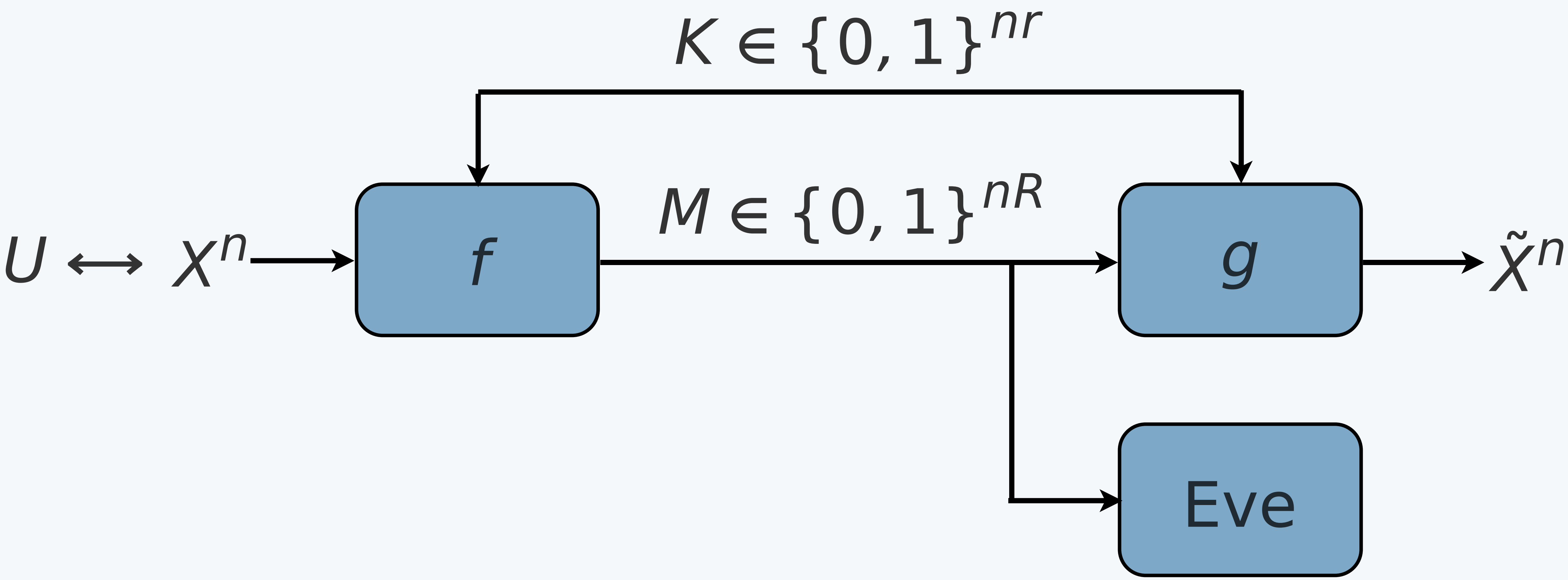
Achievability for E



Achievability for Eavesdropper

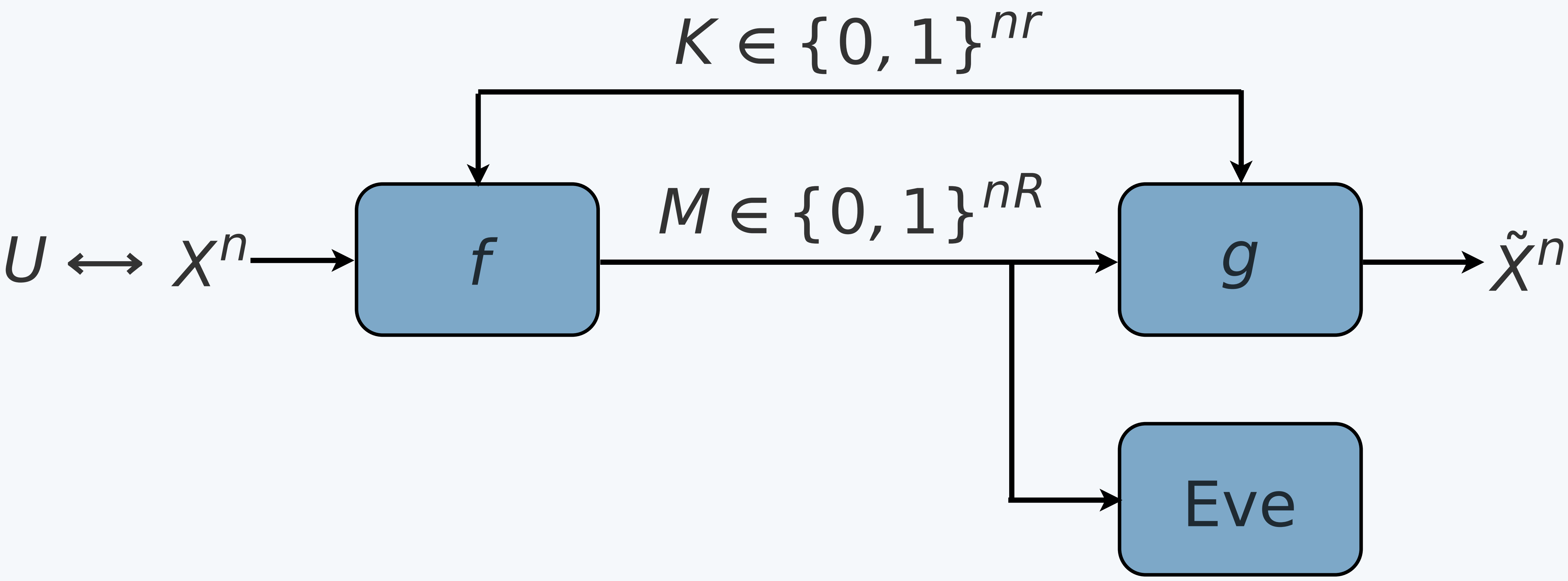


Achievability for Eavesdropper



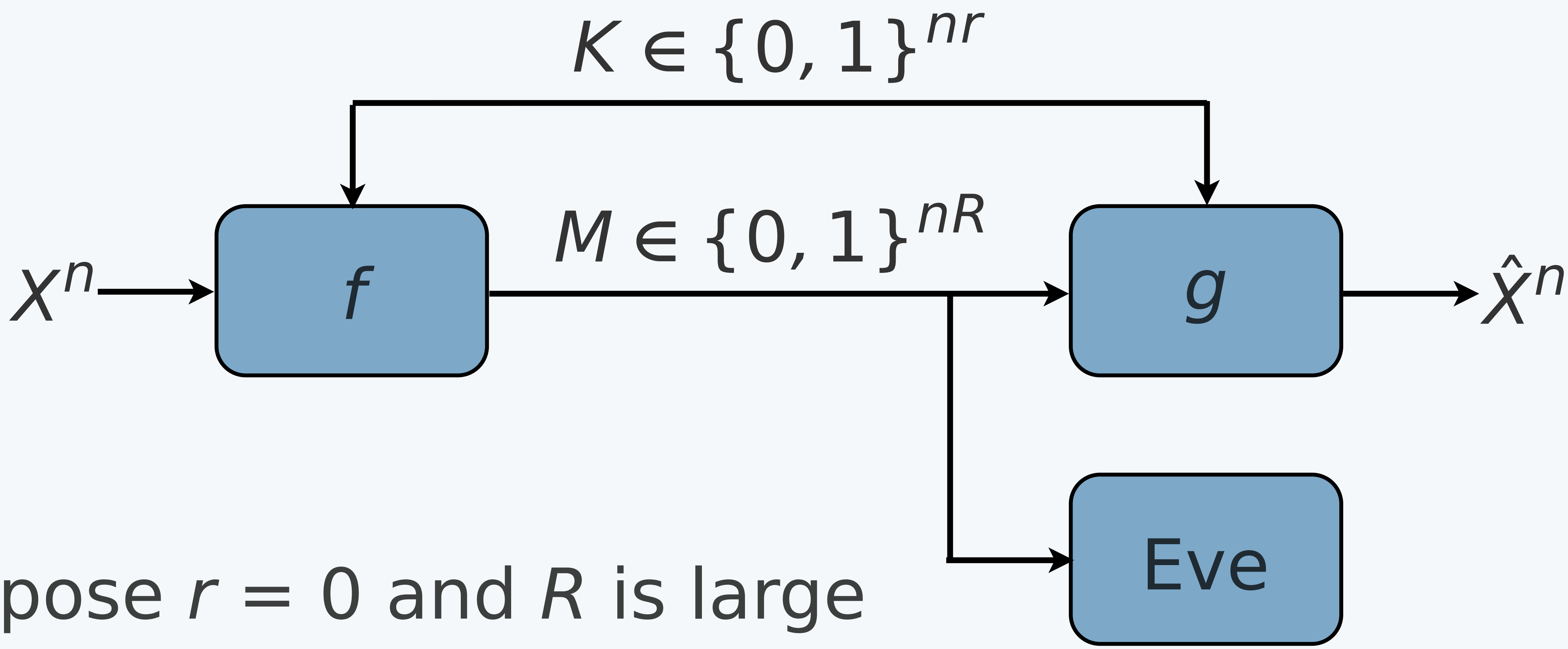
4. Pick X^n uniformly at random from within distortion ball around $\tilde{X}^n$.

Achievability for Eavesdropper



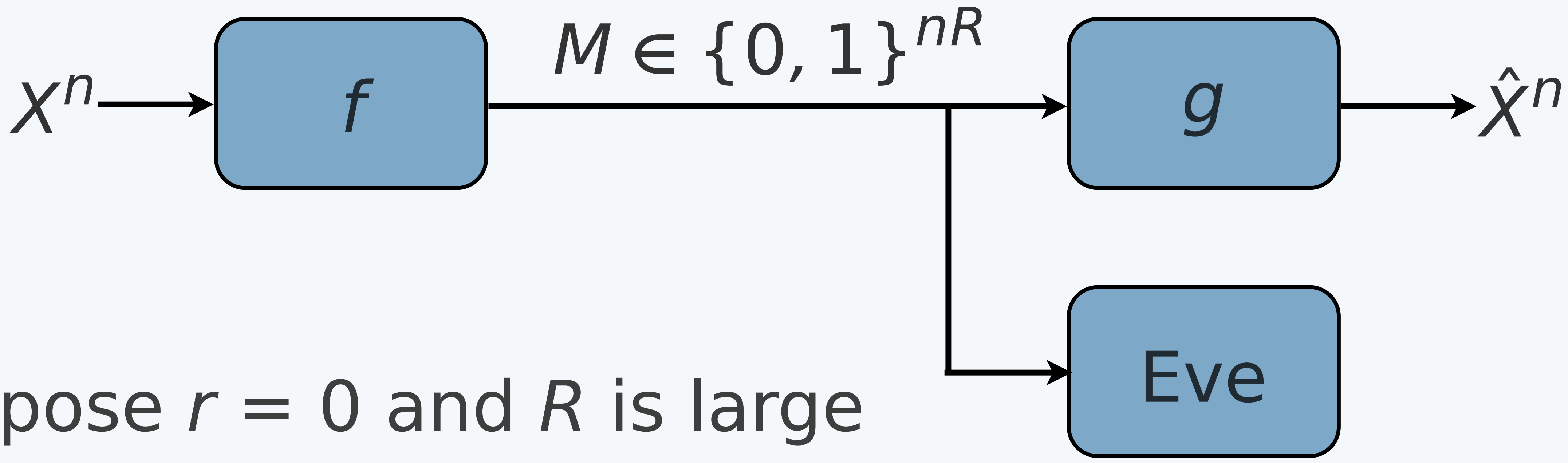
5. Generate U from X^n.

Quantization vs. Adding Noise



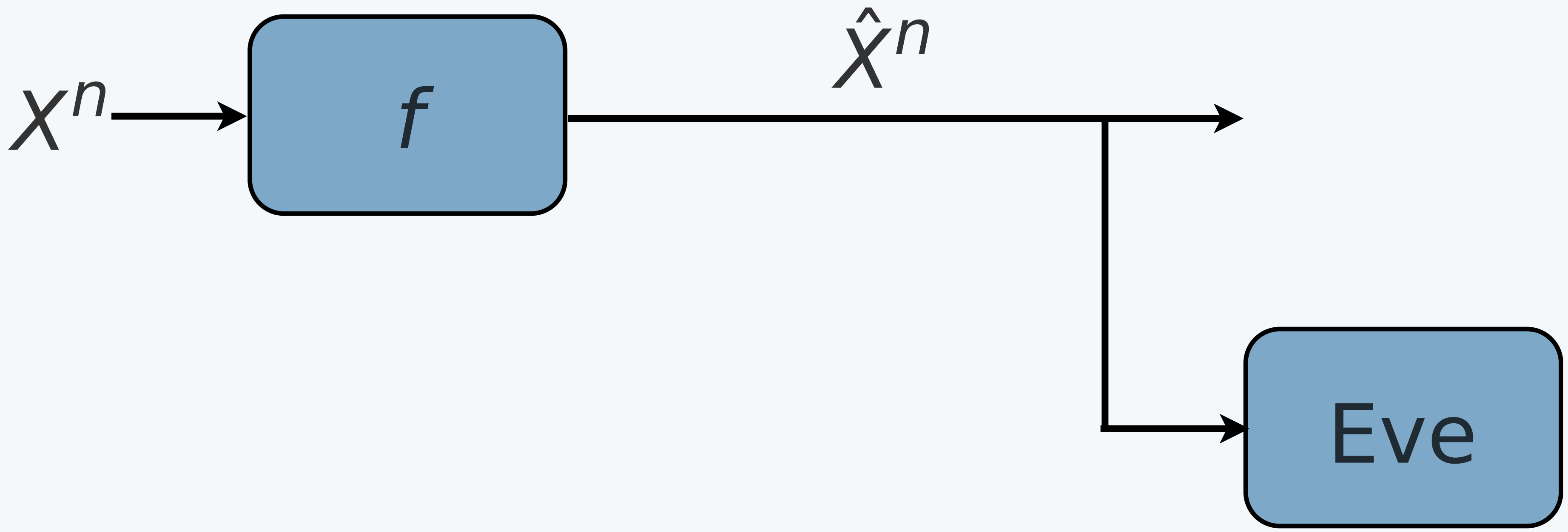
Suppose $r = 0$ and R is large

Quantization vs. Adding Noise

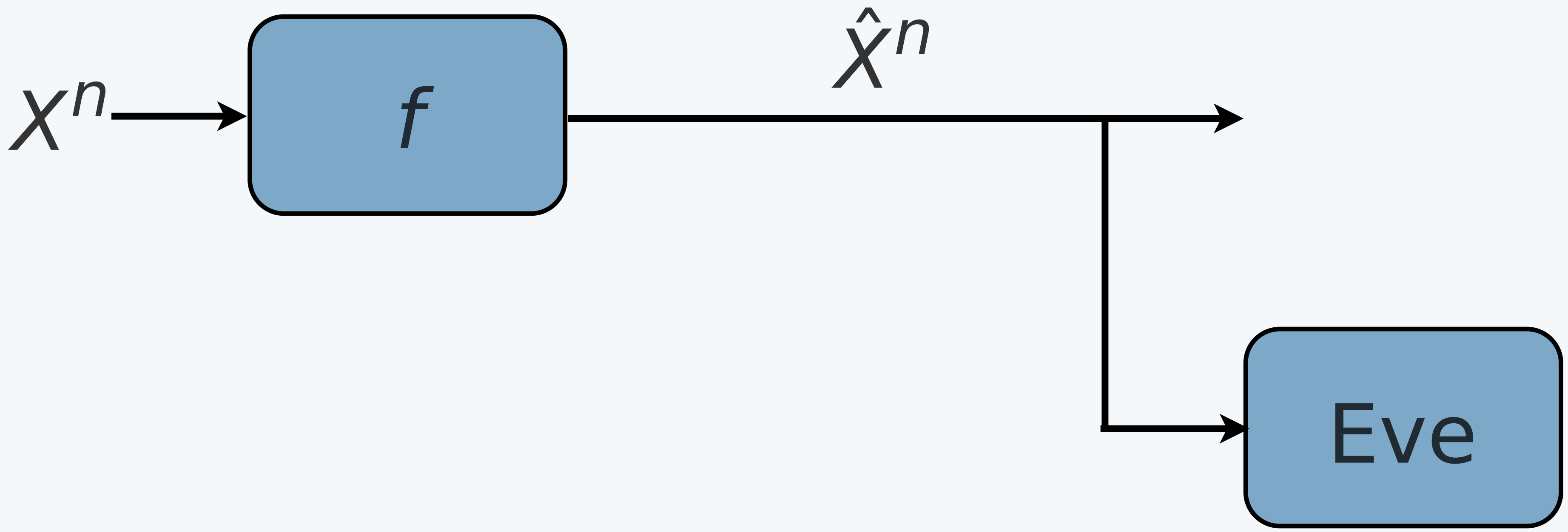


Suppose $r = 0$ and R is large

Quantization vs. Adding Noise



Quantization vs. Adding Noise

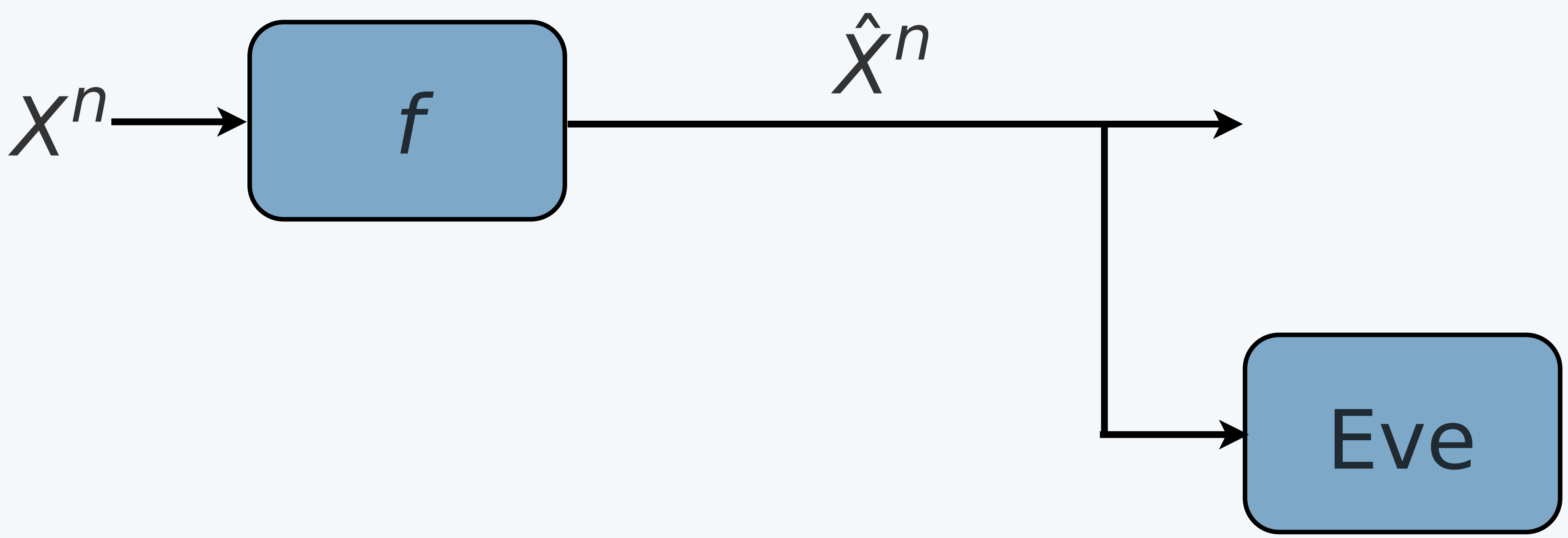


Then $\quad L_n = \min_{\hat{X}^n} \frac{1}{n} \cdot \mathcal{L}(X^n \to \hat{X}^n)$

subject to

$$\frac{1}{n} \sum_{i=1}^{n} E[d(X_i, \hat{X}_i)] \leq D$$

Quantization vs. Adding Noise

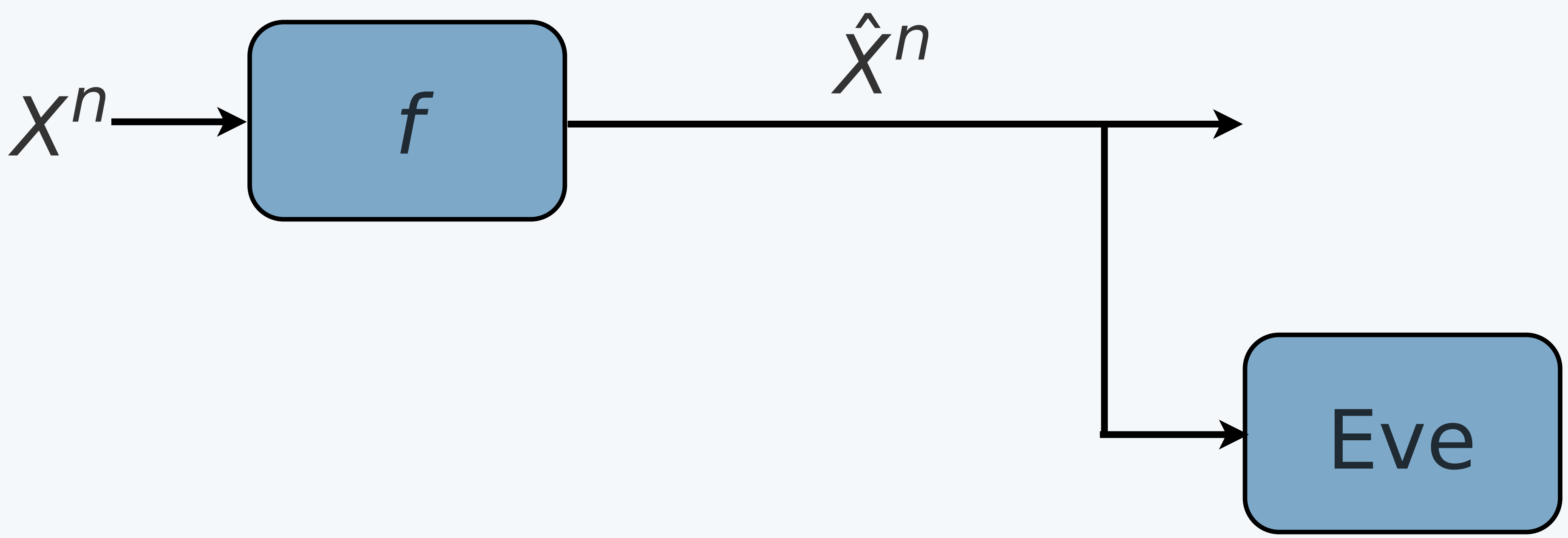


Then $\quad L_n = \min_{\hat{X}^n} \frac{1}{n} \cdot \mathcal{L}(X^n \to \hat{X}^n)$

subject to

$$L = \lim_{n\to\infty} L_n \qquad \frac{1}{n}\sum_{i=1}^{n} E[d(X_i, \hat{X}_i)] \leq D$$

Quantization vs. Adding Noise



Then $\quad L_n = \min_{\hat{X}^n} \frac{1}{n} \cdot \mathcal{L}(X^n \to \hat{X}^n)$

subject to

[side channel]

$$L = \lim_{n\to\infty} L_n \qquad \frac{1}{n}\sum_{i=1}^{n} E[d(X_i, \hat{X}_i)] \le D$$

Quantization vs. Adding Noise

Quantization vs. Adding Noise

Optimal scheme:

Quantization vs. Adding Noise

Optimal scheme:

- Compress X^n optimally to rate $R(D)$, then decompress.

Quantization vs. Adding Noise

Optimal scheme:

- Compress X^n optimally to rate $R(D)$, then decompress.
- Leaks $R(D)$ bits per symbol

Quantization vs. Adding Noise

Optimal scheme:

- Compress X^n optimally to rate $R(D)$, then decompress.
- Leaks $R(D)$ bits per symbol
- Deterministic but noncausal

Quantization vs. Adding Noise

Optimal scheme:

- Compress X^n optimally to rate $R(D)$, then decompress.
- Leaks $R(D)$ bits per symbol
- Deterministic but noncausal

Memoryless scheme:

$X_1 \longrightarrow \boxed{\text{Channel}_1} \longrightarrow \hat{X}_1$

$X_2 \longrightarrow \boxed{\text{Channel}_2} \longrightarrow \hat{X}_2$

$X_3 \longrightarrow \boxed{\text{Channel}_3} \longrightarrow \hat{X}_3$

$\vdots$

$X_n \longrightarrow \boxed{\text{Channel}_n} \longrightarrow \hat{X}_n$

Quantization vs. Adding Noise

Optimal scheme:

- Compress X^n optimally to rate $R(D)$, then decompress.
- Leaks $R(D)$ bits per symbol
- Deterministic but noncausal

Memoryless scheme:

$$X_1 \longrightarrow \text{Channel}_1 \longrightarrow \hat{X}_1$$
$$X_2 \longrightarrow \text{Channel}_2 \longrightarrow \hat{X}_2$$
$$X_3 \longrightarrow \text{Channel}_3 \longrightarrow \hat{X}_3$$
$$\vdots$$
$$X_n \longrightarrow \text{Channel}_n \longrightarrow \hat{X}_n$$

Memoryless scheme is causal but suboptimal.

Quantization vs. Adding Noise

Optimal scheme:

- Compress X^n optimally to rate $R(D)$, then decompress.
- Leaks $R(D)$ bits per symbol
- Deterministic but noncausal

Memoryless scheme:

$X_1 \rightarrow \text{Channel}_1 \rightarrow \hat{X}_1$

$X_2 \rightarrow \text{Channel}_2 \rightarrow \hat{X}_2$

$X_3 \rightarrow \text{Channel}_3 \rightarrow \hat{X}_3$

$\vdots$

$X_n \rightarrow \text{Channel}_n \rightarrow \hat{X}_n$

Memoryless scheme is causal but suboptimal.

[quantization is preferable to "adding noise"]

Quantization vs. Adding Noise

Optimal scheme:

- Compress X^n optimally to rate $R(D)$, then decompress.
- Leaks $R(D)$ bits per symbol
- Deterministic but noncausal

Memoryless scheme:

$X_1 \longrightarrow \text{Channel}_1 \longrightarrow \hat{X}_1$

$X_2 \longrightarrow \text{Channel}_2 \longrightarrow \hat{X}_2$

$X_3 \longrightarrow \text{Channel}_3 \longrightarrow \hat{X}_3$

$\vdots$

$X_n \longrightarrow \text{Channel}_n \longrightarrow \hat{X}_n$

Memoryless scheme is causal but suboptimal.

[quantization is preferable to "adding noise"]

[cf. mutual info.]

Extension: Approx. Guessing

Def (Issa-Kamath-Wagner)**:** For any metric space $\mathcal{U}$,

$$\mathcal{L}_{\mathcal{U}}(X \to Y) = \sup_{\substack{U: U \leftrightarrow X \leftrightarrow Y \\ \exists u: \Pr(U \in B(u)) > 0}} \log \frac{\sup_{\hat{u}(.)} \Pr(U \in B(\hat{u}(Y)))}{\sup_{\hat{u}} \Pr(U \in B(\hat{u}))}$$

Extension: Approx. Guessing

Def (Issa-Kamath-Wagner)**:** For any metric space $\mathcal{U}$,

$$\mathcal{L}_{\mathcal{U}}(X \to Y) = \sup_{\substack{U: U \leftrightarrow X \leftrightarrow Y \\ \exists u: \Pr(U \in B(u)) > 0}} \log \frac{\sup_{\hat{u}(.)} \Pr(U \in B(\hat{u}(Y)))}{\sup_{\hat{u}} \Pr(U \in B(\hat{u}))}$$

Theorem (Issa-Kamath-Wagner)**:** For any metric space $\mathcal{U}$,

$$\mathcal{L}_{\mathcal{U}}(X \to Y) \leq \mathcal{L}(X \to Y)$$

with equality if $\mathcal{U}$ has countably many points no two of which are contained in the same unit ball.

Extension: General Gains

Def (Issa-Kamath-Wagner)**:**

$$\mathcal{L}_G(X \to Y) = \sup_{\substack{U : U \leftrightarrow X \leftrightarrow Y \\ g(\cdot,\cdot) : \mathcal{U} \times \hat{\mathcal{U}} \mapsto [0,\infty) : \\ \sup_{\hat{u}} E[g(U,\hat{u})] > 0}} \log \frac{\sup_{\hat{u}(\cdot)} E[g(U, \hat{u}(Y))]}{\sup_{\hat{u}} E[g(U, \hat{u})]}$$

Extension: General Gains

Def (Issa-Kamath-Wagner)**:**

$$\mathcal{L}_G(X \to Y) = \sup_{\substack{U: U \leftrightarrow X \leftrightarrow Y \\ g(\cdot,\cdot): \mathcal{U} \times \hat{\mathcal{U}} \mapsto [0,\infty): \\ \sup_{\hat{u}} E[g(U,\hat{u})] > 0}} \log \frac{\sup_{\hat{u}(\cdot)} E[g(U, \hat{u}(Y))]}{\sup_{\hat{u}} E[g(U, \hat{u})]}$$

Theorem (Issa-Kamath-Wagner)**:** If X and Y are discrete, then

$$\mathcal{L}_G(X \to Y) = \mathcal{L}(X \to Y).$$

Opportunistic Attacks

Definition: The opportunistic maximal leakage is

$$\mathcal{L}_O(X \to Y) = \log E_Y \left[\sup_{U \leftrightarrow X \leftrightarrow Y} \frac{\sup_{\tilde{u}} P_{U|Y}(\tilde{u}|y)}{\sup_{\tilde{u}} P(\tilde{u})} \right]$$

Opportunistic Attacks

Definition: The opportunistic maximal leakage is

$$\mathcal{L}_O(X \to Y) = \log E_Y \left[\sup_{U \leftrightarrow X \leftrightarrow Y} \frac{\sup_{\tilde{u}} P_{U|Y}(\tilde{u}|y)}{\sup_{\tilde{u}} P(\tilde{u})} \right]$$

Theorem (Issa-Wagner): For any joint distribution P_{XY} on finite alphabets

$$\mathcal{L}_O(X \to Y) = \mathcal{L}(X \to Y)$$

Extension: General Alphabet

Corollary (IKW)**:** If X and Y are jointly continuous then

$$\mathcal{L}(X \to Y) = \log \int \sup_{x: f_X(x) > 0} f_{Y|X}(y|x)\, dy$$

Extension: General Alphabet

Corollary (IKW)**:** If X and Y are jointly continuous then

$$\mathcal{L}(X \to Y) = \log \int \sup_{x : f_X(x) > 0} f_{Y|X}(y|x)\, dy$$

Corollary (IKW)**:** If X and Y are jointly Gaussian then

$$\mathcal{L}(X \to Y) = \begin{cases} 0 & \text{if } X, Y \text{ indep.} \\ \infty & \text{otherwise} \end{cases}$$

Extension: General Alphabet

Extension: General Alphabet

Theorem (IKW '17)**:** Let $(\mathcal{X} \times \mathcal{Y}, \sigma_{X \times Y}, P_{XY})$ be a prob. space with associated prob. spaces $(\mathcal{X}, \sigma_X, P_X)$ and $(\mathcal{Y}, \sigma_Y, P_Y)$.

Extension: General Alphabet

Theorem (IKW '17)**:** Let $(\mathcal{X} \times \mathcal{Y}, \sigma_{X \times Y}, P_{XY})$ be a prob. space with associated prob. spaces $(\mathcal{X}, \sigma_X, P_X)$ and $(\mathcal{Y}, \sigma_Y, P_Y)$.

- If $P_{XY} \ll P_X \times P_Y$ and $\sigma_{\mathcal{X}}$ is generated by a countable set then

$$\mathcal{L}(X \to Y) = \log \int_{\mathcal{Y}} \operatorname{ess\,sup}_x \left\{ \frac{dP_{XY}}{dP_X \times dP_Y}(x, y) \right\} dP_Y$$

Extension: General Alphabet

Theorem (IKW ’17)**:** Let $(\mathcal{X} \times \mathcal{Y}, \sigma_{X \times Y}, P_{XY})$ be a prob. space with associated prob. spaces $(\mathcal{X}, \sigma_X, P_X)$ and $(\mathcal{Y}, \sigma_Y, P_Y)$.

- If $P_{XY} \ll P_X \times P_Y$ and $\sigma_{\mathcal{X}}$ is generated by a countable set then

$$\mathcal{L}(X \to Y) = \log \int_{\mathcal{Y}} \operatorname{ess\,sup}_x \left\{ \frac{dP_{XY}}{dP_X \times dP_Y}(x, y) \right\} dP_Y$$

Extension: General Alphabet

Theorem (IKW '17): Let $(\mathcal{X} \times \mathcal{Y}, \sigma_{X \times Y}, P_{XY})$ be a prob. space with associated prob. spaces $(\mathcal{X}, \sigma_X, P_X)$ and $(\mathcal{Y}, \sigma_Y, P_Y)$.

- If $P_{XY} \ll P_X \times P_Y$ and $\sigma_{\mathcal{X}}$ is generated by a countable set then

$$\mathcal{L}(X \to Y) = \log \int_{\mathcal{Y}} \text{ess sup}_x \left\{ \frac{dP_{XY}}{dP_X \times dP_Y}(x, y) \right\} dP_Y$$

Extension: General Alphabet

Theorem (IKW '17)**:** Let $(\mathcal{X} \times \mathcal{Y}, \sigma_{X \times Y}, P_{XY})$ be a prob. space with associated prob. spaces $(\mathcal{X}, \sigma_X, P_X)$ and $(\mathcal{Y}, \sigma_Y, P_Y)$.

- If $P_{XY} \ll P_X \times P_Y$ and $\sigma_{\mathcal{X}}$ is generated by a countable set then

$$\mathcal{L}(X \to Y) = \log \int_{\mathcal{Y}} \text{ess sup}_x \left\{ \frac{dP_{XY}}{dP_X \times dP_Y}(x, y) \right\} dP_Y$$

Extension: General Alphabet

Theorem (IKW '17)**:** Let $(\mathcal{X} \times \mathcal{Y}, \sigma_{X \times Y}, P_{XY})$ be a prob. space with associated prob. spaces $(\mathcal{X}, \sigma_X, P_X)$ and $(\mathcal{Y}, \sigma_Y, P_Y)$.

- If $P_{XY} \ll P_X \times P_Y$ and $\sigma_{\mathcal{X}}$ is generated by a countable set then

$$\mathcal{L}(X \to Y) = \log \int_{\mathcal{Y}} \text{ess sup}_x \left\{ \frac{dP_{XY}}{dP_X \times dP_Y}(x, y) \right\} dP_Y$$

Extension: General Alphabet

Theorem (IKW '17)**:** Let $(\mathcal{X} \times \mathcal{Y}, \sigma_{X \times Y}, P_{XY})$ be a prob. space with associated prob. spaces $(\mathcal{X}, \sigma_X, P_X)$ and $(\mathcal{Y}, \sigma_Y, P_Y)$.

- If $P_{XY} \ll P_X \times P_Y$ and $\sigma_{\mathcal{X}}$ is generated by a countable set then

$$\mathcal{L}(X \to Y) = \log \int_{\mathcal{Y}} \text{ess sup}_x \left\{ \frac{dP_{XY}}{dP_X \times dP_Y}(x, y) \right\} dP_Y$$

- If then

Extension: General Alphabet

Theorem (IKW '17)**:** Let $(\mathcal{X} \times \mathcal{Y}, \sigma_{X\times Y}, P_{XY})$ be a prob. space with associated prob. spaces $(\mathcal{X}, \sigma_X, P_X)$ and $(\mathcal{Y}, \sigma_Y, P_Y)$.

- If $P_{XY} \ll P_X \times P_Y$ and $\sigma_{\mathcal{X}}$ is generated by a countable set then

$$\mathcal{L}(X \to Y) = \log \int_{\mathcal{Y}} \text{ess sup}_x \left\{ \frac{dP_{XY}}{dP_X \times dP_Y}(x, y) \right\} dP_Y$$

- If $P_{XY} \not\ll P_X \times P_y$ then

$$\mathcal{L}(X \to Y) = \infty$$

Extension: General Alphabet

Extension: General Alphabet

Theorem (IKW '17)**:** Let $(\mathcal{X} \times \mathcal{Y}, \sigma_{X \times Y}, P_{XY})$ be a prob. space with associated prob. spaces $(\mathcal{X}, \sigma_X, P_X)$ and $(\mathcal{Y}, \sigma_Y, P_Y)$.

Extension: General Alphabet

Theorem (IKW '17)**:** Let $(\mathcal{X} \times \mathcal{Y}, \sigma_{X \times Y}, P_{XY})$ be a prob. space with associated prob. spaces $(\mathcal{X}, \sigma_X, P_X)$ and $(\mathcal{Y}, \sigma_Y, P_Y)$.

- If $P_{XY} \ll P_X \times P_Y$ and $\sigma_{\mathcal{X}}$ is generated by a countable set then

$$\mathcal{L}(X \to Y) = \log \int_{\mathcal{Y}} \operatorname{ess\,sup}_x \left\{ \frac{dP_{XY}}{dP_X \times dP_Y}(x, y) \right\} dP_Y$$

Extension: General Alphabet

Theorem (IKW '17)**:** Let $(\mathcal{X} \times \mathcal{Y}, \sigma_{X \times Y}, P_{XY})$ be a prob. space with associated prob. spaces $(\mathcal{X}, \sigma_X, P_X)$ and $(\mathcal{Y}, \sigma_Y, P_Y)$.

- If $P_{XY} \ll P_X \times P_Y$ and $\sigma_{\mathcal{X}}$ is generated by a countable set then

$$\mathcal{L}(X \to Y) = \log \int_{\mathcal{Y}} \text{ess sup}_x \left\{ \frac{dP_{XY}}{dP_X \times dP_Y}(x, y) \right\} dP_Y$$

Extension: General Alphabet

Theorem (IKW '17): Let $(\mathcal{X} \times \mathcal{Y}, \sigma_{X \times Y}, P_{XY})$ be a prob. space with associated prob. spaces $(\mathcal{X}, \sigma_X, P_X)$ and $(\mathcal{Y}, \sigma_Y, P_Y)$.

- If $P_{XY} \ll P_X \times P_Y$ and $\sigma_{\mathcal{X}}$ is generated by a countable set then

$$\mathcal{L}(X \to Y) = \log \int_{\mathcal{Y}} \text{ess sup}_x \left\{ \frac{dP_{XY}}{dP_X \times dP_Y}(x, y) \right\} dP_Y$$

Extension: General Alphabet

Theorem (IKW '17)**:** Let $(\mathcal{X} \times \mathcal{Y}, \sigma_{X \times Y}, P_{XY})$ be a prob. space with associated prob. spaces $(\mathcal{X}, \sigma_X, P_X)$ and $(\mathcal{Y}, \sigma_Y, P_Y)$.

- If $P_{XY} \ll P_X \times P_Y$ and $\sigma_{\mathcal{X}}$ is generated by a countable set then

$$\mathcal{L}(X \to Y) = \log \int_{\mathcal{Y}} \text{ess sup}_x \left\{ \frac{dP_{XY}}{dP_X \times dP_Y}(x, y) \right\} dP_Y$$

Extension: General Alphabet

Theorem (IKW '17)**:** Let $(\mathcal{X} \times \mathcal{Y}, \sigma_{X \times Y}, P_{XY})$ be a prob. space with associated prob. spaces $(\mathcal{X}, \sigma_X, P_X)$ and $(\mathcal{Y}, \sigma_Y, P_Y)$.

- If $P_{XY} \ll P_X \times P_Y$ and $\sigma_{\mathcal{X}}$ is generated by a countable set then

$$\mathcal{L}(X \to Y) = \log \int_{\mathcal{Y}} \text{ess sup}_x \left\{ \frac{dP_{XY}}{dP_X \times dP_Y}(x, y) \right\} dP_Y$$

Extension: General Alphabet

Theorem (IKW '17)**:** Let $(\mathcal{X} \times \mathcal{Y}, \sigma_{X\times Y}, P_{XY})$ be a prob. space with associated prob. spaces $(\mathcal{X}, \sigma_X, P_X)$ and $(\mathcal{Y}, \sigma_Y, P_Y)$.

- If $P_{XY} \ll P_X \times P_Y$ and $\sigma_{\mathcal{X}}$ is generated by a countable set then

$$\mathcal{L}(X \to Y) = \log \int_{\mathcal{Y}} \text{ess sup}_x \left\{ \frac{dP_{XY}}{dP_X \times dP_Y}(x, y) \right\} dP_Y$$

- If then

Extension: General Alphabet

Theorem (IKW '17)**:** Let $(\mathcal{X}\times\mathcal{Y}, \sigma_{X\times Y}, P_{XY})$ be a prob. space with associated prob. spaces $(\mathcal{X}, \sigma_X, P_X)$ and $(\mathcal{Y}, \sigma_Y, P_Y)$.

- If $P_{XY} \ll P_X \times P_Y$ and $\sigma_{\mathcal{X}}$ is generated by a countable set then

$$\mathcal{L}(X \to Y) = \log \int_{\mathcal{Y}} \operatorname{ess\,sup}_x \left\{ \frac{dP_{XY}}{dP_X \times dP_Y}(x,y) \right\} dP_Y$$

- If $P_{XY} \not\ll P_X \times P_y$ then

$$\mathcal{L}(X \to Y) = \infty$$