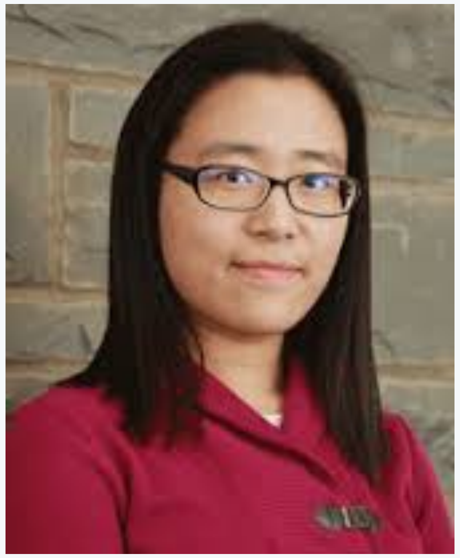
Must One Learn the Channel to Communicate at Capacity?

Aaron Wagner
Cornell University



Joint work with Yuguang Gao

The Capacity Theorem

Discrete memoryless channel: $W(y|x)$ $x \in \mathcal{X}$

$y \in \mathcal{Y}$

$$W(y|x)$$

The Capacity Theorem

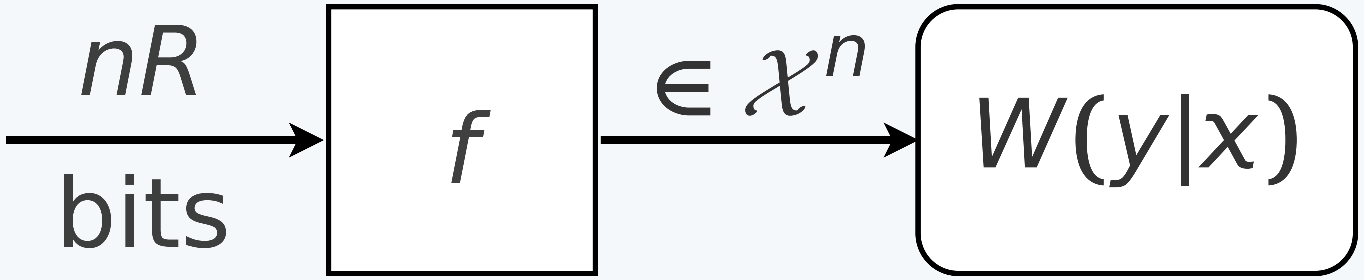
Discrete memoryless channel: $W(y|x)$ $x \in \mathcal{X}$

$y \in \mathcal{Y}$

The Capacity Theorem

Discrete memoryless channel: $W(y|x)$ $x \in \mathcal{X}$

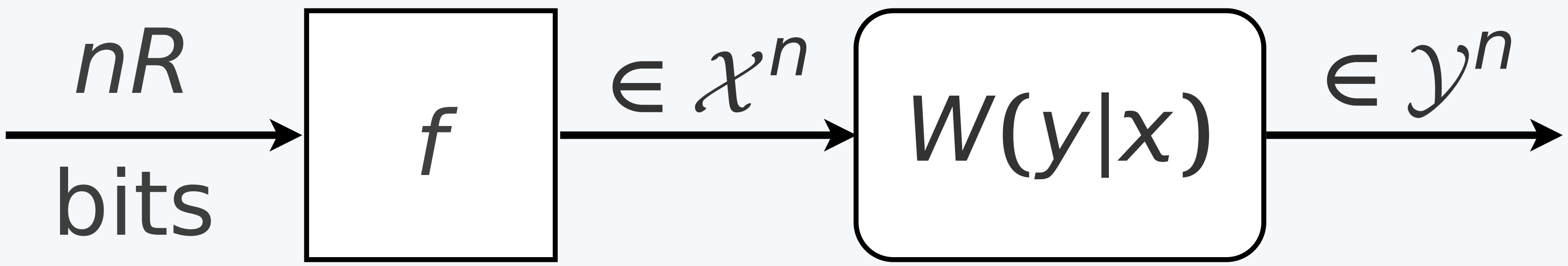
$y \in \mathcal{Y}$



The Capacity Theorem

Discrete memoryless channel: $W(y|x)$ $\quad x \in \mathcal{X}$

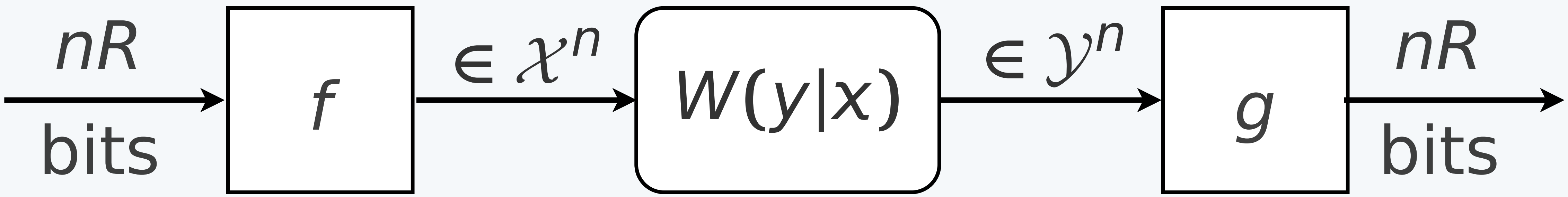
$y \in \mathcal{Y}$



The Capacity Theorem

Discrete memoryless channel: $W(y|x)$ $x \in \mathcal{X}$

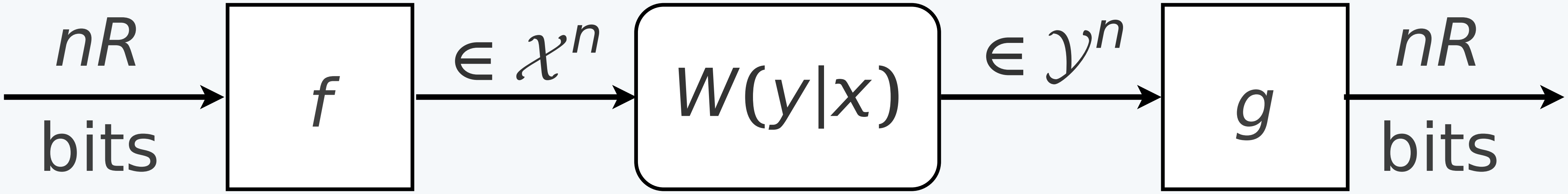
$y \in \mathcal{Y}$

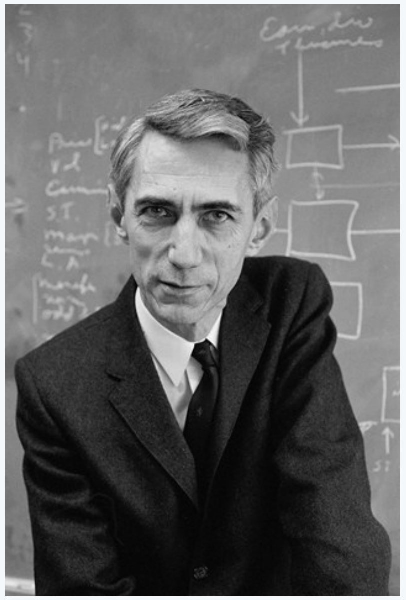


The Capacity Theorem

Discrete memoryless channel: $W(y|x)$ $x \in \mathcal{X}$

$y \in \mathcal{Y}$



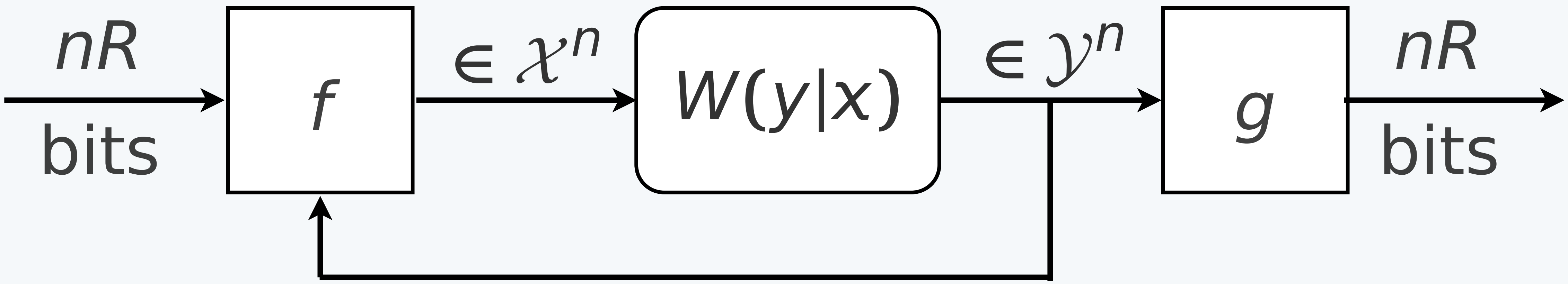


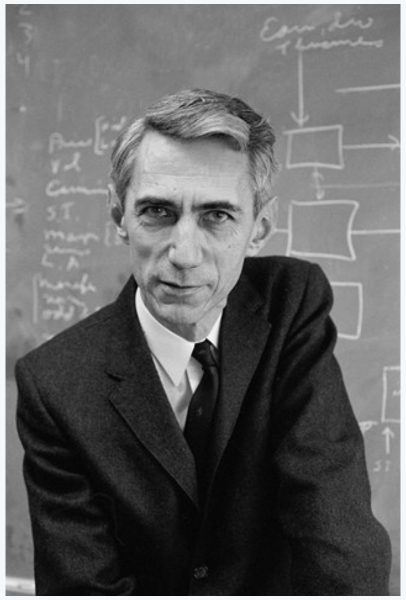
capacity $C = \max_{p(x)} I(p(x); W(y|x))$

The Capacity Theorem

Discrete memoryless channel: $W(y|x)$ $\quad x \in \mathcal{X}$

$y \in \mathcal{Y}$



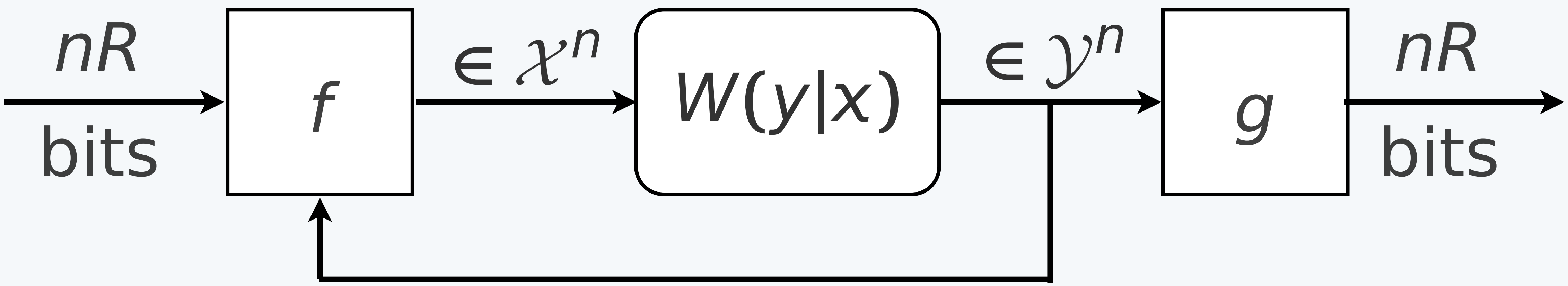


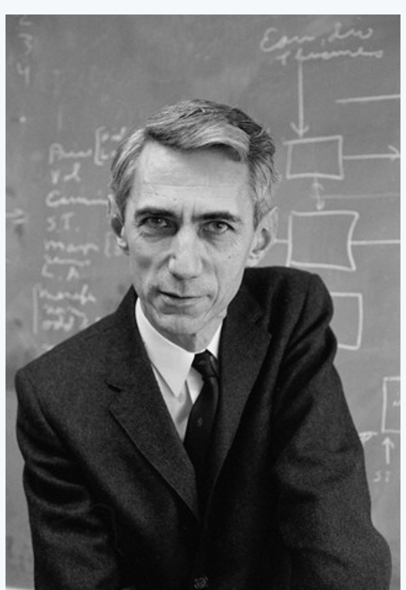
capacity $C = \max_{p(x)} I(p(x); W(y|x))$

The Capacity Theorem

Discrete memoryless channel: $W(y|x)$ $\quad x \in \mathcal{X}$

$y \in \mathcal{Y}$

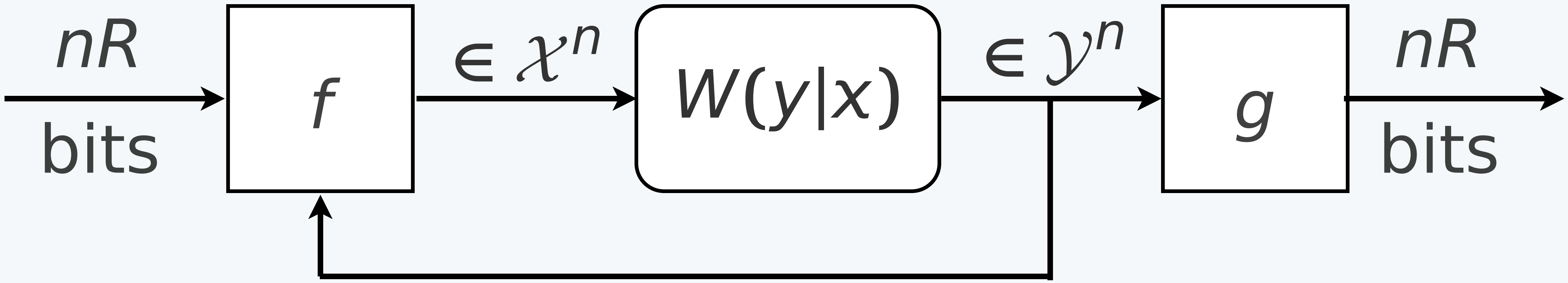




capacity $C = \max_{p(x)} I(p(x); W(y|x))$

[known channel]

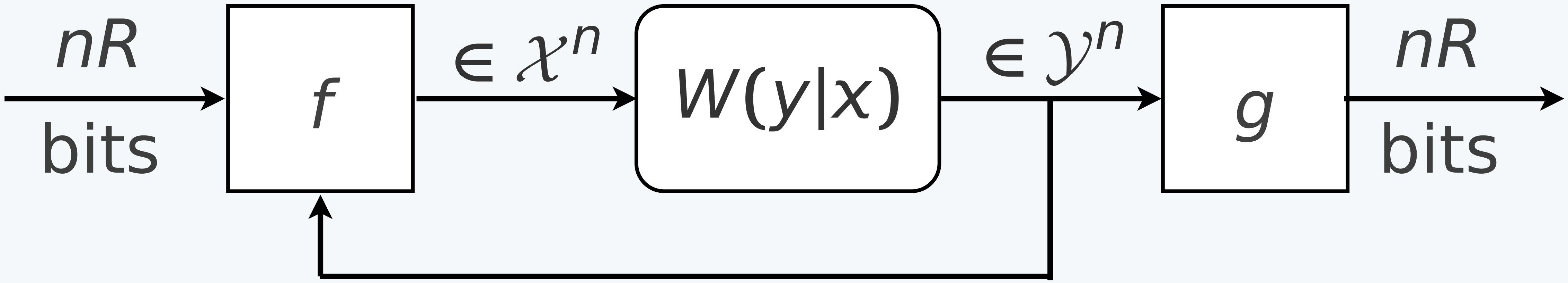
Channel Coding as Minimax



Theorem: (folk)

For any $\mathcal{X}$, $\mathcal{Y}$, R, and $\epsilon > 0$,

Channel Coding as Minimax

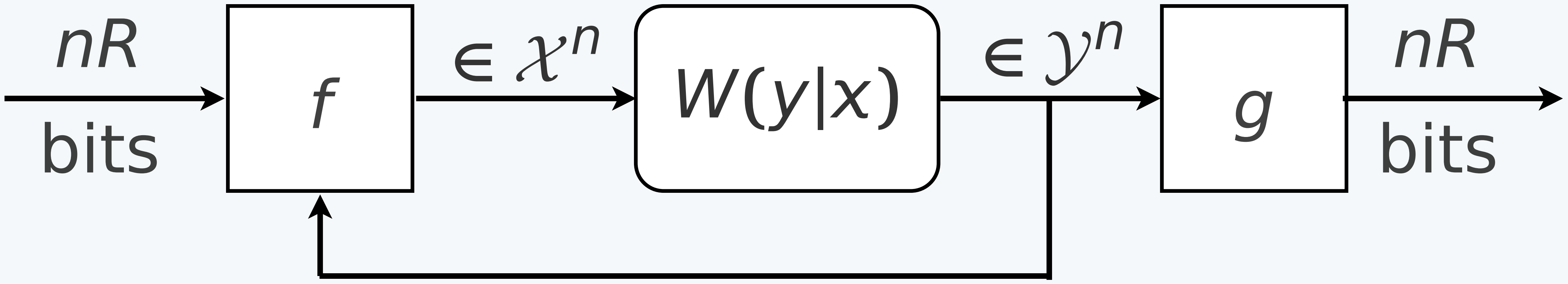


Theorem: (folk)

For any $\mathcal{X}$, $\mathcal{Y}$, R, and $\epsilon > 0$,

$$\inf_{f,g:\text{rate} \geq R - \epsilon}$$

Channel Coding as Minimax

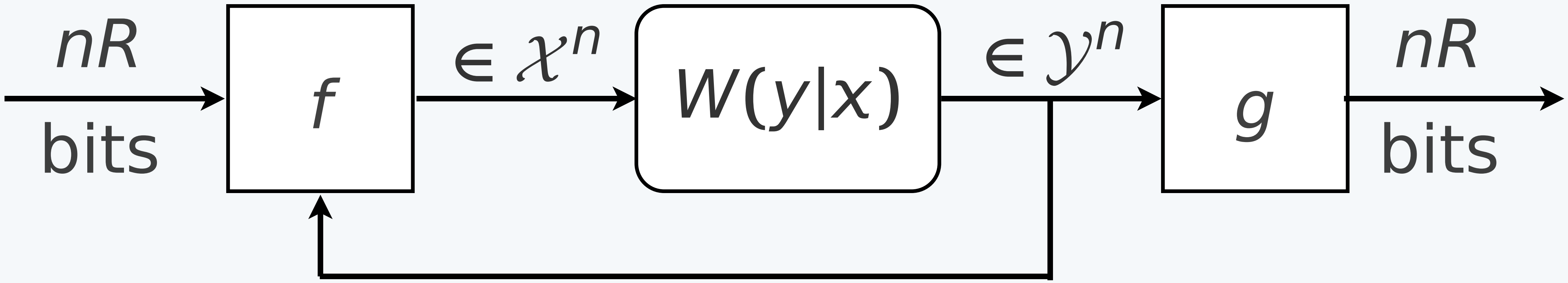


Theorem: (folk)

For any $\mathcal{X}$, $\mathcal{Y}$, R, and $\epsilon > 0$,

$$\inf_{f,g:\text{rate}\geq R-\epsilon} \qquad \mathrm{P_e}(f, g, W)$$

Channel Coding as Minimax

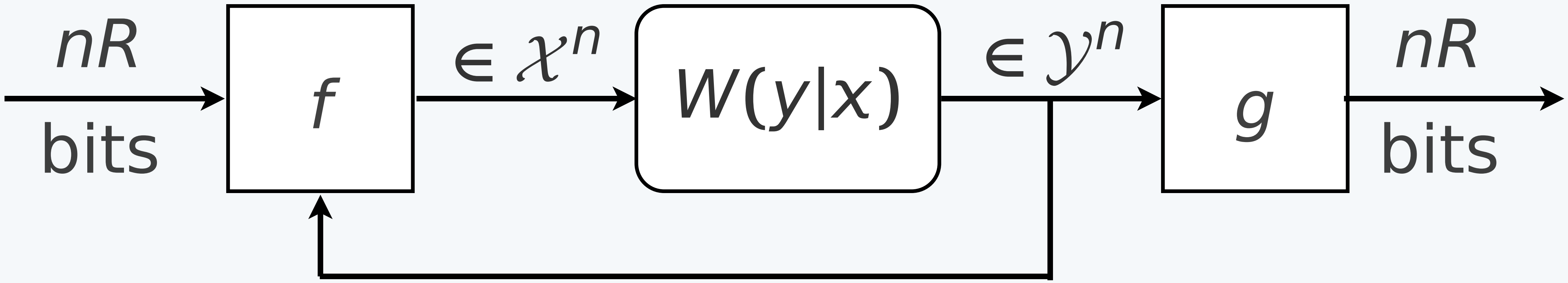


Theorem: (folk)

For any $\mathcal{X}$, $\mathcal{Y}$, R, and $\epsilon > 0$,

$$\inf_{f,g:\text{rate}\geq R-\epsilon} \quad \sup_{W:\mathcal{X}\to\mathcal{Y}:C(W)\geq R} \quad \mathrm{P_e}(f, g, W)$$

Channel Coding as Minimax

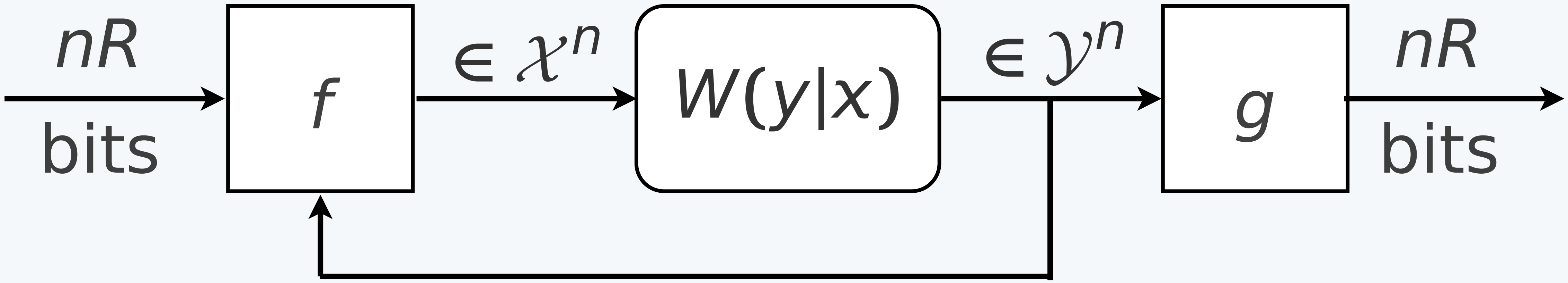


Theorem: (folk)

For any $\mathcal{X}$, $\mathcal{Y}$, R, and $\epsilon > 0$,

$$\lim_{n\to\infty} \inf_{f,g:\text{rate}\geq R-\epsilon} \sup_{W:\mathcal{X}\to\mathcal{Y}:C(W)\geq R} \mathrm{P_e}(f,g,W) = 0$$

Channel Coding as Minimax



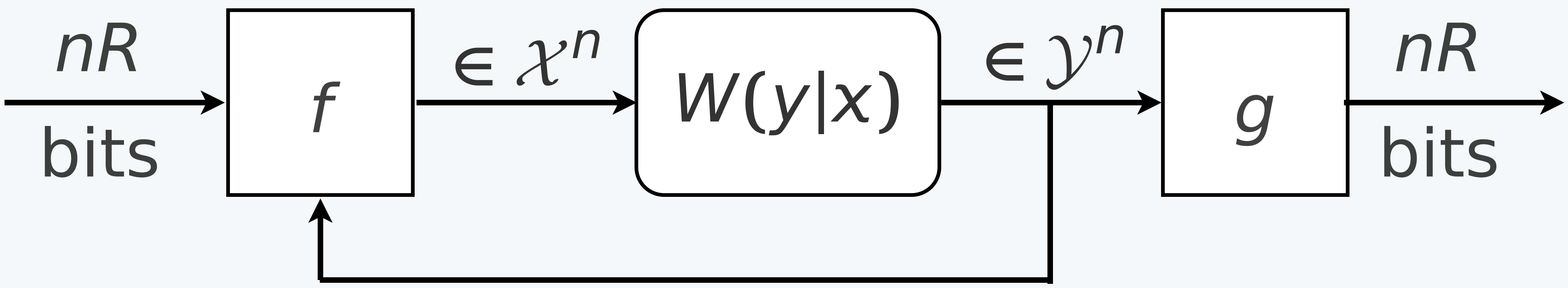
Theorem: (folk)

For any $\mathcal{X}$, $\mathcal{Y}$, R, and $\epsilon > 0$,

$$\lim_{n\to\infty} \inf_{f,g:\text{rate}\geq R-\epsilon} \sup_{W:\mathcal{X}\to\mathcal{Y}:C(W)\geq R} \mathrm{P_e}(f,g,W) = 0$$

Proof: Learn the channel via training symbols.

Channel Coding as Minimax



Theorem: (folk)

For any $\mathcal{X}$, $\mathcal{Y}$, R, and $\epsilon > 0$,

$$\lim_{n\to\infty} \inf_{f,g:\text{rate}\geq R-\epsilon} \sup_{W:\mathcal{X}\to\mathcal{Y}:C(W)\geq R} \mathrm{P_e}(f,g,W) = 0$$

Proof: Learn the channel via training symbols.

How large must n be? (as a function of $|\mathcal{X}|$ and $|\mathcal{Y}|$)

Formulation

Definition:

$\{(\mathcal{X}_n, \mathcal{Y}_n)\}_{n=1}^{\infty}$ *supports universal communication* if for all $\epsilon > 0$ and all $\{R_n\}_{n=1}^{\infty}$

$$\lim_{n\to\infty} \inf_{f,g:\text{rate}\geq e^{-\epsilon}R_n-\epsilon} \sup_{W:\mathcal{X}_n\to\mathcal{Y}_n:C(W)\geq R_n} \mathrm{P_e}(f, g, W) = 0$$

Formulation

Definition:

$\{(\mathcal{X}_n, \mathcal{Y}_n)\}_{n=1}^{\infty}$ *supports universal communication* if for all $\epsilon > 0$ and all $\{R_n\}_{n=1}^{\infty}$

$$\lim_{n\to\infty} \inf_{f,g:\text{rate}\geq e^{-\epsilon}R_n-\epsilon} \sup_{W:\mathcal{X}_n\to\mathcal{Y}_n:C(W)\geq R_n} \mathrm{P_e}(f,g,W) = 0$$

- $\mathcal{X}_n = \mathcal{X}$ and $\mathcal{Y}_n = \mathcal{Y}$ ✓

Formulation

Definition:

$\{(\mathcal{X}_n, \mathcal{Y}_n)\}_{n=1}^{\infty}$ *supports universal communication* if for all $\epsilon > 0$ and all $\{R_n\}_{n=1}^{\infty}$

$$\lim_{n\to\infty} \inf_{f,g:\text{rate}\geq e^{-\epsilon}R_n-\epsilon} \sup_{W:\mathcal{X}_n\to\mathcal{Y}_n:C(W)\geq R_n} \mathrm{P_e}(f, g, W) = 0$$

- $\mathcal{X}_n = \mathcal{X}$ and $\mathcal{Y}_n = \mathcal{Y}$ ✓
- How fast can $\mathcal{X}_n$ and $\mathcal{Y}_n$ grow with n?

Formulation

Definition:

$\{(\mathcal{X}_n, \mathcal{Y}_n)\}_{n=1}^{\infty}$ *supports universal communication* if for all $\epsilon > 0$ and all $\{R_n\}_{n=1}^{\infty}$

$$\lim_{n\to\infty} \inf_{f,g:\text{rate}\geq e^{-\epsilon}R_n-\epsilon} \sup_{W:\mathcal{X}_n\to\mathcal{Y}_n:C(W)\geq R_n} \mathrm{P_e}(f, g, W) = 0$$

- cf. universal lossless compression
  - Orlitsky et al.
  - Boucheron et al.
  - Shamir
  - Szpankowski and Weinberger
  - ...

Formulation

Definition:

$\{(\mathcal{X}_n, \mathcal{Y}_n)\}_{n=1}^{\infty}$ *supports universal communication* if for all $\epsilon > 0$ and all $\{R_n\}_{n=1}^{\infty}$

$$\lim_{n\to\infty} \inf_{f,g:\text{rate}\geq e^{-\epsilon}R_n-\epsilon} \sup_{W:\mathcal{X}_n\to\mathcal{Y}_n:C(W)\geq R_n} \mathrm{P_e}(f,g,W) = 0$$

Assume $|\mathcal{X}_n| = n^{\alpha}$
$|\mathcal{Y}_n| = n^{\beta}$

Learning the Channel

Theorem (Gao and Wagner '13): The sequence $\{(\mathcal{X}_n, \mathcal{Y}_n)\}$ supports universal communication if

$$|\mathcal{X}_n| = n^{\alpha}$$
$$|\mathcal{Y}_n| = n^{\beta}$$

for some $\alpha + \beta < 1$.

Learning the Channel

Theorem (Gao and Wagner '13): The sequence $\{(\mathcal{X}_n, \mathcal{Y}_n)\}$ supports universal communication if

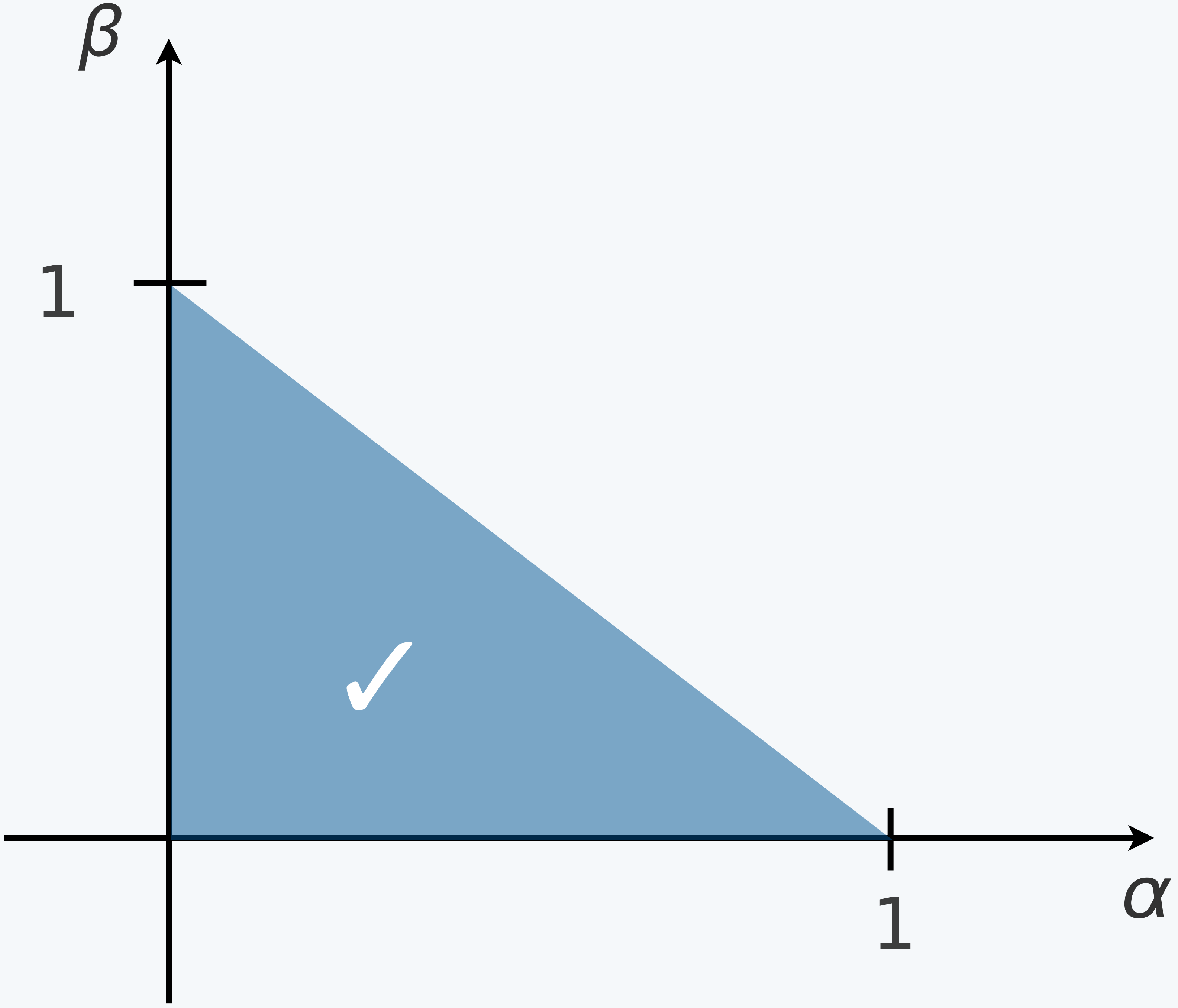
$$|\mathcal{X}_n| = n^{\alpha}$$

$$|\mathcal{Y}_n| = n^{\beta}$$

for some $\alpha + \beta < 1$.

$|\mathcal{X}_n| \cdot |\mathcal{Y}_n| = \#$ of parameters in the channel

Learning the Channel

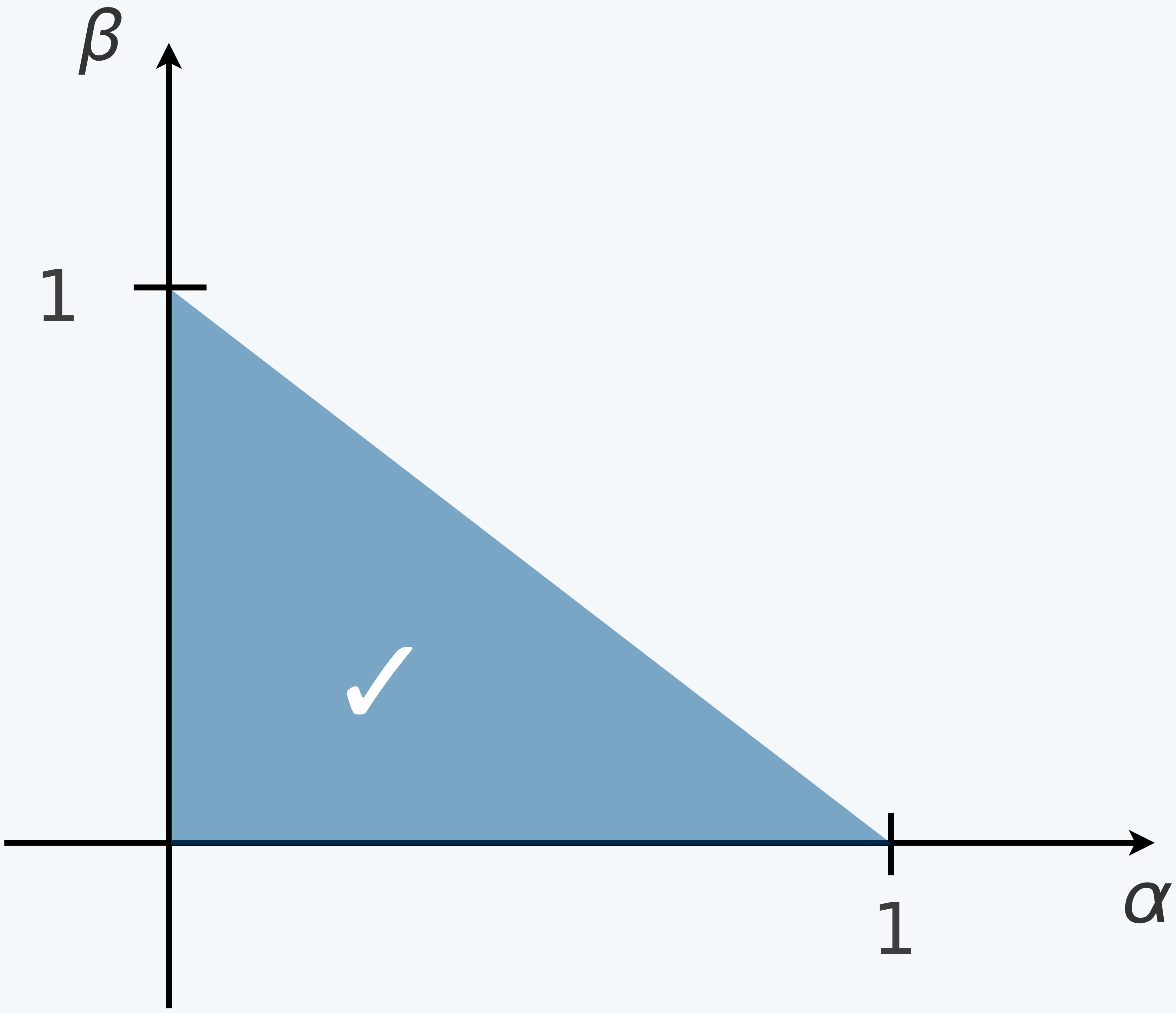


n : blocklength

$|\mathcal{X}_n| = n^\alpha$

$|\mathcal{Y}_n| = n^\beta$

Learning the Channel



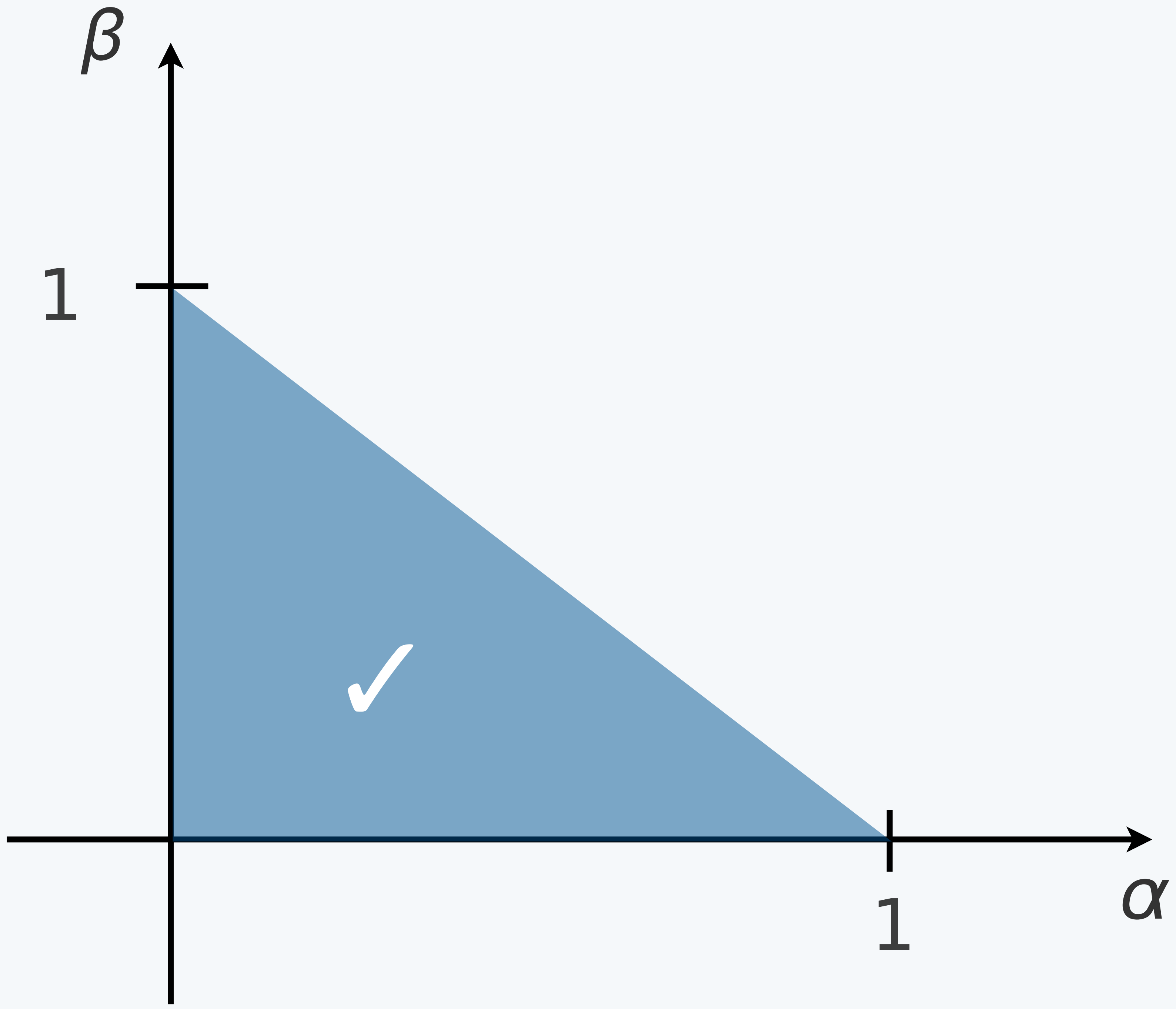
n : blocklength

$|\mathcal{X}_n| = n^{\alpha}$

$|\mathcal{Y}_n| = n^{\beta}$

Is this optimal?

Learning the Channel



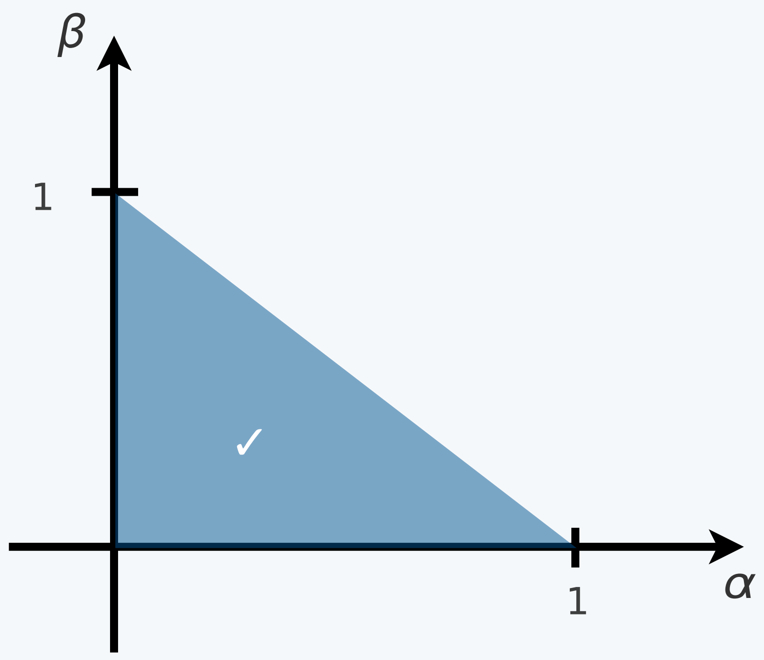
n : blocklength

$|\mathcal{X}_n| = n^{\alpha}$

$|\mathcal{Y}_n| = n^{\beta}$

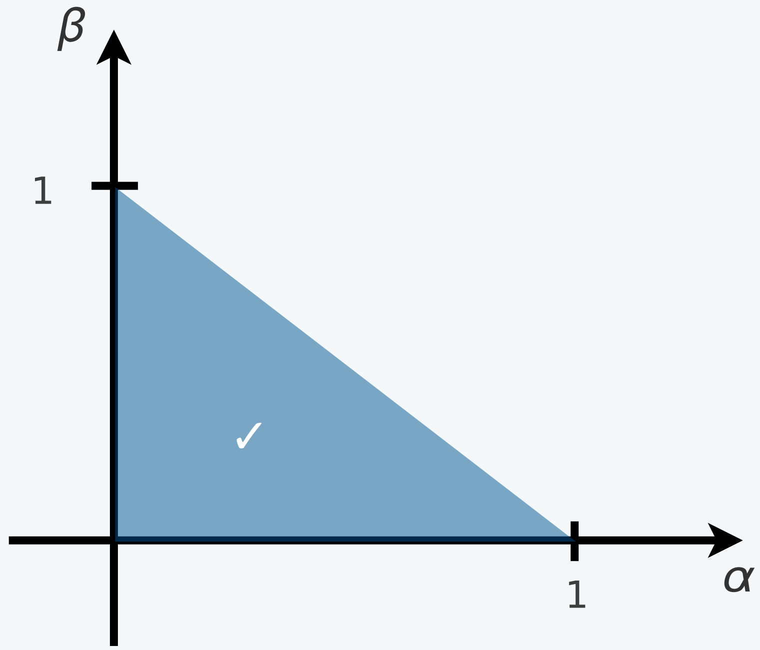
It is if one insists upon learning the channel.

Learning the Channel



Must one [be able to] learn the channel in order to communicate at capacity?

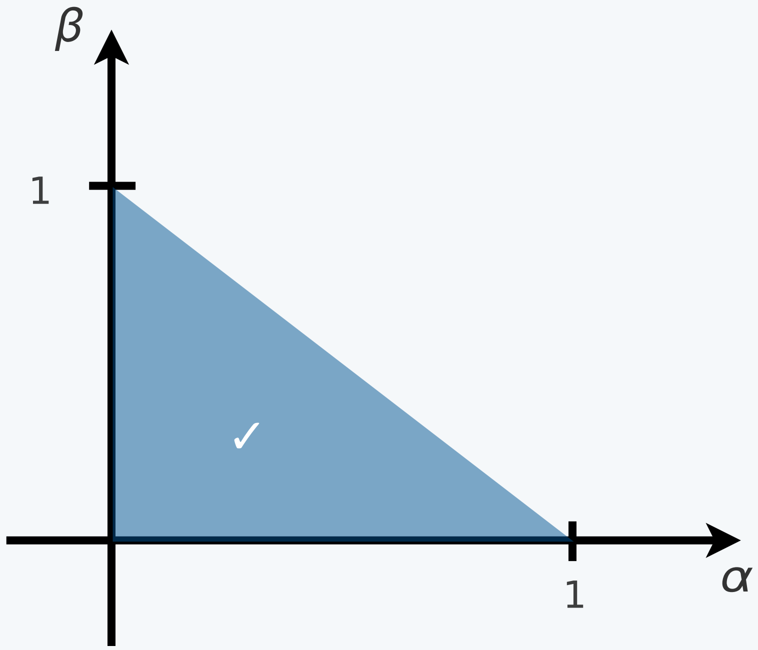
Learning the Channel



Must one [be able to] learn the channel in order to communicate at capacity?

- Training-based scheme suggests maybe not:

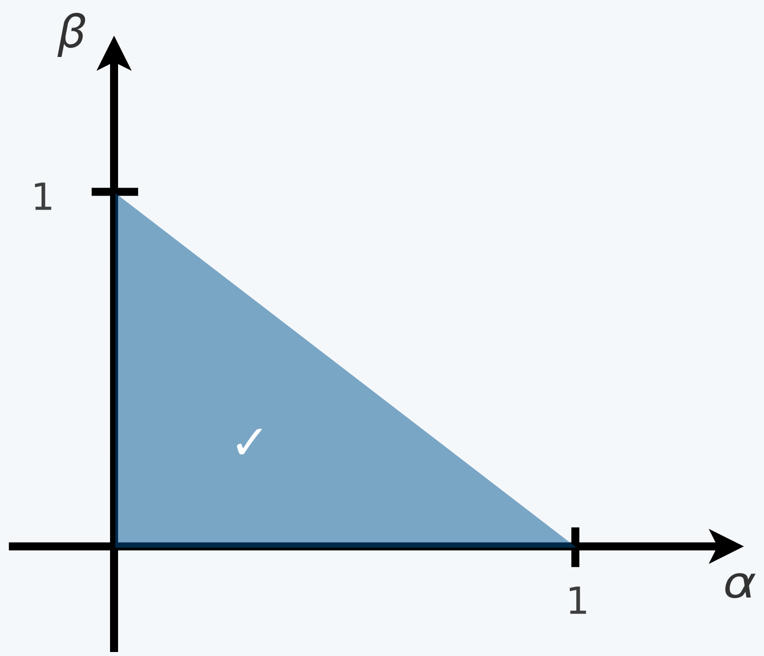
Learning the Channel



Must one [be able to] learn the channel in order to communicate at capacity?

- Training-based scheme suggests maybe not:
  - Encoder only needs arg min $I(p;W)$

Learning the Channel



Must one [be able to] learn the channel in order to communicate at capacity?

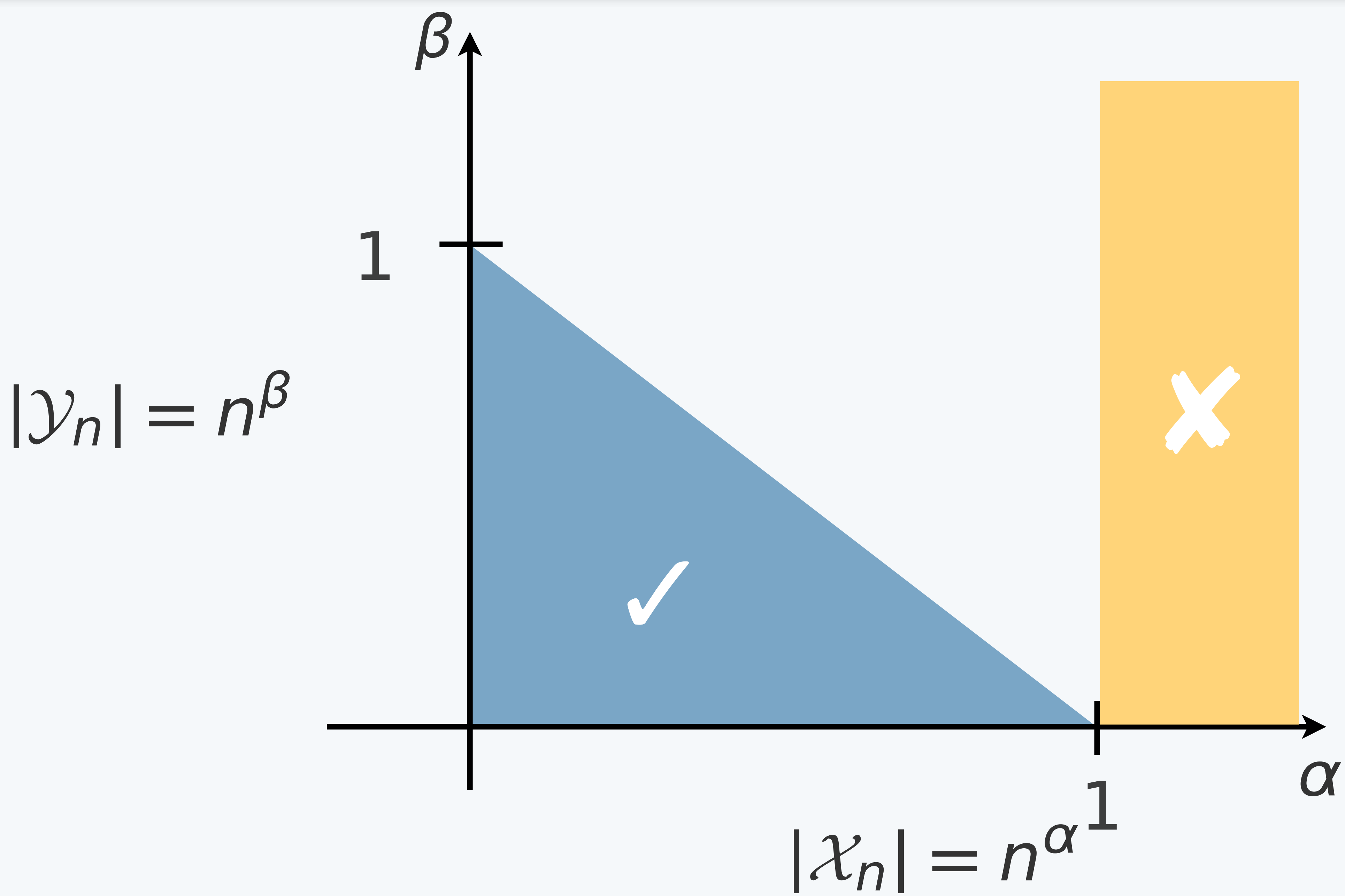
- Training-based scheme suggests maybe not:
  - Encoder only needs arg min $I(p;W)$
  - Decoder doesn't use the training

Impossibility Result #1

Theorem (Gao and Wagner '14): if $|\mathcal{X}_n| = n^{\alpha}$ and $\alpha \geq 1$ then $(\{\mathcal{X}_n\}, \{\mathcal{Y}_n\})$ does not support universal communication.

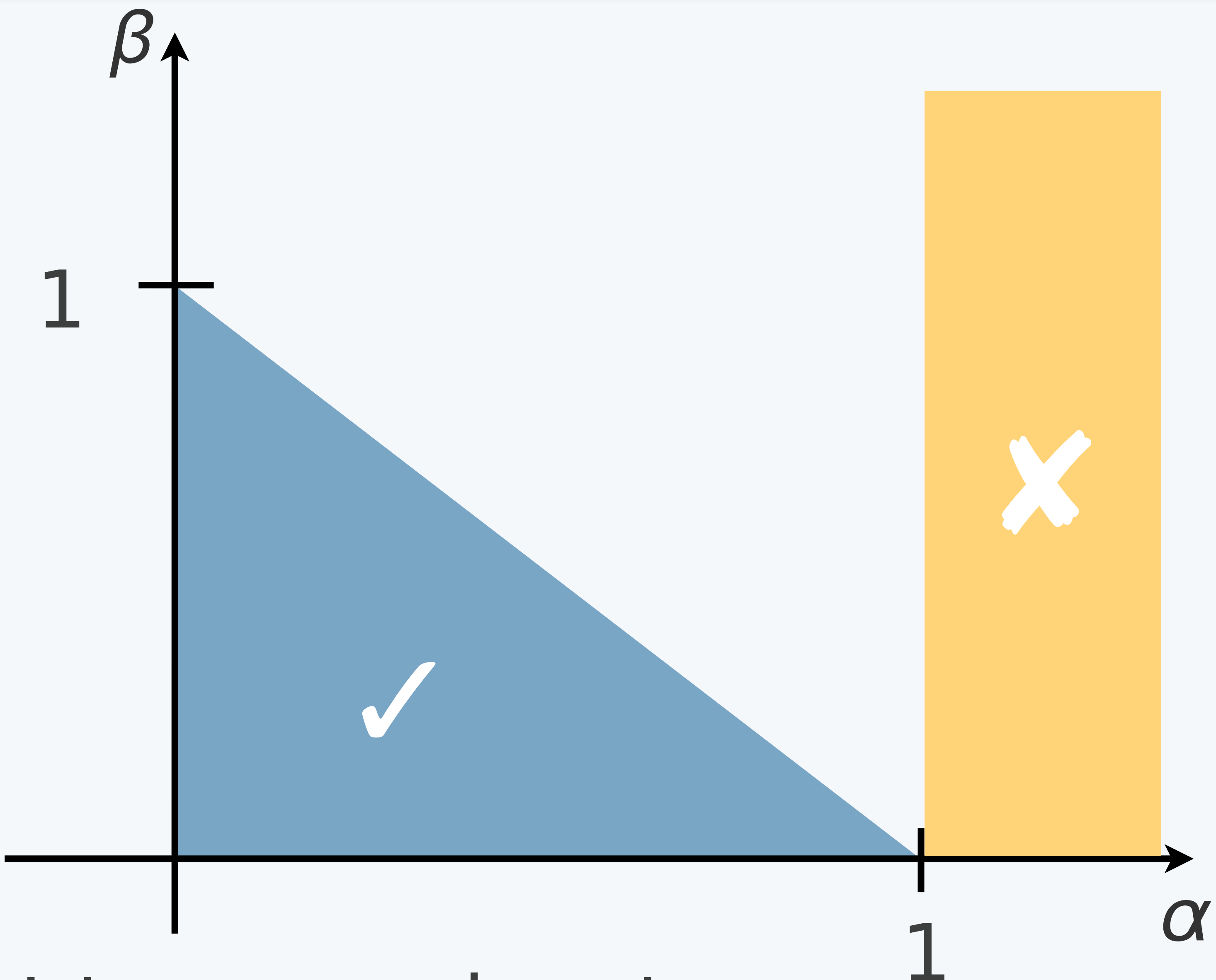
Impossibility Result #1

Theorem (Gao and Wagner '14): if $|\mathcal{X}_n| = n^{\alpha}$ and α ≥ 1 then $(\{\mathcal{X}_n\}, \{\mathcal{Y}_n\})$ does not support universal communication.



Impossibility Result #1

Theorem (Gao and Wagner '14): if $|\mathcal{X}_n| = n^{\alpha}$ and α ≥ 1 then $(\{\mathcal{X}_n\}, \{\mathcal{Y}_n\})$ does not support universal communication.



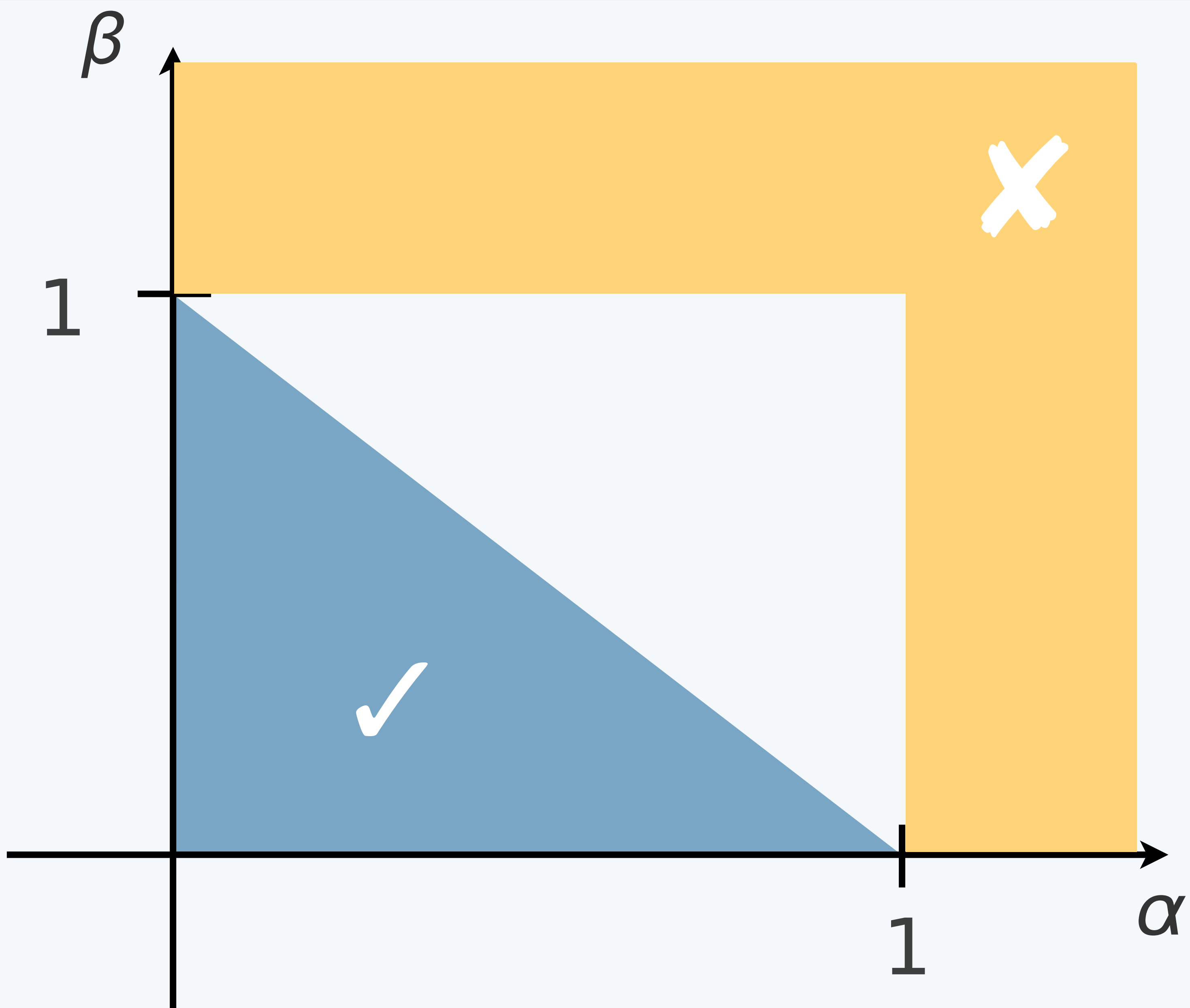
Proof: Encoder must try every input.

Impossibility Result #2

Theorem (Gao and Wagner '14): if $|\mathcal{Y}_n| = n^\beta$ and $\beta \geq 1$ then $(\{\mathcal{X}_n\}, \{\mathcal{Y}_n\})$ does not support universal communication.

Impossibility Result #2

Theorem (Gao and Wagner '14): if $|\mathcal{Y}_n| = n^\beta$ and $\beta \geq 1$ then $(\{\mathcal{X}_n\}, \{\mathcal{Y}_n\})$ does not support universal communication.



Impossibility Result #2

“Channel” minimax lower bounding techniques

Impossibility Result #2

“Channel” minimax lower bounding techniques

$$\begin{bmatrix} \frac{2}{n^\beta} & \cdots & \frac{2}{n^\beta} & 0 & \cdots & 0 \\ 0 & \cdots & 0 & \frac{2}{n^\beta} & \cdots & \frac{2}{n^\beta} \end{bmatrix}$$

Impossibility Result #2

"Channel" minimax lower bounding techniques

$$\begin{bmatrix} \frac{2}{n^\beta} & \cdots & \frac{2}{n^\beta} & 0 & \cdots & 0 \\ 0 & \cdots & 0 & \frac{2}{n^\beta} & \cdots & \frac{2}{n^\beta} \end{bmatrix}$$

- C = log 2

Impossibility Result #2

“Channel” minimax lower bounding techniques

$$\begin{bmatrix} \frac{2}{n^\beta} & \cdots & \frac{2}{n^\beta} & 0 & \cdots & 0 \\ 0 & \cdots & 0 & \frac{2}{n^\beta} & \cdots & \frac{2}{n^\beta} \end{bmatrix}$$

- C = log 2
- Achieved by a uniform input

Impossibility Result #2

$$\begin{bmatrix} \frac{2}{n^\beta} & \cdots & \frac{2}{n^\beta} & 0 & \cdots & 0 \\ 0 & \cdots & 0 & \frac{2}{n^\beta} & \cdots & \frac{2}{n^\beta} \end{bmatrix}$$

Impossibility Result #2

$$\begin{bmatrix} \frac{2}{n^\beta} & \cdots & \frac{2}{n^\beta} & 0 & \cdots & 0 \\ 0 & \cdots & 0 & \frac{2}{n^\beta} & \cdots & \frac{2}{n^\beta} \end{bmatrix}$$

- Let $\mathcal{W}_n$ be the set of all channels obtained via column permutations

Impossibility Result #2

$$\begin{bmatrix} \frac{2}{n^\beta} & \cdots & \frac{2}{n^\beta} & 0 & \cdots & 0 \\ 0 & \cdots & 0 & \frac{2}{n^\beta} & \cdots & \frac{2}{n^\beta} \end{bmatrix}$$

- Let $\mathcal{W}_n$ be the set of all channels obtained via column permutations
- For any $W \in \mathcal{W}_n$, C = log 2

Impossibility Result #2

$$\begin{bmatrix} \frac{2}{n^\beta} & \cdots & \frac{2}{n^\beta} & 0 & \cdots & 0 \\ 0 & \cdots & 0 & \frac{2}{n^\beta} & \cdots & \frac{2}{n^\beta} \end{bmatrix}$$

- Let $\mathcal{W}_n$ be the set of all channels obtained via column permutations
- For any $W \in \mathcal{W}_n$, C = log 2
- Consider the block mixture channel:

$$\bar{W}(y^n|x^n) = \frac{1}{|\mathcal{W}_n|} \sum_{W \in \mathcal{W}_n} W^n(y^n|x^n)$$

Impossibility Result #2

$$\begin{bmatrix} \frac{2}{n^\beta} & \cdots & \frac{2}{n^\beta} & 0 & \cdots & 0 \\ 0 & \cdots & 0 & \frac{2}{n^\beta} & \cdots & \frac{2}{n^\beta} \end{bmatrix}$$

- Let $\mathcal{W}_n$ be the set of all channels obtained via column permutations
- For any $W \in \mathcal{W}_n$, C = log 2
- Consider the block mixture channel:

$$\bar{W}(y^n|x^n) = \frac{1}{|\mathcal{W}_n|} \sum_{W \in \mathcal{W}_n} W^n(y^n|x^n)$$

- If the alphabets admit a universal code then the capacity of this channel must be log 2,

Impossibility Result #2

$$\begin{bmatrix} \frac{2}{n^\beta} & \cdots & \frac{2}{n^\beta} & 0 & \cdots & 0 \\ 0 & \cdots & 0 & \frac{2}{n^\beta} & \cdots & \frac{2}{n^\beta} \end{bmatrix}$$

- Let $\mathcal{W}_n$ be the set of all channels obtained via column permutations
- For any $W \in \mathcal{W}_n$, C = log 2
- Consider the block mixture channel:

$$\bar{W}(y^n|x^n) = \frac{1}{|\mathcal{W}_n|} \sum_{W \in \mathcal{W}_n} W^n(y^n|x^n)$$

- If the alphabets admit a universal code then the capacity of this channel must be log 2, *but it actually lower.*

Impossibility Result #2

Block-mixture channel:

$$\bar{W}(y^n|x^n) = \frac{1}{|\mathcal{W}_n|} \sum_{W \in \mathcal{W}_n} W^n(y^n|x^n)$$

Impossibility Result #2

Block-mixture channel: $$\bar{W}(y^n|x^n) = \frac{1}{|\mathcal{W}_n|} \sum_{W \in \mathcal{W}_n} W^n(y^n|x^n)$$

$$I(M;Y^n) = H(Y^n) - H(Y^n|M)$$

$$\leq n \log n^{\beta} - \sum_{i=1}^{n} H(Y_i|M, Y_1, \ldots, Y_{i-1})$$

Impossibility Result #2

Block-mixture channel: $$\bar{W}(y^n|x^n) = \frac{1}{|\mathcal{W}_n|} \sum_{W \in \mathcal{W}_n} W^n(y^n|x^n)$$

$$I(M; Y^n) = H(Y^n) - H(Y^n|M)$$

$$\leq n \log n^\beta - \sum_{i=1}^{n} H(Y_i|M, Y_1, \ldots, Y_{i-1})$$

Consider $H(Y_1|M)$

Impossibility Result #2

Block-mixture channel: $$\bar{W}(y^n|x^n) = \frac{1}{|\mathcal{W}_n|} \sum_{W \in \mathcal{W}_n} W^n(y^n|x^n)$$

$$I(M;Y^n) = H(Y^n) - H(Y^n|M)$$
$$\leq n \log n^{\beta} - \sum_{i=1}^{n} H(Y_i|M, Y_1, \ldots, Y_{i-1})$$

Consider $H(Y_1|M)$ dist. of Y_1 given M is uniform

Impossibility Result #2

Block-mixture channel: $\bar{W}(y^n|x^n) = \frac{1}{|\mathcal{W}_n|} \sum_{W \in \mathcal{W}_n} W^n(y^n|x^n)$

$$I(M; Y^n) = H(Y^n) - H(Y^n|M)$$

$$\leq n \log n^\beta - \sum_{i=1}^{n} H(Y_i|M, Y_1, \ldots, Y_{i-1})$$

Consider $H(Y_1|M)$ dist. of Y_1 given M is uniform

$$\rightarrow H(Y_1|M) = \log n^\beta$$

Impossibility Result #2

Block-mixture channel: $$\bar{W}(y^n|x^n) = \frac{1}{|\mathcal{W}_n|} \sum_{W \in \mathcal{W}_n} W^n(y^n|x^n)$$

$$I(M; Y^n) = H(Y^n) - H(Y^n|M)$$
$$\leq n \log n^\beta - \sum_{i=1}^{n} H(Y_i|M, Y_1, \ldots, Y_{i-1})$$

Consider $H(Y_2|M, Y_1)$

Impossibility Result #2

Block-mixture channel: $$\bar{W}(y^n|x^n) = \frac{1}{|\mathcal{W}_n|} \sum_{W \in \mathcal{W}_n} W^n(y^n|x^n)$$

$$I(M; Y^n) = H(Y^n) - H(Y^n|M)$$

$$\leq n \log n^\beta - \sum_{i=1}^{n} H(Y_i|M, Y_1, \ldots, Y_{i-1})$$

Consider $H(Y_2|M, Y_1)$

dist. of Y_2 given (M, Y_1) is (if $X_1 = X_2 = 0$)

Impossibility Result #2

Block-mixture channel: $$\bar{W}(y^n|x^n) = \frac{1}{|\mathcal{W}_n|} \sum_{W \in \mathcal{W}_n} W^n(y^n|x^n)$$

$$I(M; Y^n) = H(Y^n) - H(Y^n|M)$$
$$\leq n \log n^\beta - \sum_{i=1}^{n} H(Y_i|M, Y_1, \ldots, Y_{i-1})$$

Consider $H(Y_2|M, Y_1)$

dist. of Y_2 given (M, Y_1) is (if $X_1 = X_2 = 0$)

$$\left[q \;\; q \;\; \cdots \;\; q \;\; \frac{2}{n^\beta} \;\; q \cdots q \right]$$

Impossibility Result #2

Block-mixture channel: $\bar{W}(y^n|x^n) = \frac{1}{|\mathcal{W}_n|} \sum_{W \in \mathcal{W}_n} W^n(y^n|x^n)$

$$I(M;Y^n) = H(Y^n) - H(Y^n|M)$$

$$\leq n \log n^\beta - \sum_{i=1}^{n} H(Y_i|M, Y_1, \ldots, Y_{i-1})$$

Consider $H(Y_2|M, Y_1)$

dist. of Y_2 given (M, Y_1) is (if $X_1 = X_2 = 0$)

$$\rightarrow H(Y_2|M, Y_1) \approx \log n^\beta$$

Impossibility Result #2

Block-mixture channel: $$\bar{W}(y^n|x^n) = \frac{1}{|\mathcal{W}_n|} \sum_{W \in \mathcal{W}_n} W^n(y^n|x^n)$$

$$I(M; Y^n) = H(Y^n) - H(Y^n|M)$$

$$\leq n \log n^\beta - \sum_{i=1}^{n} H(Y_i|M, Y_1, \ldots, Y_{i-1})$$

Consider $H(Y_2|M, Y_1)$

dist. of Y_2 given (M, Y_1) is (if $X_1 = X_2 = 0$)

$$\rightarrow H(Y_2|M, Y_1) \approx \log n^\beta$$

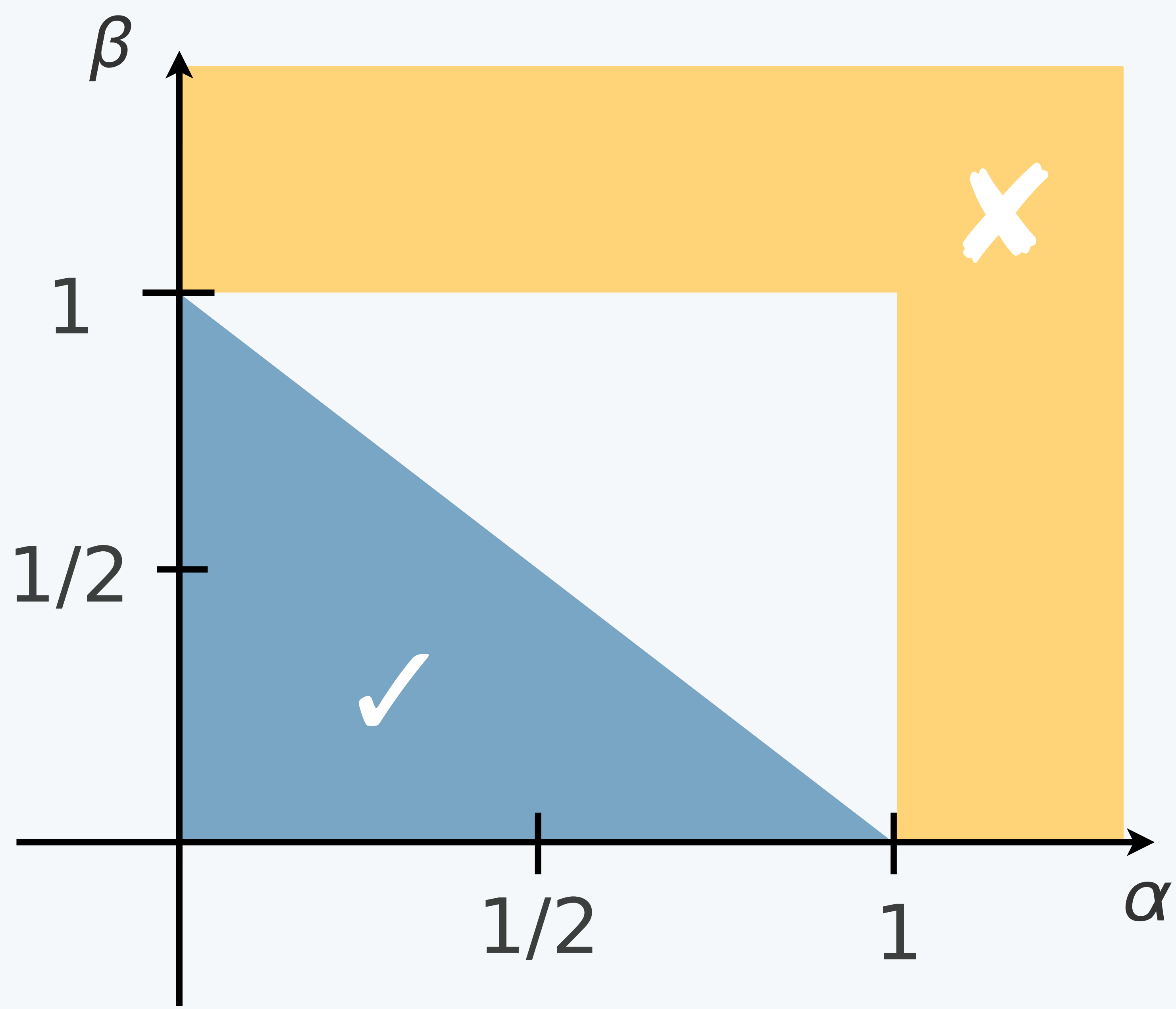
$$H(Y_i|M, Y_1, \ldots, Y_{i-1}) \approx \log n^\beta \quad \text{if } i \ll n^\beta$$

Impossibility Result #2

$$\begin{bmatrix} \frac{2}{n^\beta} & \cdots & \frac{2}{n^\beta} & 0 & \cdots & 0 \\ 0 & \cdots & 0 & \frac{2}{n^\beta} & \cdots & \frac{2}{n^\beta} \end{bmatrix}$$

- No code can achieve capacity for all channels in this class ...
- ... even though the optimal input distribution is known

Impossibility Result #2

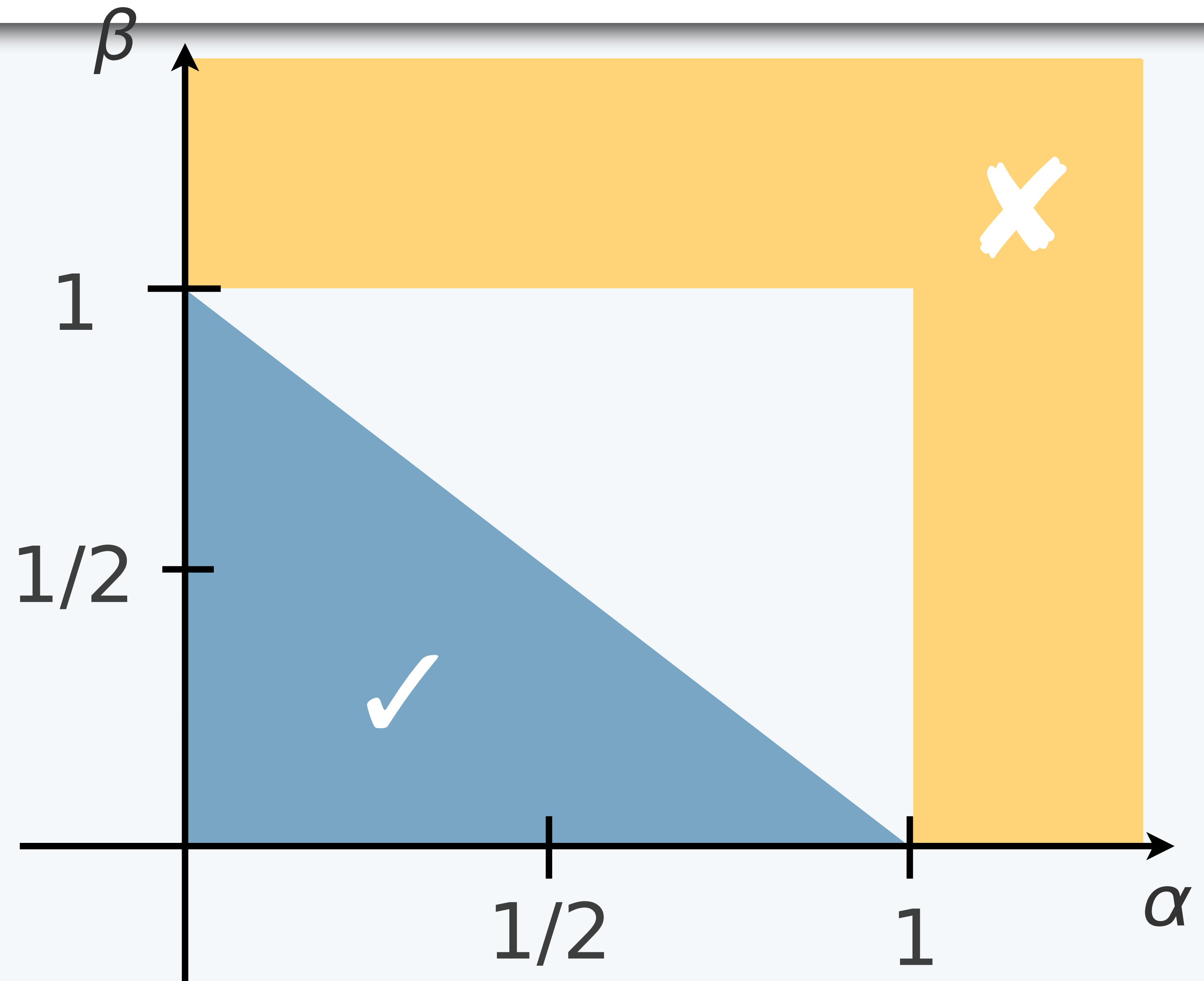


Impossibility Result #3

Theorem (Gao and Wagner '15): if $|\mathcal{X}_n| = n^\alpha$ and $|\mathcal{Y}_n| = n^\beta$ with $\alpha + \beta > 1$ and $\beta > 1/2$ then $\{(\mathcal{X}_n, \mathcal{Y}_n)\}$ does not support universal communication.

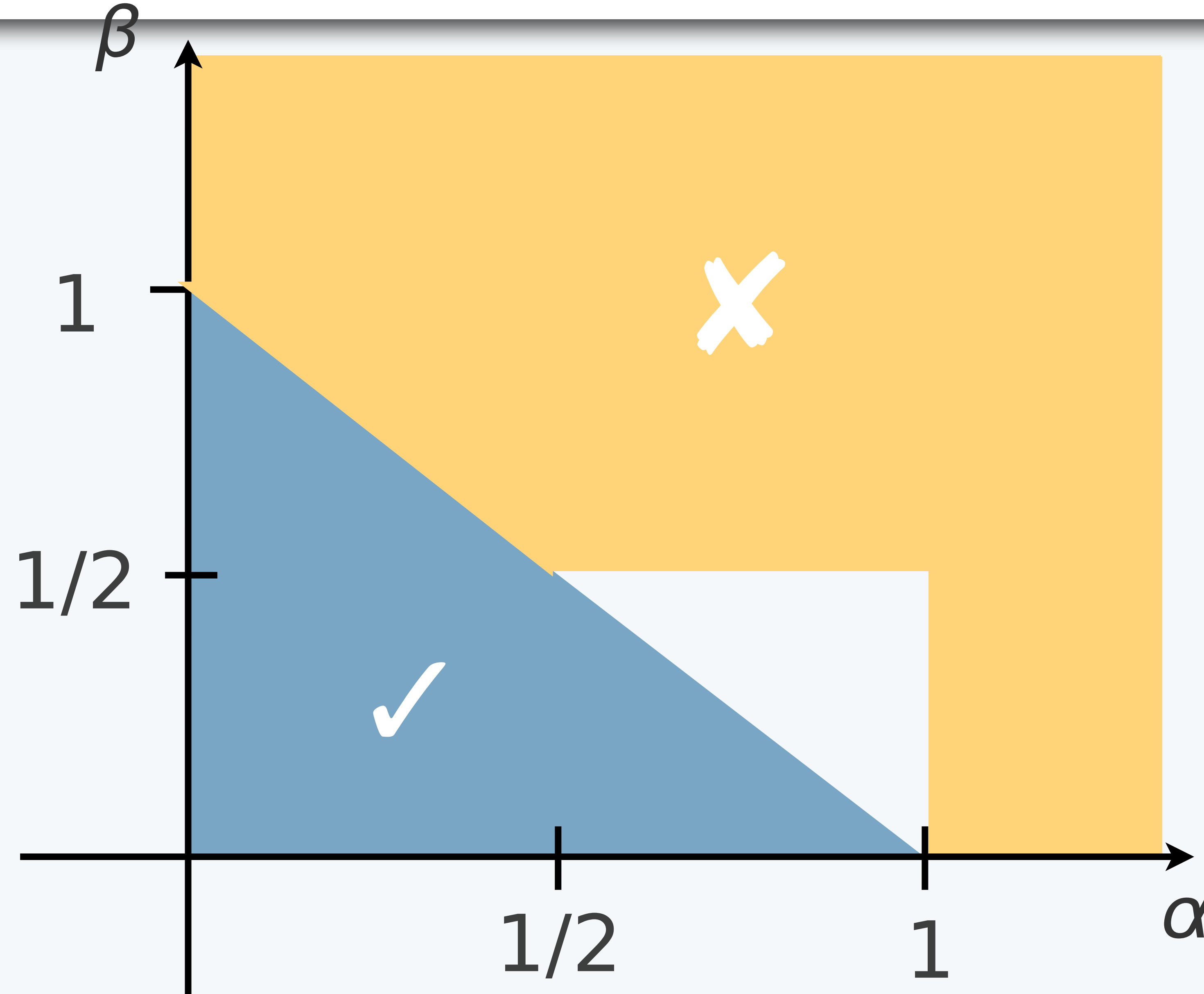
Impossibility Result #3

Theorem (Gao and Wagner '15): if $|\mathcal{X}_n| = n^{\alpha}$ and $|\mathcal{Y}_n| = n^{\beta}$ with $\alpha + \beta > 1$ and $\beta > 1/2$ then $\{(\mathcal{X}_n, \mathcal{Y}_n)\}$ does not support universal communication.



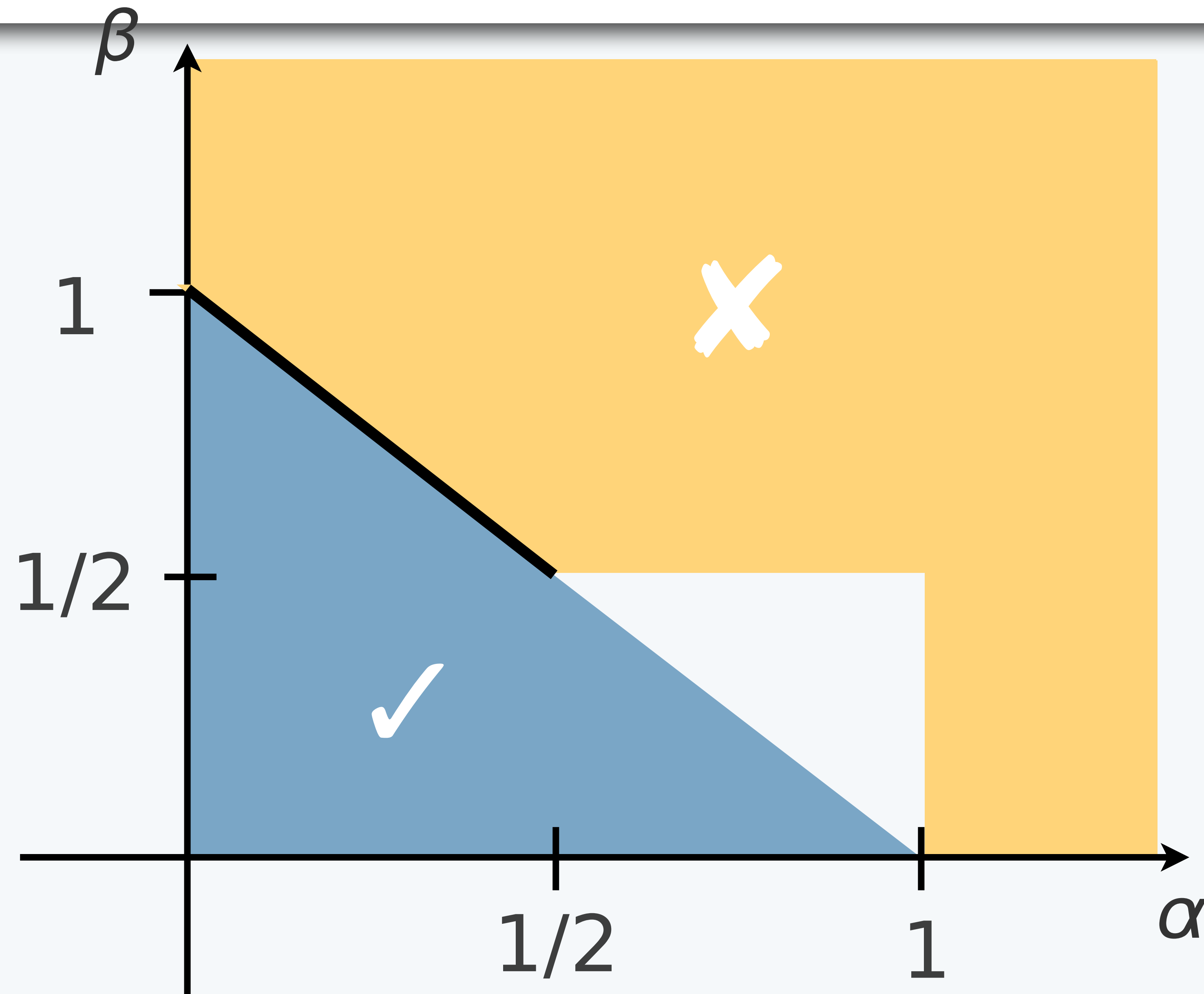
Impossibility Result #3

Theorem (Gao and Wagner ’15): if $|\mathcal{X}_n| = n^{\alpha}$ and $|\mathcal{Y}_n| = n^{\beta}$ with $\alpha + \beta > 1$ and $\beta > 1/2$ then $\{(\mathcal{X}_n, \mathcal{Y}_n)\}$ does not support universal communication.



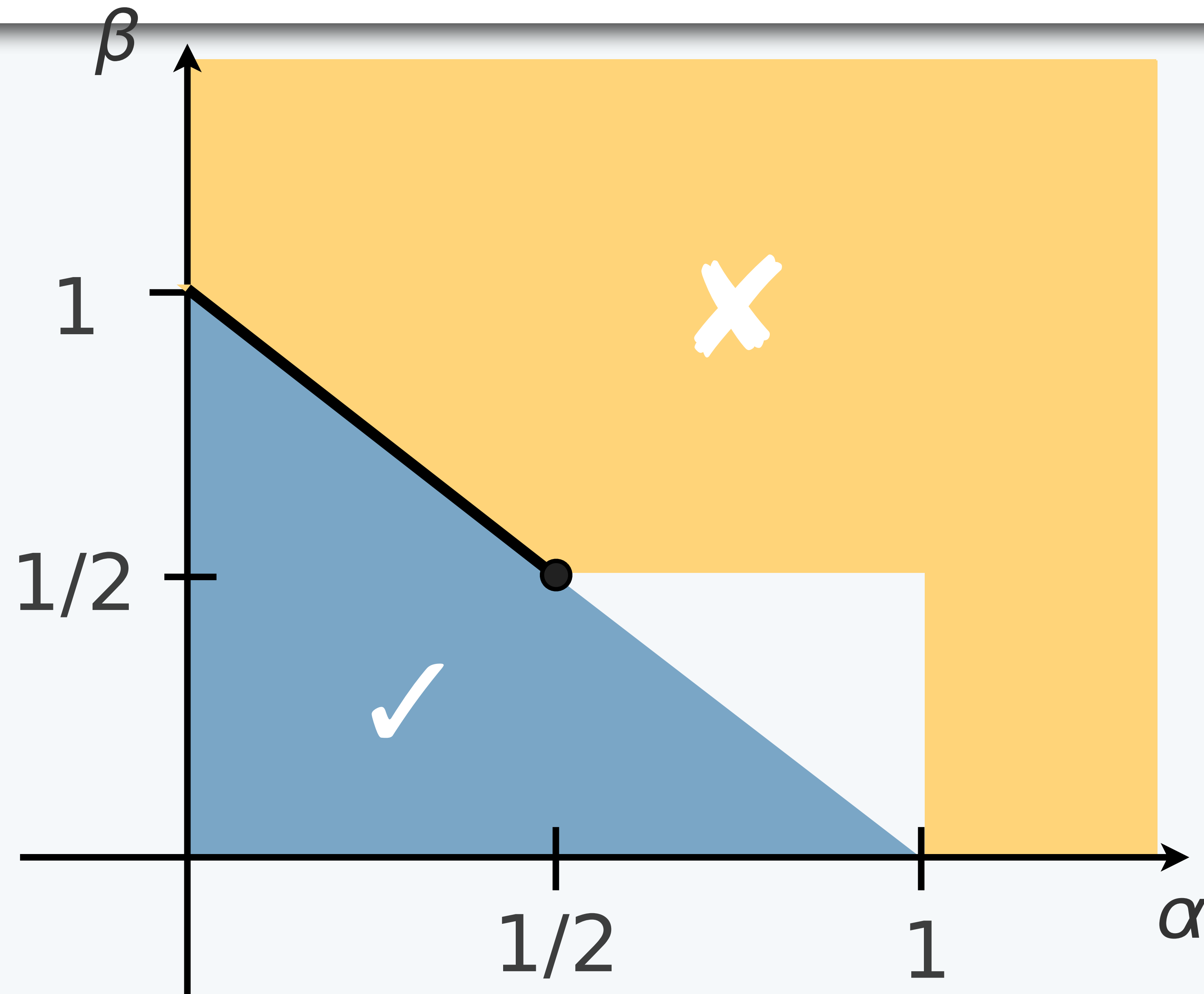
Impossibility Result #3

Theorem (Gao and Wagner '15): if $|\mathcal{X}_n| = n^{\alpha}$ and $|\mathcal{Y}_n| = n^{\beta}$ with $\alpha + \beta > 1$ and $\beta > 1/2$ then $\{(\mathcal{X}_n, \mathcal{Y}_n)\}$ does not support universal communication.



Impossibility Result #3

Theorem (Gao and Wagner '15): if $|\mathcal{X}_n| = n^{\alpha}$ and $|\mathcal{Y}_n| = n^{\beta}$ with $\alpha + \beta > 1$ and $\beta > 1/2$ then $\{(\mathcal{X}_n, \mathcal{Y}_n)\}$ does not support universal communication.



Impossibility Result #3

- Suppose $|\mathcal{X}_n| = |\mathcal{Y}_n| = n^{1/2}$
- Consider a channel from the following class:

$$\begin{bmatrix} 1/2 & & & & & \\ & 1/2 & & & & \\ & & 1/2 & & & \\ & & & \ddots & & \\ & & & & 1/2 & \\ & & & & & 1/2 \end{bmatrix}$$

Impossibility Result #3

- Suppose $|\mathcal{X}_n| = |\mathcal{Y}_n| = n^{1/2}$
- Consider a channel from the following class:

$$\begin{bmatrix} 1/2 & & & & & \\ & 1/2 & & & & \\ & & 1/2 & & & \\ & & & \ddots & & \\ & & & & 1/2 & \\ & & & & & 1/2 \end{bmatrix}$$

randomly choose n^{γ} per row

$$\gamma < \frac{1}{2}$$

Impossibility Result #3

- Suppose $|\mathcal{X}_n| = |\mathcal{Y}_n| = n^{1/2}$
- Consider a channel from the following class:

$$\begin{bmatrix} 1/2 & & & & & \\ & 1/2 & & & & \\ & & 1/2 & & & \\ & & & \ddots & & \\ & & & & 1/2 & \\ & & & & & 1/2 \end{bmatrix}$$

randomly choose n^γ $\blacksquare$ per row

$$\gamma < \frac{1}{2}$$

$$\blacksquare = \frac{1}{2n^\gamma}$$

Impossibility Result #3

- Suppose $|\mathcal{X}_n| = |\mathcal{Y}_n| = n^{1/2}$
- Consider a channel from the following class:

$$\begin{bmatrix} 1/2 & \blacksquare & & \blacksquare & & \blacksquare \\ & 1/2 & & & & \\ & & 1/2 & & & \\ & & & \ddots & & \\ & & & & 1/2 & \\ & & & & & 1/2 \end{bmatrix}$$

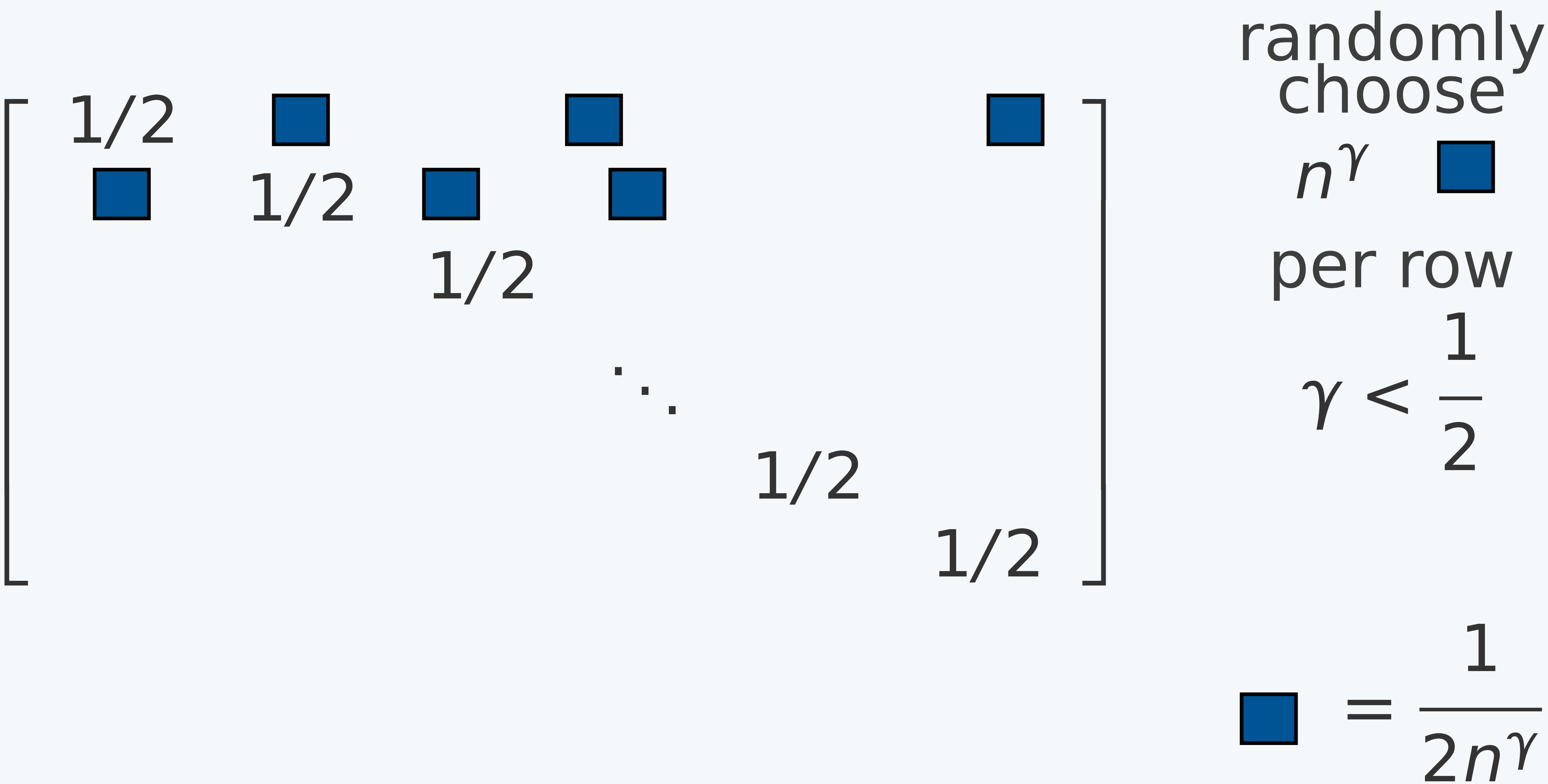
randomly choose n^{γ} $\blacksquare$ per row

$\gamma < \frac{1}{2}$

$\blacksquare = \frac{1}{2n^{\gamma}}$

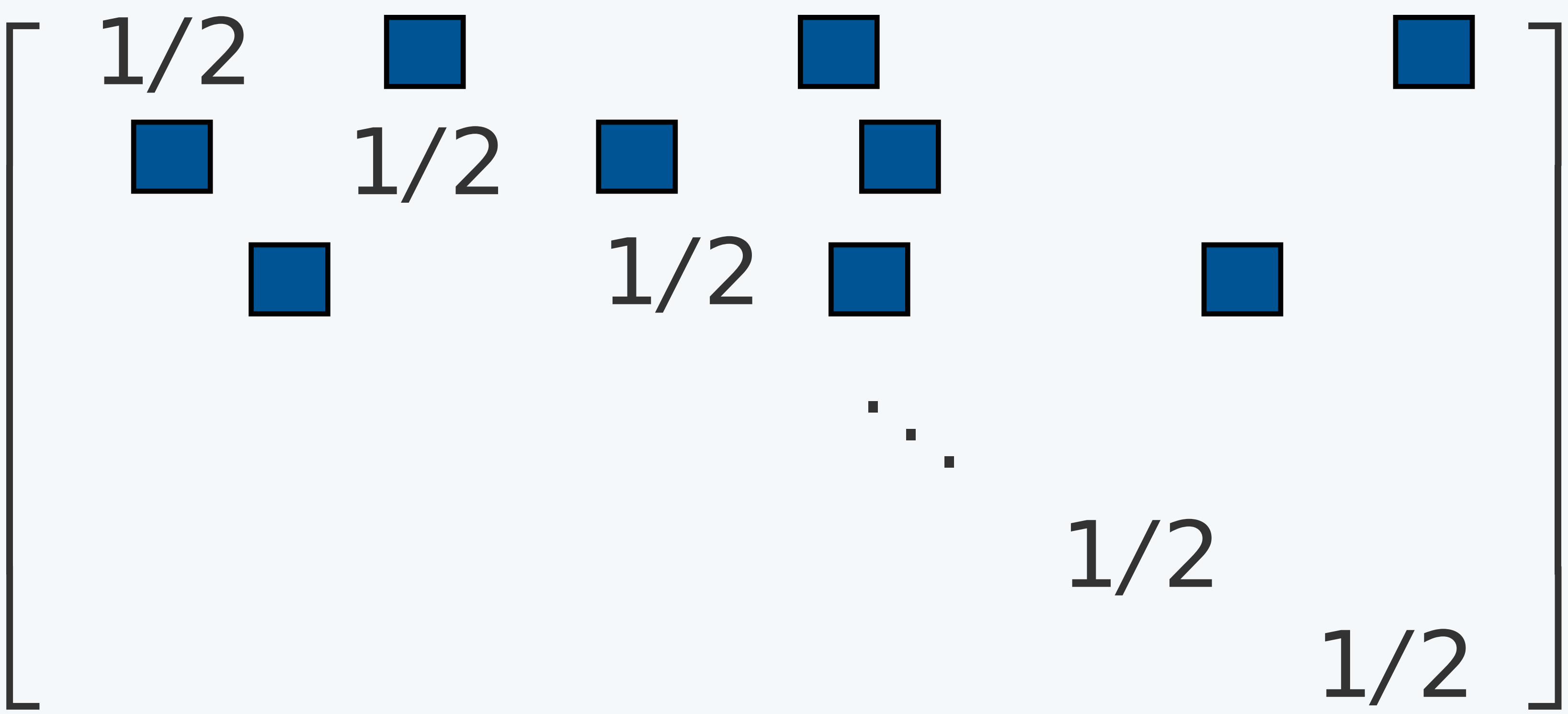
Impossibility Result #3

- Suppose $|\mathcal{X}_n| = |\mathcal{Y}_n| = n^{1/2}$
- Consider a channel from the following class:



Impossibility Result #3

- Suppose $|\mathcal{X}_n| = |\mathcal{Y}_n| = n^{1/2}$
- Consider a channel from the following class:



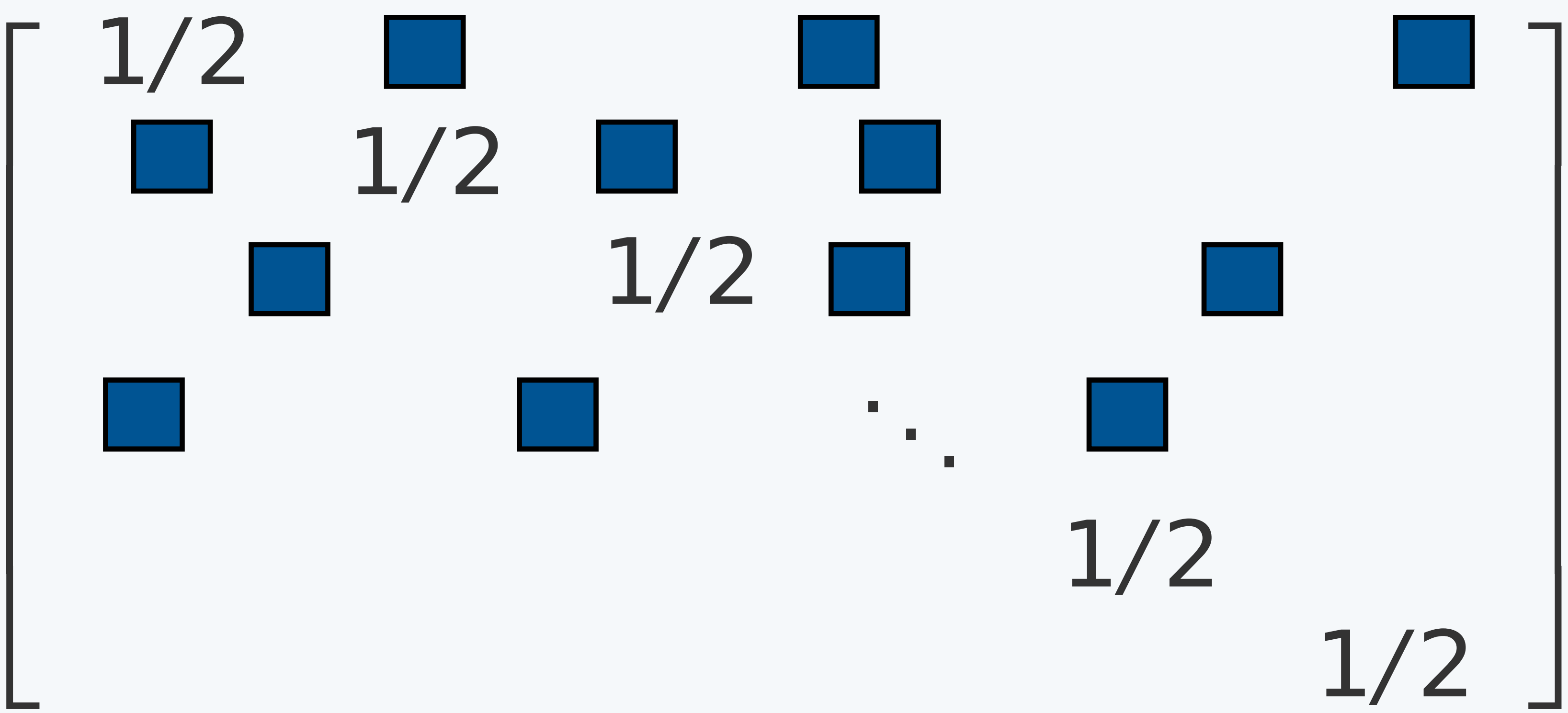
randomly choose n^{γ} ■ per row

$\gamma < \frac{1}{2}$

■ $= \frac{1}{2n^{\gamma}}$

Impossibility Result #3

- Suppose $|\mathcal{X}_n| = |\mathcal{Y}_n| = n^{1/2}$
- Consider a channel from the following class:



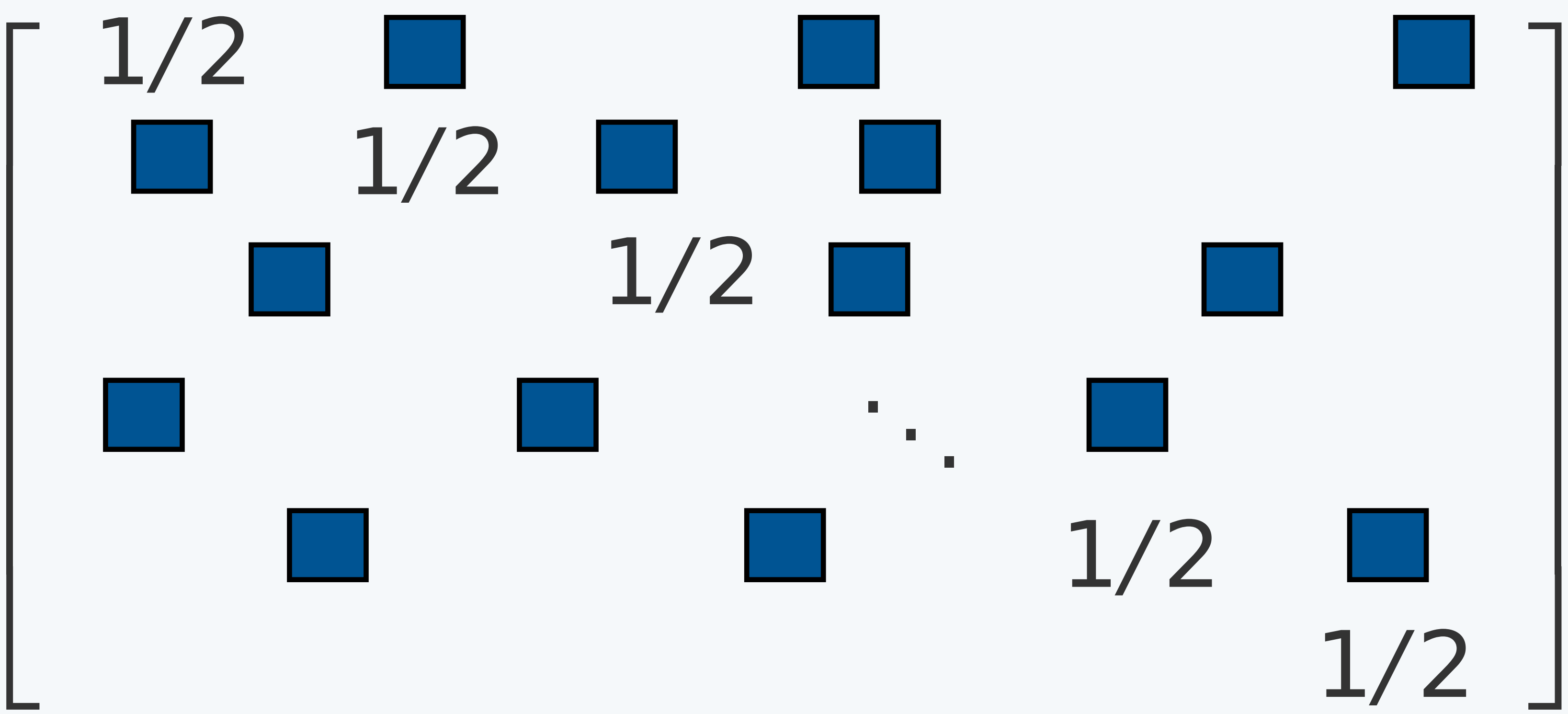
randomly choose n^γ ■ per row

$\gamma < \frac{1}{2}$

■ $= \frac{1}{2n^\gamma}$

Impossibility Result #3

- Suppose $|\mathcal{X}_n| = |\mathcal{Y}_n| = n^{1/2}$
- Consider a channel from the following class:



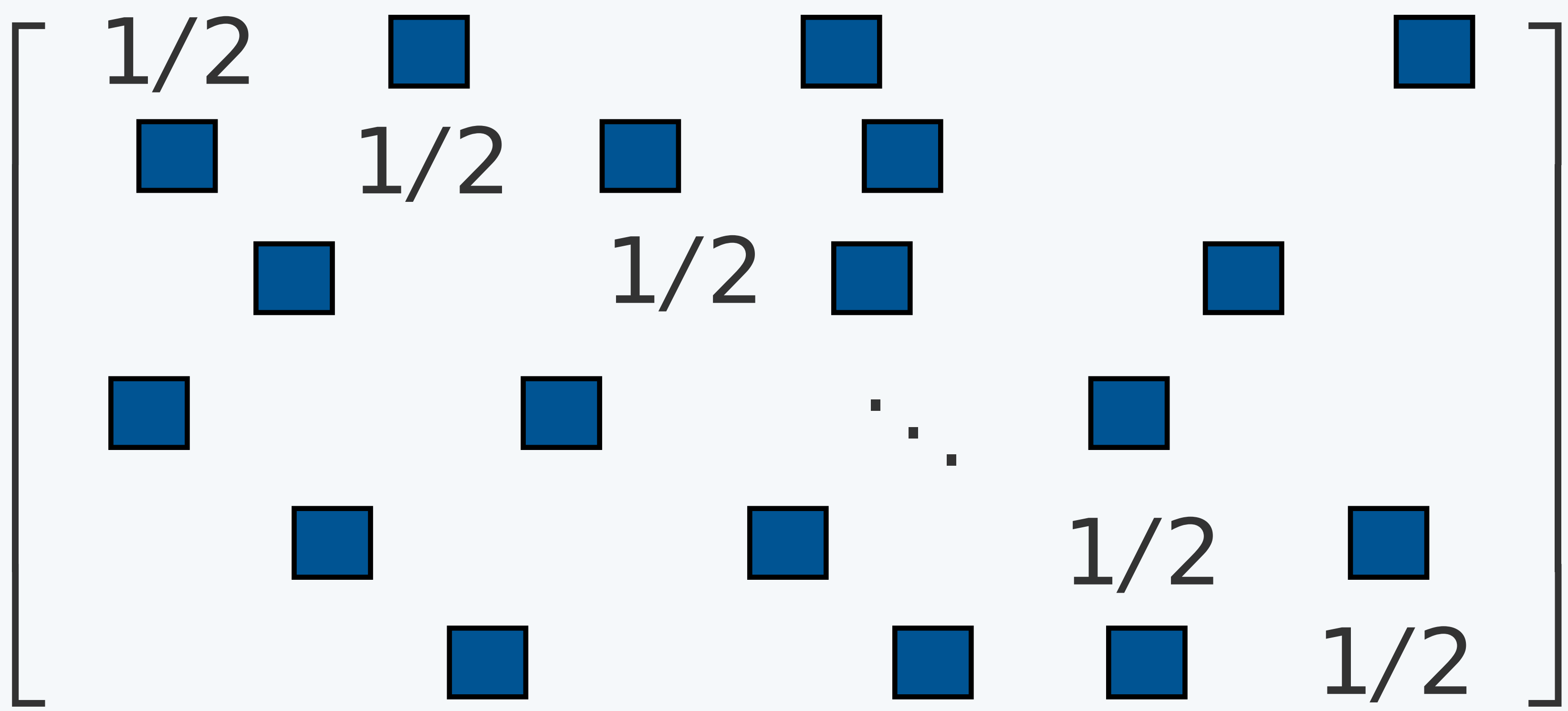
randomly choose n^γ ■ per row

$\gamma < \frac{1}{2}$

■ $= \frac{1}{2n^\gamma}$

Impossibility Result #3

- Suppose $|\mathcal{X}_n| = |\mathcal{Y}_n| = n^{1/2}$
- Consider a channel from the following class:

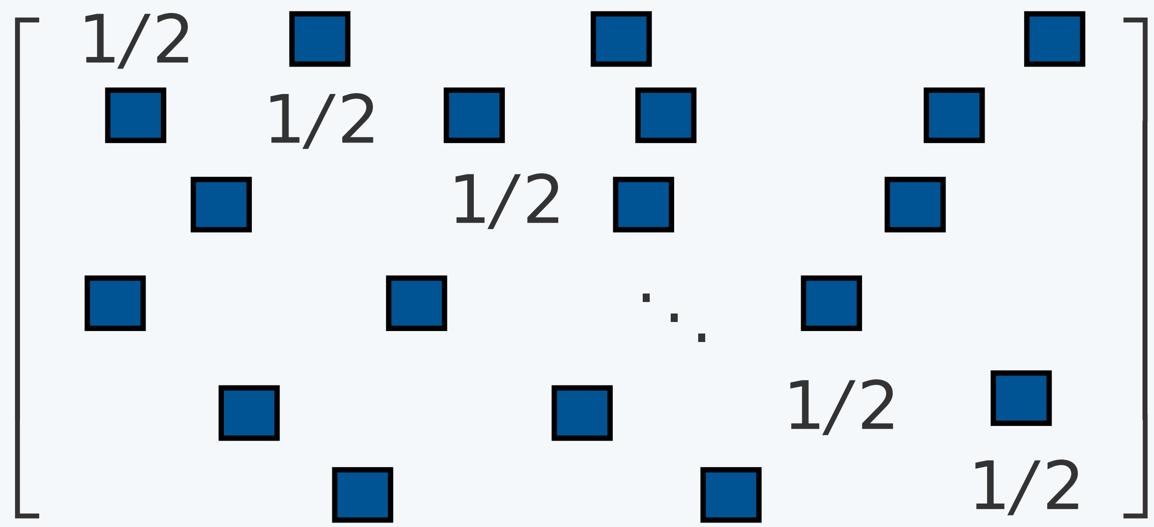


randomly choose n^γ ■ per row

$\gamma < \frac{1}{2}$

■ $= \frac{1}{2n^\gamma}$

Impossibility Result #3



- Capacity of typical channel realization is

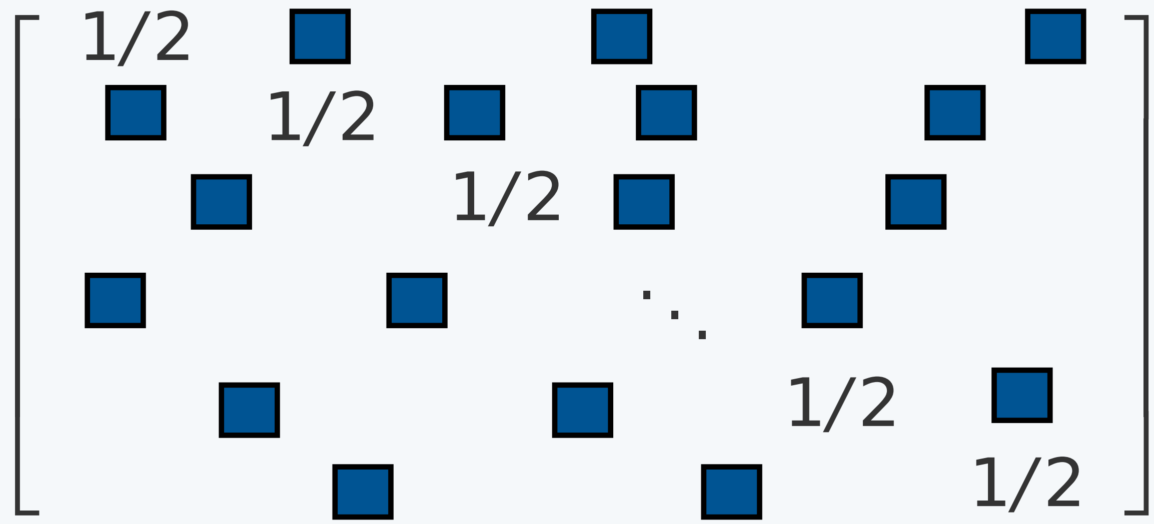
$$\max_{p(x)} I(X;Y) = \max_{p(x)} H(Y) - H(Y|X)$$

$$\leq \log n^{\beta} - \left[\frac{1}{2}\log 2 + \frac{1}{2}\log 2n^{\gamma}\right]$$

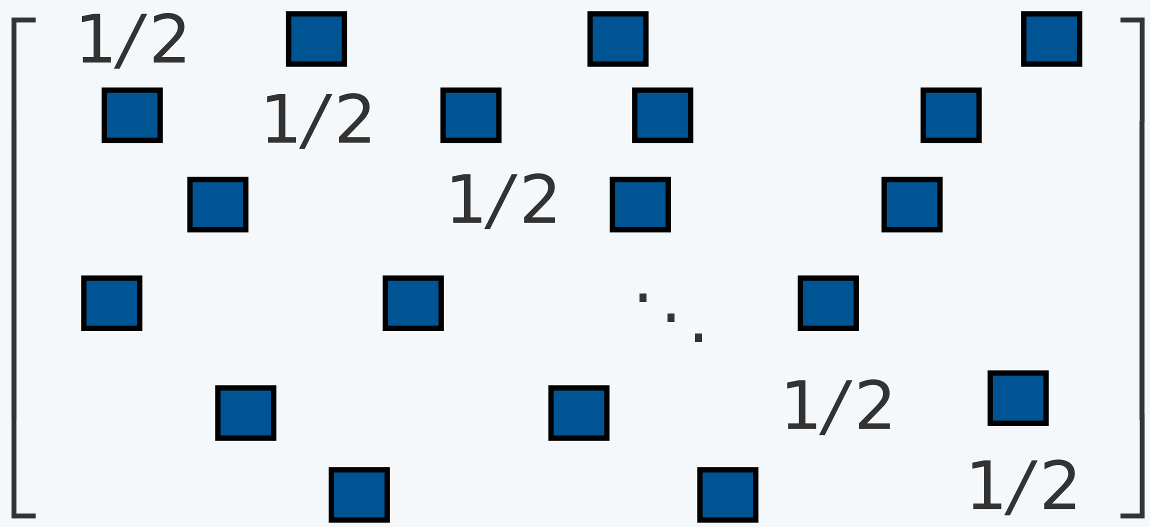
$$= (\beta - \gamma/2)\log n - \log 2$$

and is nearly achieved by a uniform input

Impossibility Result #3

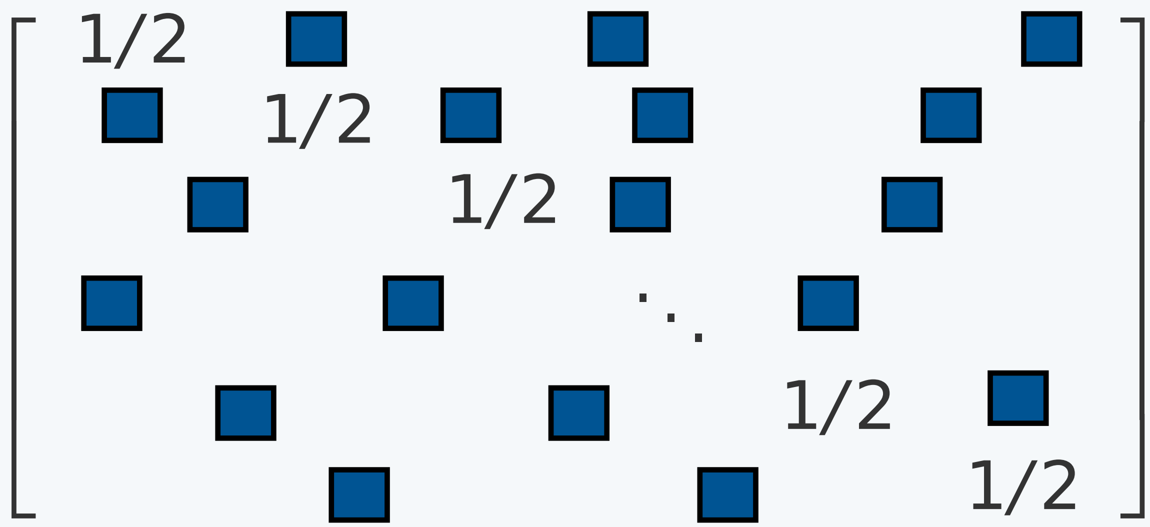


Impossibility Result #3



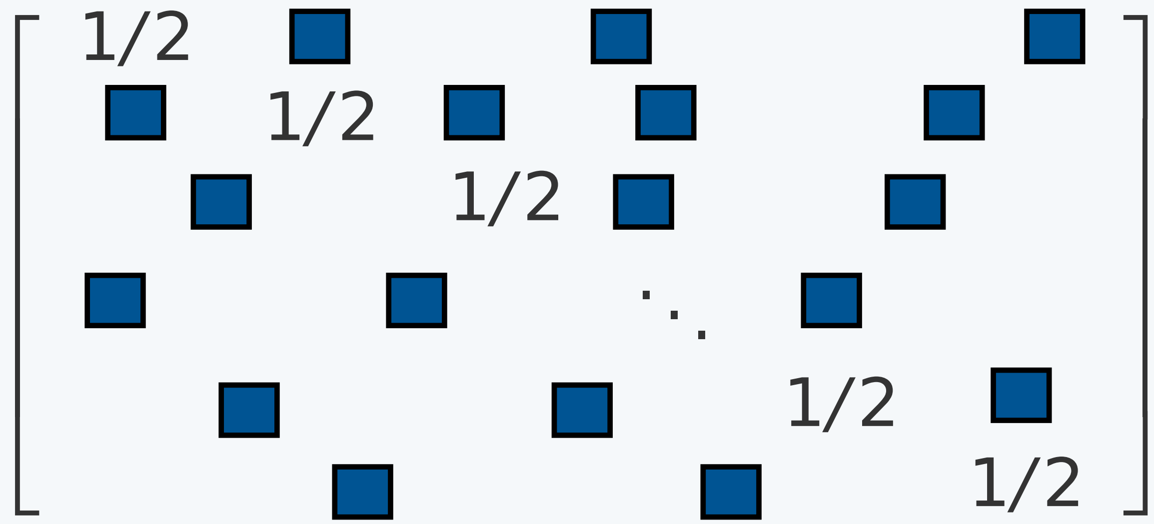
- Similarly to #2, capacity of mixture channel is $\ll (\beta - \gamma/2)\log n - \log 2$

Impossibility Result #3



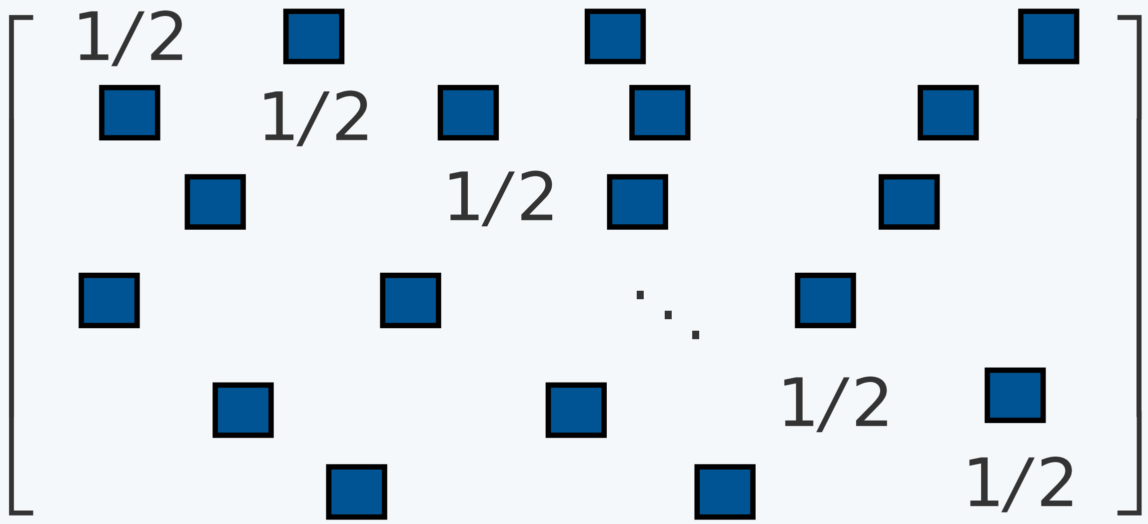
- Similarly to #2, capacity of mixture channel is $\ll (\beta - \gamma/2)\log n - \log 2$
  - Decoder must learn locations of the ■ to hit this rate

Impossibility Result #3



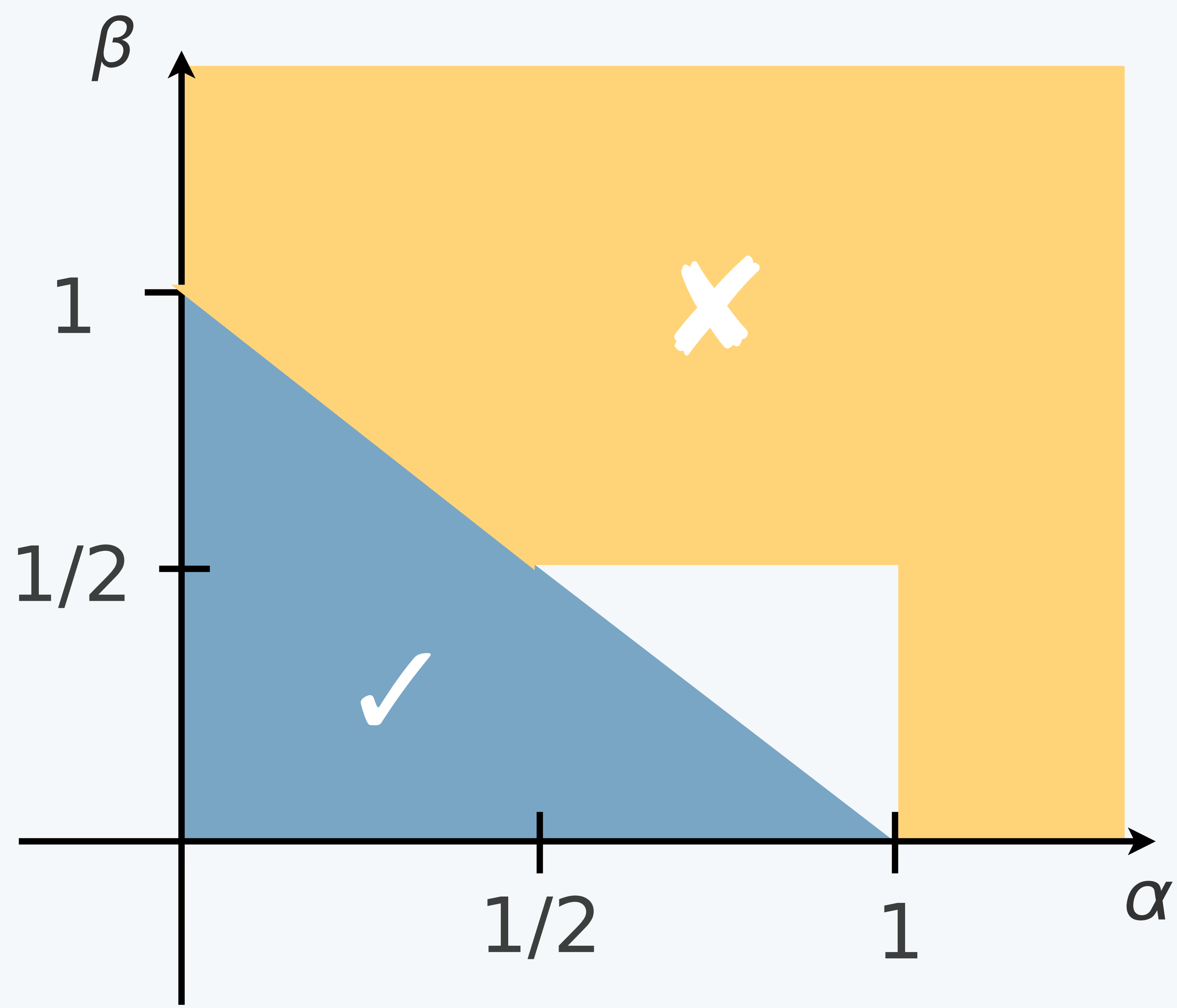
- Similarly to #2, capacity of mixture channel is $\ll (\beta - \gamma/2) \log n - \log 2$
  - Decoder must learn locations of the ■ to hit this rate
  - Can learn locations for some rows …

Impossibility Result #3



- Similarly to #2, capacity of mixture channel is $\ll (\beta - \gamma/2)\log n - \log 2$
  - Decoder must learn locations of the  to hit this rate
  - Can learn locations for some rows ...
  - ... but not for all of them.

Summary of Results



Maximin Formulation

Definition:

$\{(\mathcal{X}_n, \mathcal{Y}_n)\}_{n=1}^{\infty}$ *support communication at capacity* if for all $\epsilon > 0$ and all $\{R_n\}_{n=1}^{\infty}$

$$\lim_{n\to\infty} \sup_{W:\mathcal{X}_n\to\mathcal{Y}_n:C(W)\geq R_n} \inf_{f,g:\text{rate}\geq e^{-\epsilon}R_n-\epsilon} \mathrm{P_e}(f, g, W) = 0.$$

Maximin Formulation

Definition:

$\{(\mathcal{X}_n, \mathcal{Y}_n)\}_{n=1}^{\infty}$ *support communication at capacity* if for all $\epsilon > 0$ and all $\{R_n\}_{n=1}^{\infty}$

$$\lim_{n\to\infty} \sup_{W:\mathcal{X}_n\to\mathcal{Y}_n:C(W)\geq R_n} \inf_{f,g:\text{rate}\geq e^{-\epsilon}R_n-\epsilon} \mathrm{P_e}(f,g,W) = 0.$$

Theorem (Gao and Wagner '13): $\{(\mathcal{X}_n, \mathcal{Y}_n)\}_{n=1}^{\infty}$ support communication at capacity iff

$$\lim_{n\to\infty} \frac{\log^2 \min(|\mathcal{X}_n|, |\mathcal{Y}_n|)}{n} = 0.$$

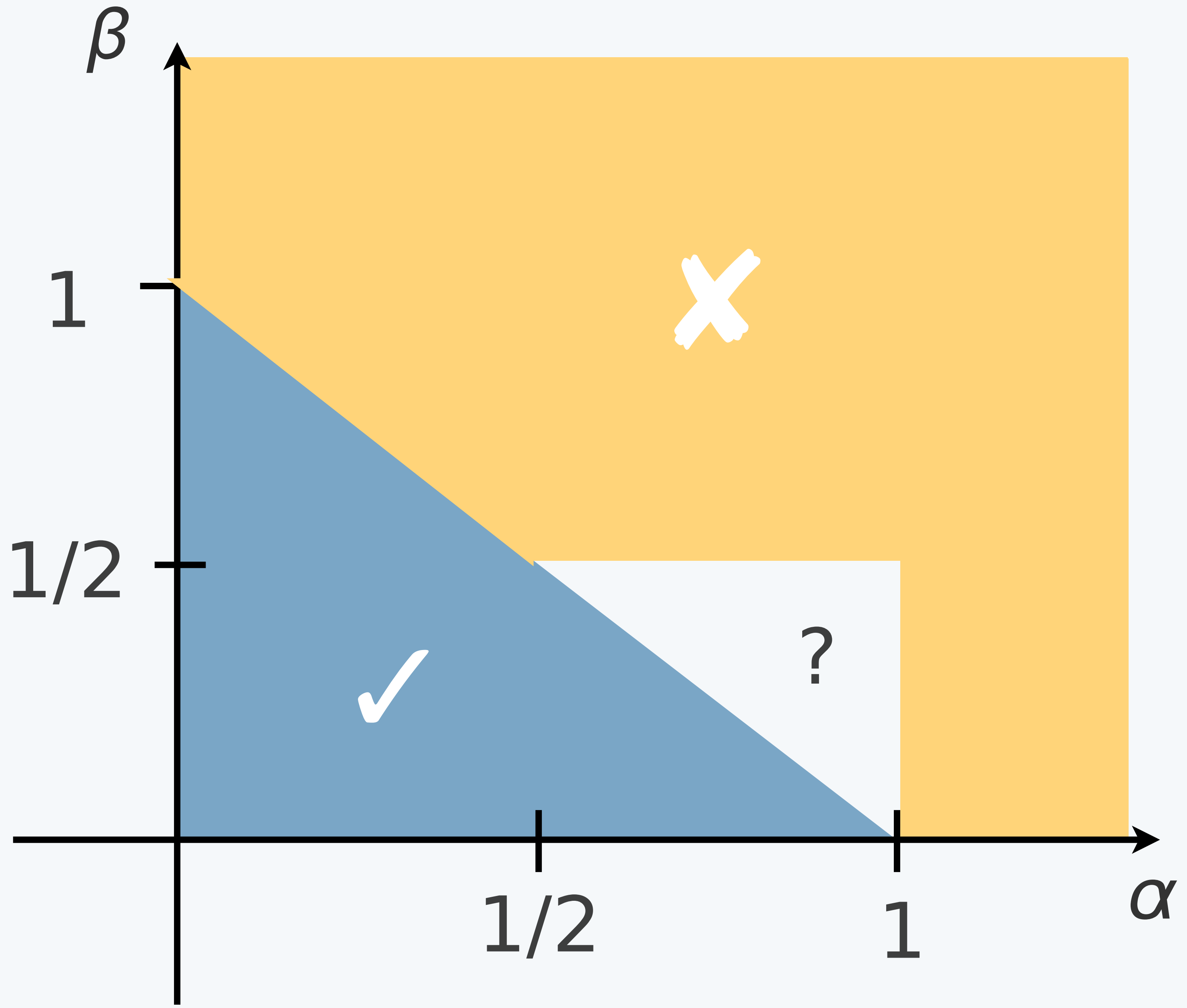
Must One Learn the Channel?

Must One Learn the Channel?

Appears so but...

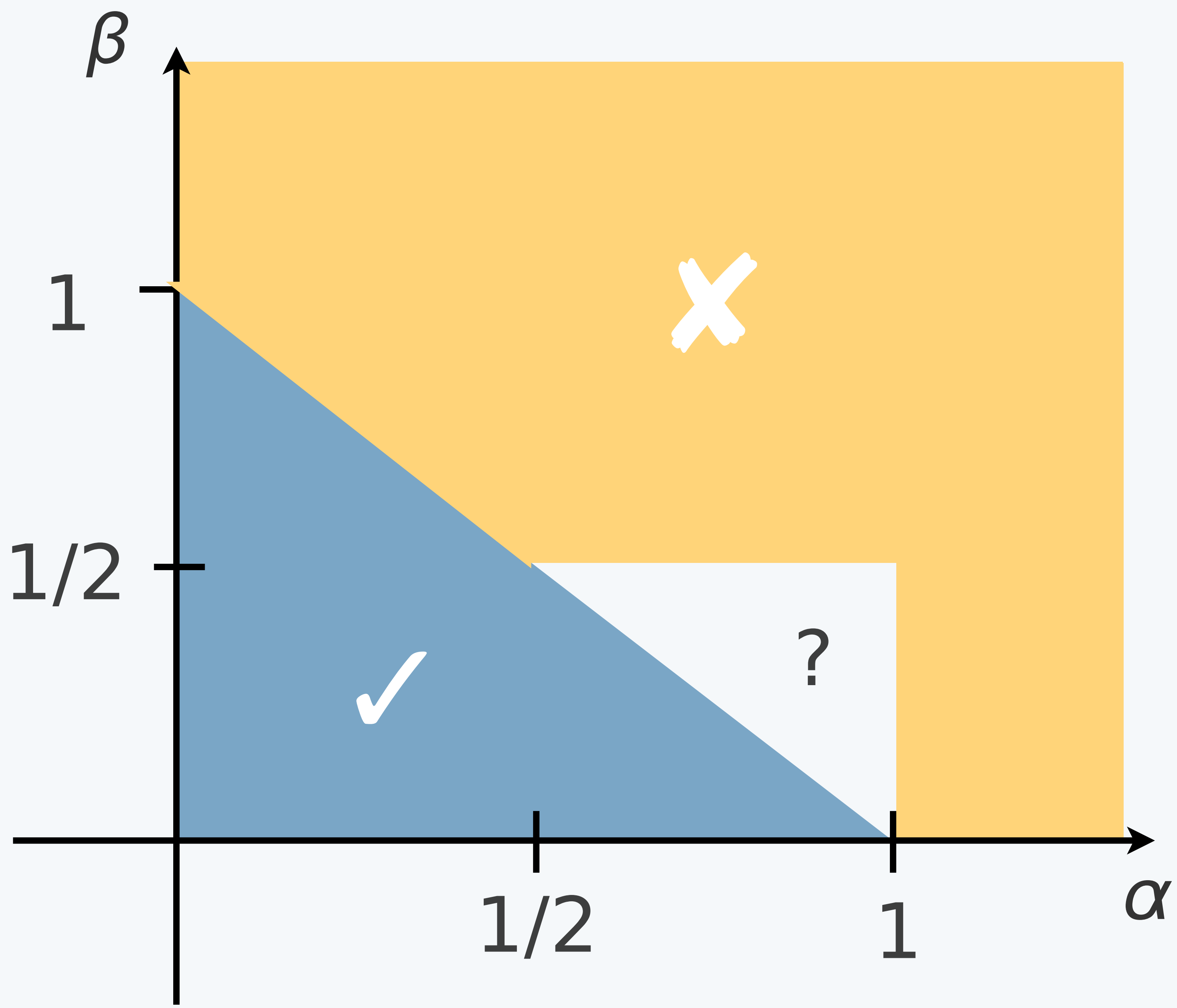
Must One Learn the Channel?

Appears so but...



Must One Learn the Channel?

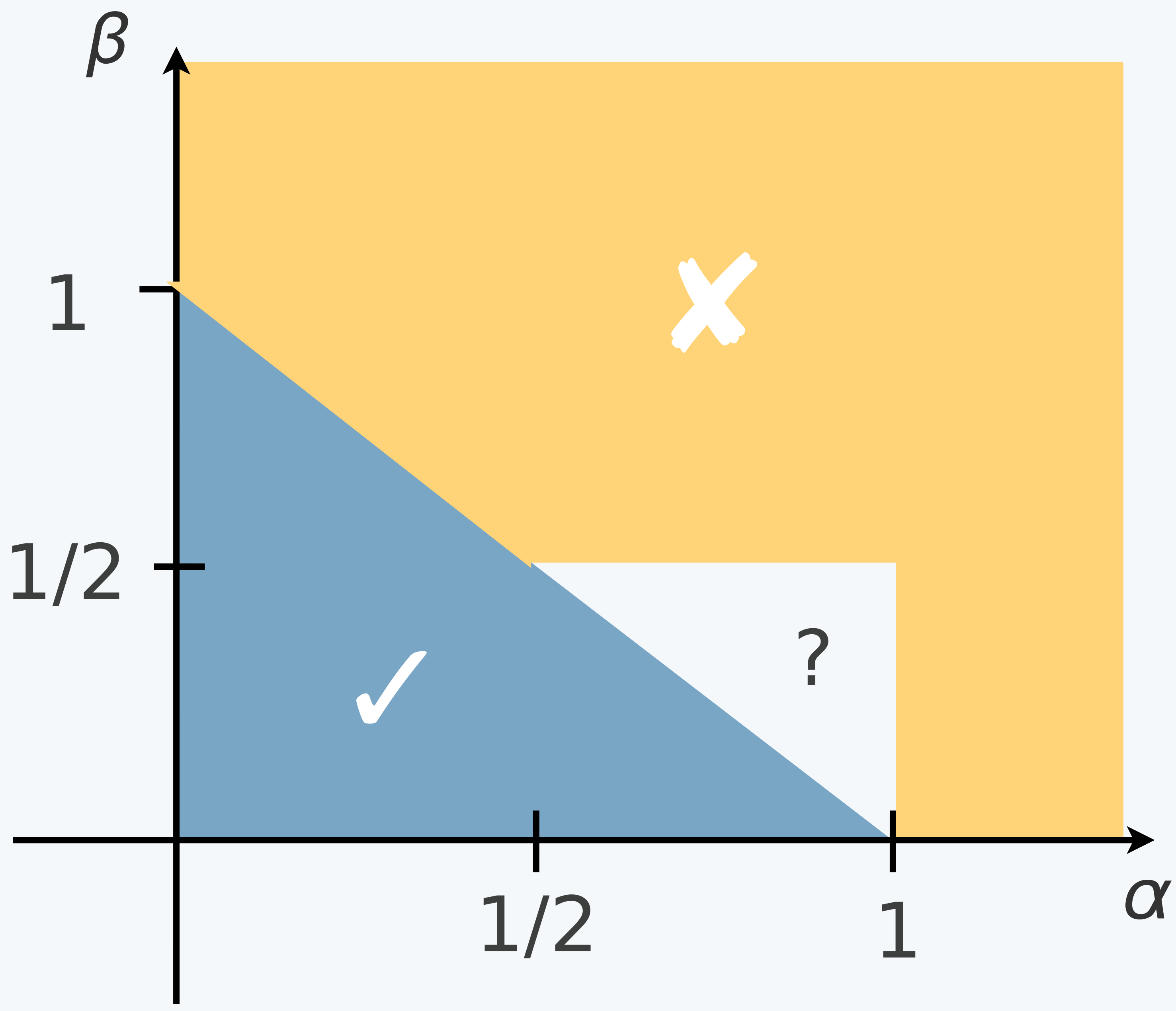
Appears so but...



- Sub-poly factors

Must One Learn the Channel?

Appears so but...



- Sub-poly factors
- Unknown channel drawn from some *class* (sim. Stéphane's talk)