Exact Asymptotics in Channel Coding

Aaron Wagner
(Joint work with Yücel Altuğ)

Cornell University
School of Electrical and Computer Engineering

I.I.D. Sums

- Let $\{X_i\}$ be i.i.d. with zero mean and unit var.

$$\text{Asymptotic behavior of } \sum_{i=1}^{N} X_i ?$$

I.I.D. Sums

- Let $\{X_i\}$ be i.i.d. with zero mean and unit var.

I.I.D. Sums

- Let $\{X_i\}$ be i.i.d. with zero mean and unit var.
- Weak Law of Large Numbers:

$$\lim_{N \to 0} \Pr\left(\sum_{i=1}^{N} X_i > \epsilon N\right) = 0 \qquad \epsilon > 0$$

I.I.D. Sums

- Let $\{X_i\}$ be i.i.d. with zero mean and unit var.

- Weak Law of Large Numbers:

$$\lim_{N\to 0} \Pr\left(\sum_{i=1}^{N} X_i > \epsilon N\right) = 0 \qquad \epsilon > 0$$

- Large Deviations:

$$\lim_{N\to 0} -\frac{1}{N} \log \Pr\left(\sum_{i=1}^{N} X_i > \epsilon N\right) = \Lambda^*(\epsilon)$$

I.I.D. Sums

- Let $\{X_i\}$ be i.i.d. with zero mean and unit var.

- Weak Law of Large Numbers:

$$\lim_{N \to 0} \Pr\left(\sum_{i=1}^{N} X_i > \epsilon N\right) = 0 \qquad \epsilon > 0$$

- Large Deviations:

$$\lim_{N \to 0} -\frac{1}{N} \log \Pr\left(\sum_{i=1}^{N} X_i > \epsilon N\right) = \Lambda^*(\epsilon)$$

- Central Limit Theorem:

$$\lim_{N \to 0} \Pr\left(\sum_{i=1}^{N} X_i > \epsilon \sqrt{N}\right) = Q(\epsilon)$$

I.I.D. Sums

I.I.D. Sums

- Exact Asymptotics [Bahadur-Rao '60]:

$$\lim_{N \to 0} \frac{\Pr\left(\sum_{i=1}^{N} X_i > N\epsilon\right)}{\frac{c}{\sqrt{N}} 2^{-N\Lambda^*(\epsilon)}} = 1$$

I.I.D. Sums

- Exact Asymptotics [Bahadur-Rao '60]:

$$\lim_{N\to 0} \frac{\Pr\left(\sum_{i=1}^{N} X_i > N\epsilon\right)}{\frac{c}{\sqrt{N}} 2^{-N\Lambda^*(\epsilon)}} = 1$$

- Moderate Deviations: if β is in (1/2, 1):

$$\lim_{N\to 0} \frac{1}{N^{2\beta-1}} \log \Pr\left(\sum_{i=1}^{N} X_i > \epsilon N^{\beta}\right) = \Lambda^*_{\mathcal{N}}(\epsilon)$$

Channel Coding

Discrete memoryless channel: *W(y|x)*

$x \in \mathcal{X}$

$y \in \mathcal{Y}$

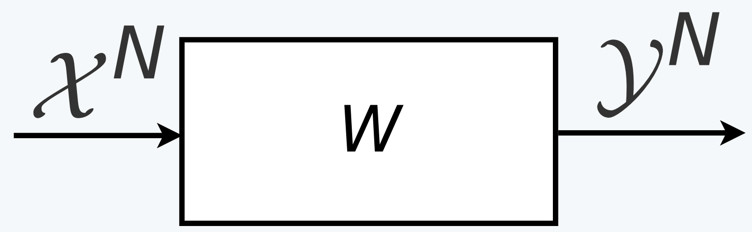
$$\boxed{W}$$

Channel Coding

Discrete memoryless channel: $W(y|x)$

$x \in \mathcal{X}$

$y \in \mathcal{Y}$

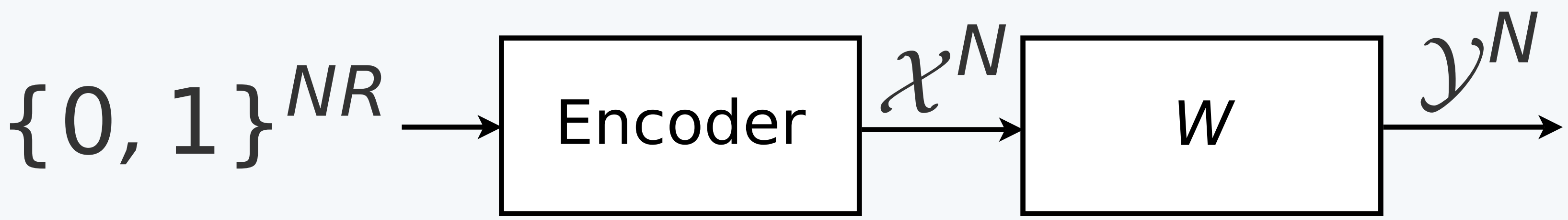


Channel Coding

Discrete memoryless channel: *W(y|x)*

$x \in \mathcal{X}$

$y \in \mathcal{Y}$



Channel Coding

Discrete memoryless channel: *W(y|x)*

$x \in \mathcal{X}$

$y \in \mathcal{Y}$

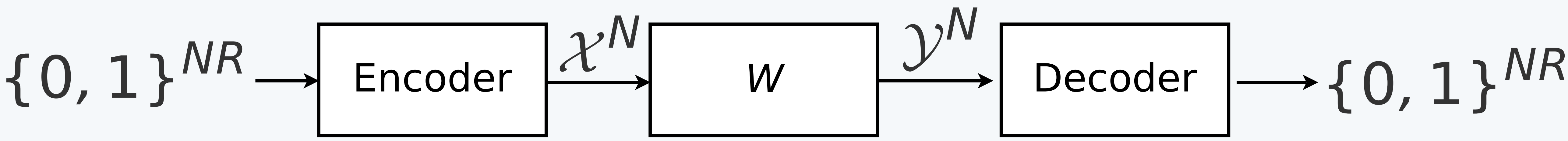
$$\{0,1\}^{NR} \longrightarrow \boxed{\text{Encoder}} \xrightarrow{\mathcal{X}^N} \boxed{W} \xrightarrow{\mathcal{Y}^N} \boxed{\text{Decoder}} \longrightarrow \{0,1\}^{NR}$$

Channel Coding

Discrete memoryless channel: *W(y|x)*

$x \in \mathcal{X}$

$y \in \mathcal{Y}$

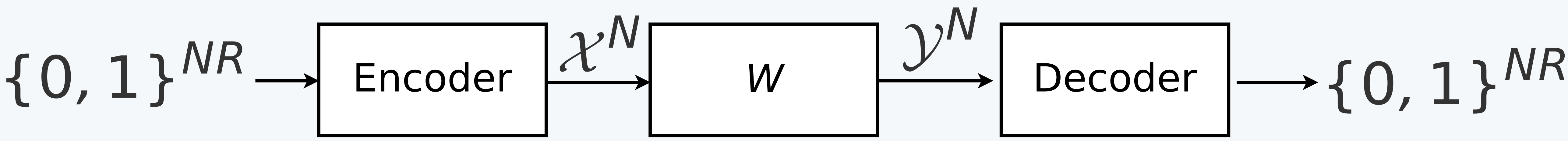


Parameters:

Channel Coding

Discrete memoryless channel: $W(y|x)$

$$x \in \mathcal{X}$$
$$y \in \mathcal{Y}$$



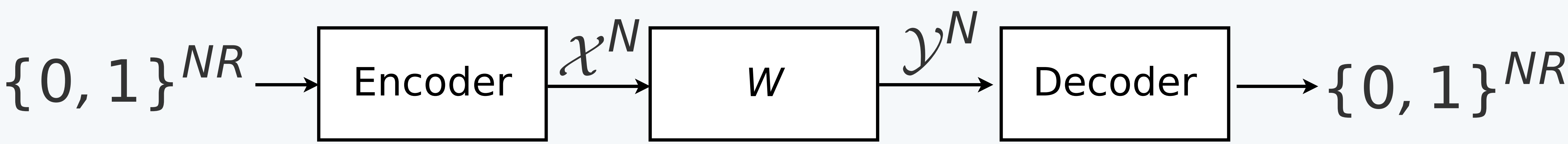
Parameters:

- N: blocklength = # channel uses

Channel Coding

Discrete memoryless channel: *W(y|x)*

$x \in \mathcal{X}$
$y \in \mathcal{Y}$



Parameters:

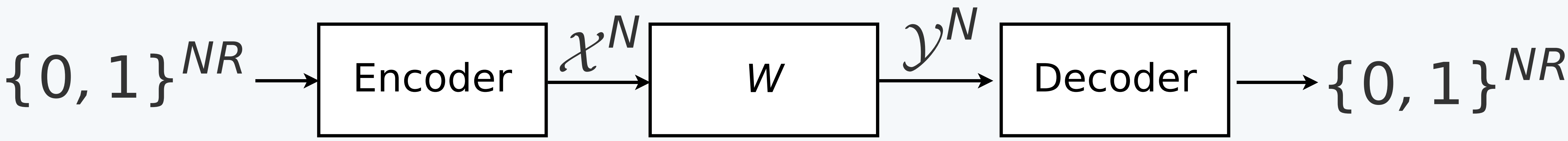
- N: blocklength = # channel uses
- R: data rate (bits per channel use)

Channel Coding

Discrete memoryless channel: *W(y|x)*

$x \in \mathcal{X}$

$y \in \mathcal{Y}$



Parameters:

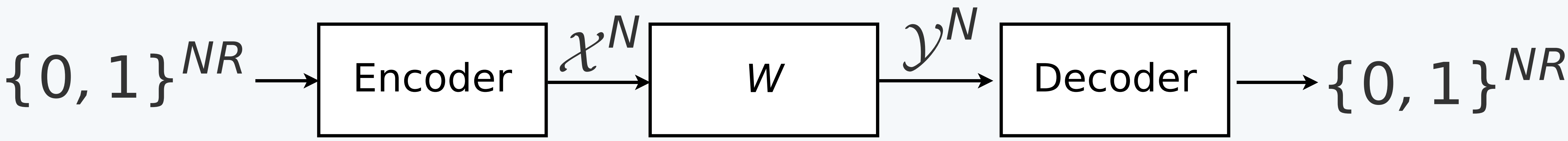
- N: blocklength = # channel uses
- R: data rate (bits per channel use)
- P_e: error probability

Channel Coding

Discrete memoryless channel: $W(y|x)$

$x \in \mathcal{X}$

$y \in \mathcal{Y}$



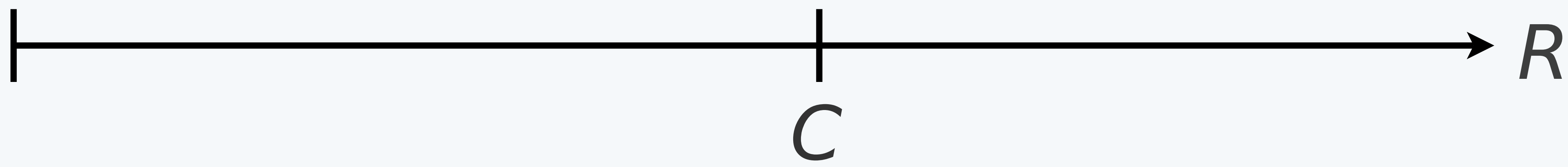
Parameters:

- N: blocklength = # channel uses
- R: data rate (bits per channel use)
- P_e: error probability

$$P_e(N,R) := \min\{P_e : \text{ code is } (N,R)\}$$

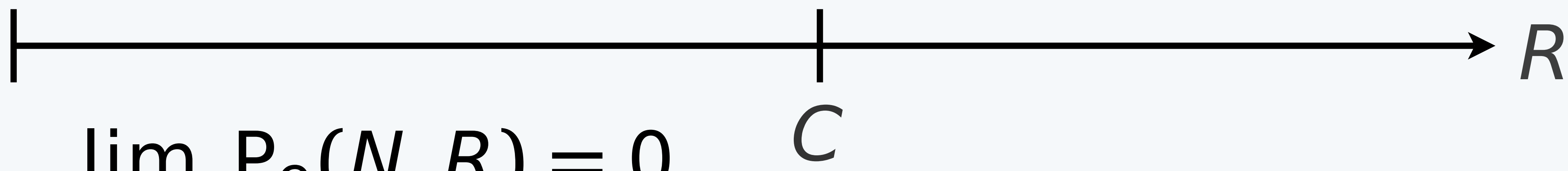
Channel Coding Theorem [Shannon '48]

Channel capacity: $C := \max_{P_X} \mathrm{I}(P_X; W)$



Channel Coding Theorem [Shannon '48]

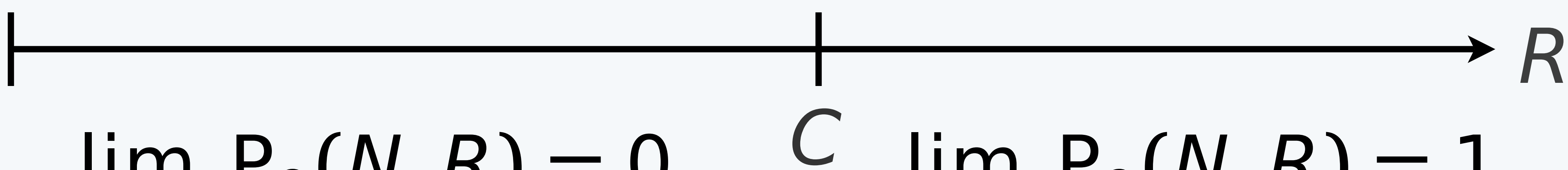
Channel capacity: $C := \max_{P_X} \mathrm{I}(P_X; W)$



$$\lim_{N\to\infty} \mathrm{P_e}(N, R) = 0$$

Channel Coding Theorem [Shannon '48]

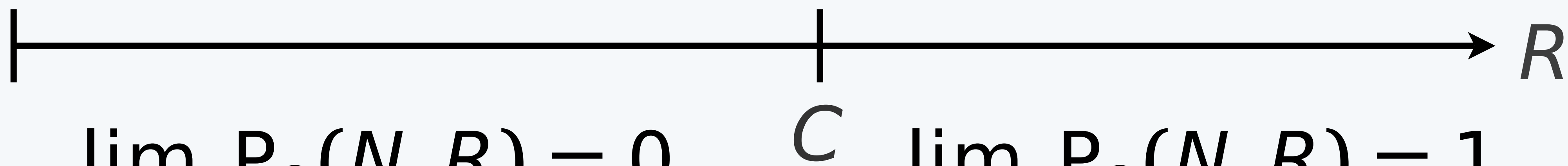
Channel capacity: $C := \max_{P_X} \mathrm{I}(P_X; W)$



$$\lim_{N\to\infty} \mathrm{P_e}(N,R) = 0 \qquad \lim_{N\to\infty} \mathrm{P_e}(N,R) = 1$$

Channel Coding Theorem [Shannon ’48]

Channel capacity: $C := \max_{P_X} \mathrm{I}(P_X; W)$



$$\lim_{N\to\infty} \mathrm{P_e}(N,R) = 0 \qquad \lim_{N\to\infty} \mathrm{P_e}(N,R) = 1$$

Speed of convergence?

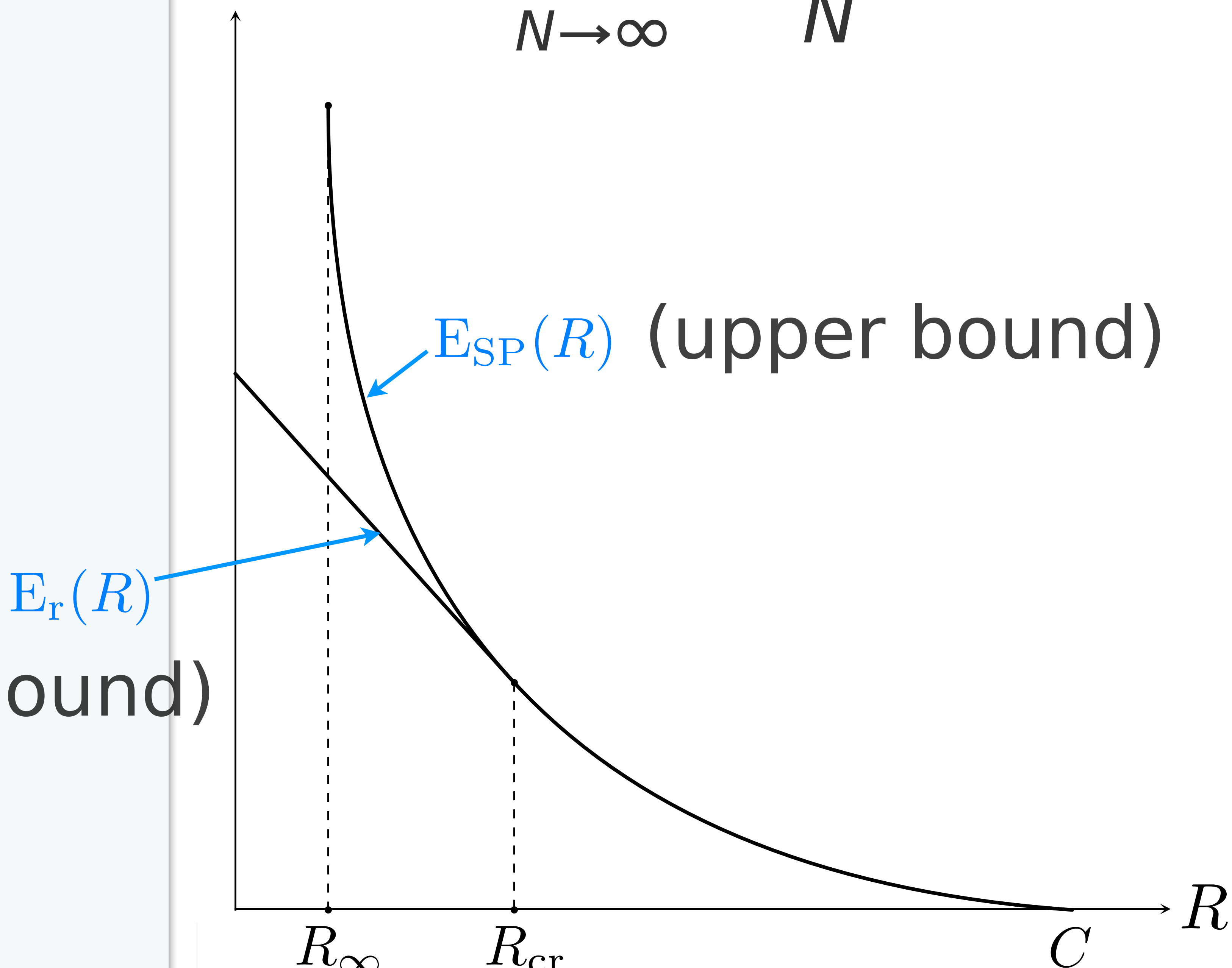
Error Exponents

$P_e(N,R)$ decays exponentially with N

Error Exponents

$P_e(N,R)$ decays exponentially with N

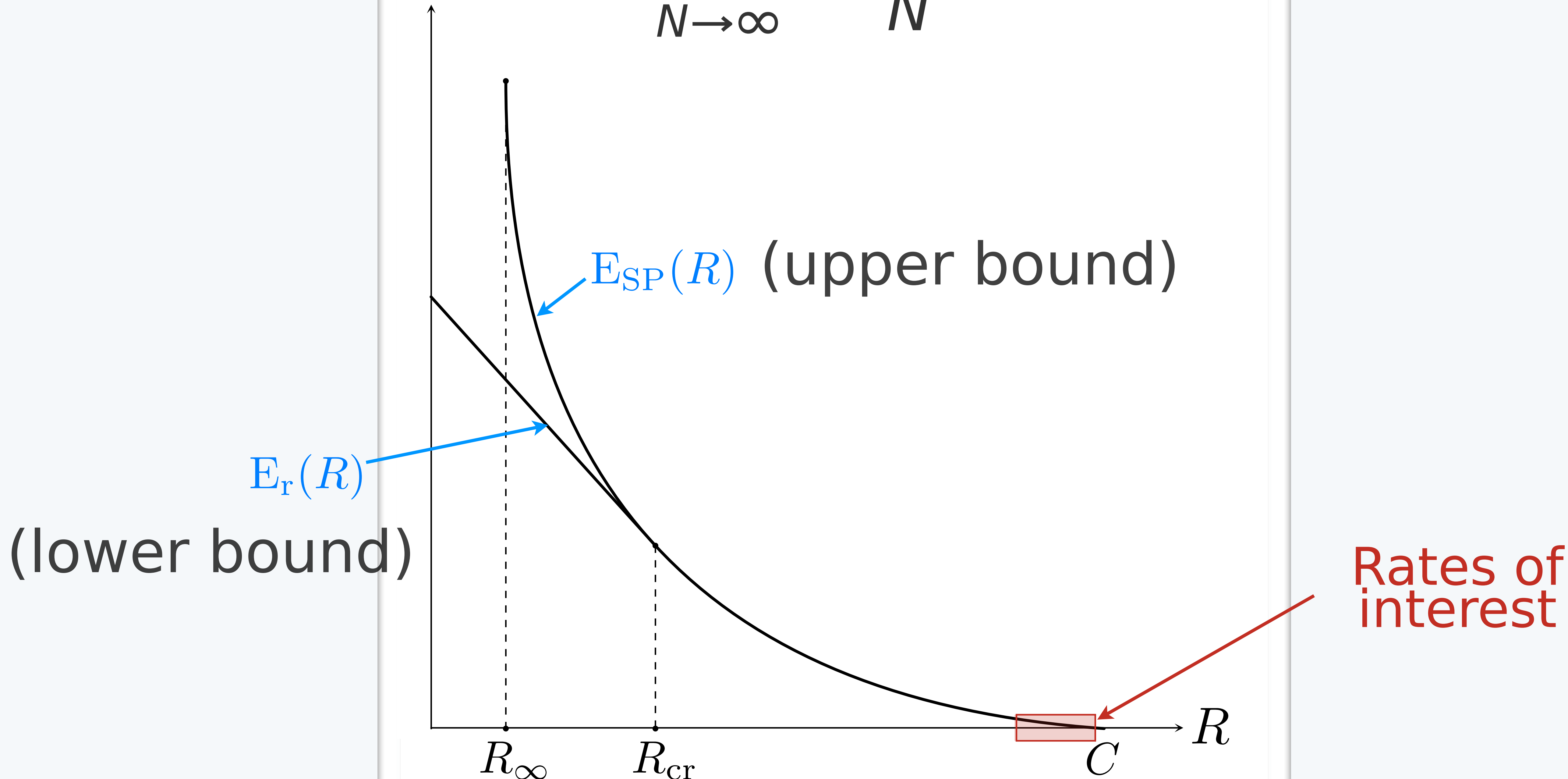
$$E(R) := \limsup_{N\to\infty} -\frac{1}{N} \log P_e(N,R)$$



Error Exponents

$\mathrm{P_e}(N,R)$ decays exponentially with N

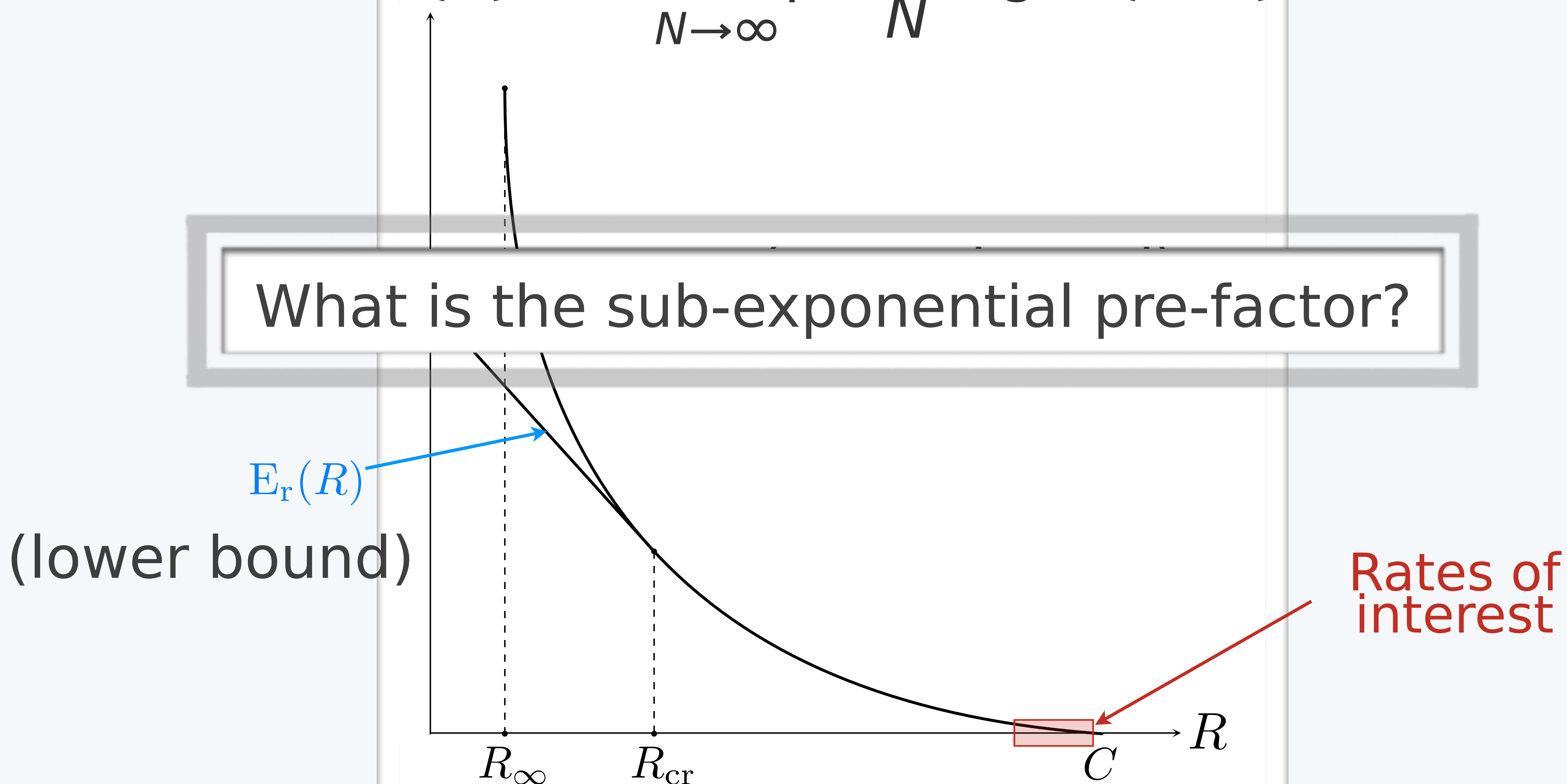
$$E(R) := \limsup_{N\to\infty} -\frac{1}{N}\log \mathrm{P_e}(N,R)$$



Error Exponents

$\mathrm{P_e}(N,R)$ decays exponentially with N

$$E(R) := \limsup_{N\to\infty} -\frac{1}{N}\log \mathrm{P_e}(N,R)$$



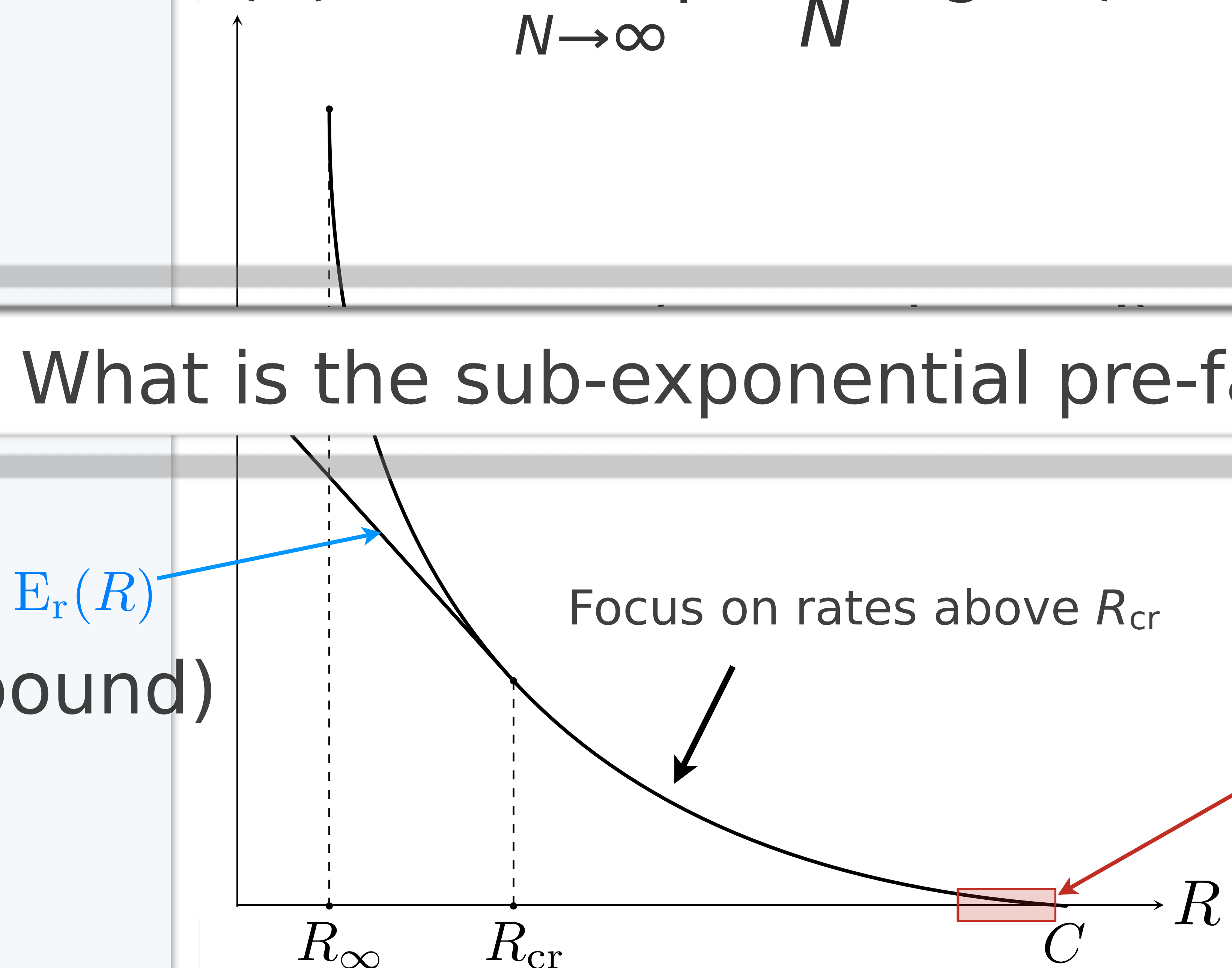
What is the sub-exponential pre-factor?

Error Exponents

$\mathrm{P_e}(N,R)$ decays exponentially with N

$$E(R) := \limsup_{N\to\infty} -\frac{1}{N}\log \mathrm{P_e}(N,R)$$

What is the sub-exponential pre-factor?



(lower bound)

Rates of interest

Pre-factors From the Literature

Upper bounds

Lower bounds

Pre-factors From the Literature

Fano '61: $2 \cdot 2^{-NE(R)}$

Upper bounds

Lower bounds

Pre-factors From the Literature

Fano '61:	$2 \cdot 2^{-NE(R)}$
Gallager '65:	$2^{-NE(R)}$

Upper bounds

Lower bounds

Pre-factors From the Literature

Csiszár *et al.* ’77: $O(N^{2|\mathcal{X}||\mathcal{Y}|})2^{-NE(R)}$ [Universal]

Fano ’61: $2 \cdot 2^{-NE(R)}$

Gallager ’65: $2^{-NE(R)}$

Upper bounds

Lower bounds

Pre-factors From the Literature

Csiszár *et al.* '77: $O(N^{2|\mathcal{X}||\mathcal{Y}|})2^{-NE(R)}$ [Universal]

Fano '61: $2 \cdot 2^{-NE(R)}$

Gallager '65: $2^{-NE(R)}$

Upper bounds

Lower bounds

Shannon *et al.* '67: $\Omega(2^{-\sqrt{N}})2^{-NE(R)}$

Pre-factors From the Literature

Csiszár *et al.* '77: $O(N^{2|\mathcal{X}||\mathcal{Y}|})2^{-NE(R)}$ [Universal]

Fano '61: $2 \cdot 2^{-NE(R)}$

Gallager '65: $2^{-NE(R)}$

Upper bounds

Lower bounds

Haroutunian '68: $\Omega(N^{-|\mathcal{X}||\mathcal{Y}|})2^{-NE(R)}$

Shannon *et al.* '67: $\Omega(2^{-\sqrt{N}})2^{-NE(R)}$

Pre-factors From the Literature

Csiszár *et al.* '77: $O(N^{2|\mathcal{X}||\mathcal{Y}|})2^{-NE(R)}$ [Universal]

Fano '61: $2 \cdot 2^{-NE(R)}$

Gallager '65: $2^{-NE(R)}$

Upper bounds

Lower bounds

Haroutunian '68: $\Omega(N^{-|\mathcal{X}||\mathcal{Y}|})2^{-NE(R)}$

Valembois-Fossorier '04: $\Omega(2^{-\sqrt{N}})2^{-NE(R)}$?

Shannon *et al.* '67: $\Omega(2^{-\sqrt{N}})2^{-NE(R)}$

Initial Guess

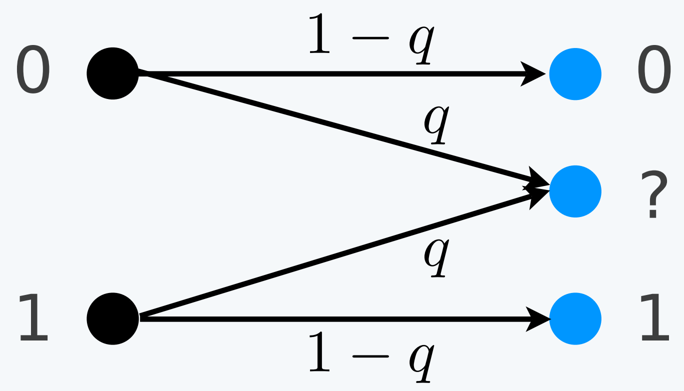
Initial Guess

- Exact Asymptotics [Bahadur-Rao ’60]:

$$\lim_{N \to 0} \frac{\Pr\left(\sum_{i=1}^{N} X_i > N\epsilon\right)}{\frac{c}{\sqrt{N}} 2^{-N\Lambda^*(\epsilon)}} = 1$$

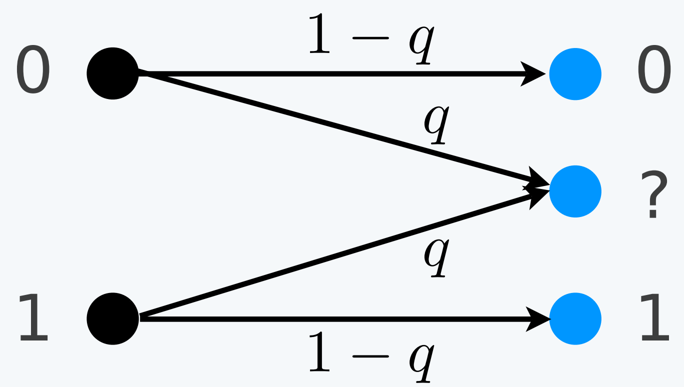
Elias '55

- Binary Erasure Channel (BEC):



Elias ’55

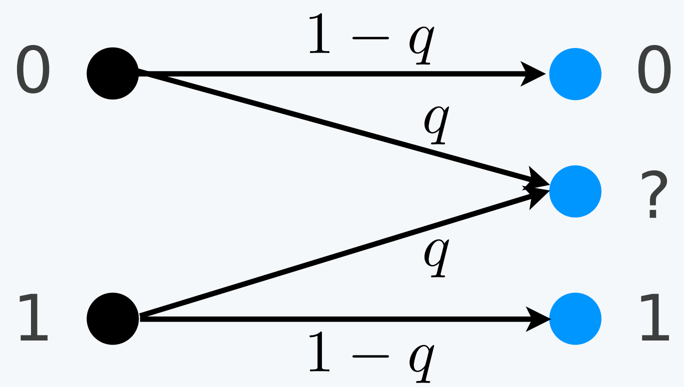
- <u>Binary Erasure Channel (BEC):</u>



$$P_e(N,R) = \Theta\left(N^{-0.5}\right) 2^{-NE(R)}$$

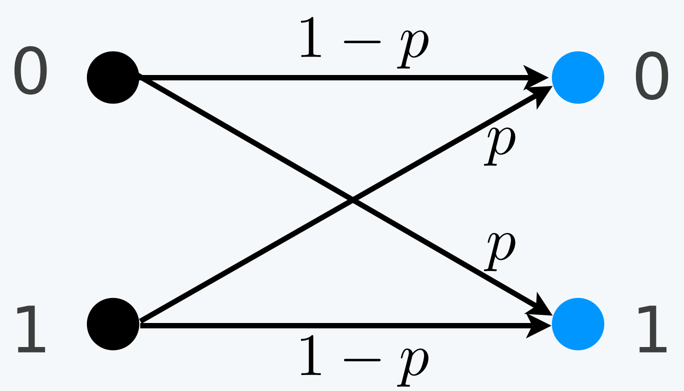
Elias '55

- Binary Erasure Channel (BEC):



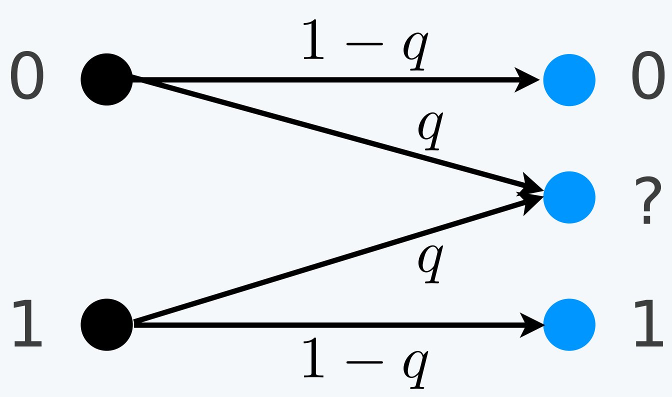
$$P_e(N, R) = \Theta\left(N^{-0.5}\right) 2^{-NE(R)}$$

- Binary Symmetric Channel (BSC):



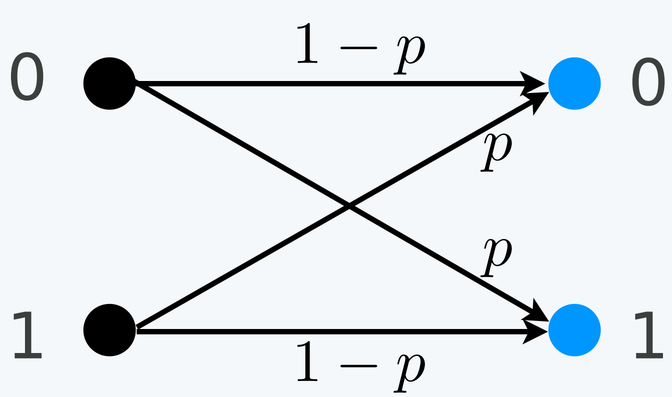
Elias ’55

- <u>Binary Erasure Channel (BEC):</u>



$$P_e(N,R) = \Theta\left(N^{-0.5}\right) 2^{-NE(R)}$$

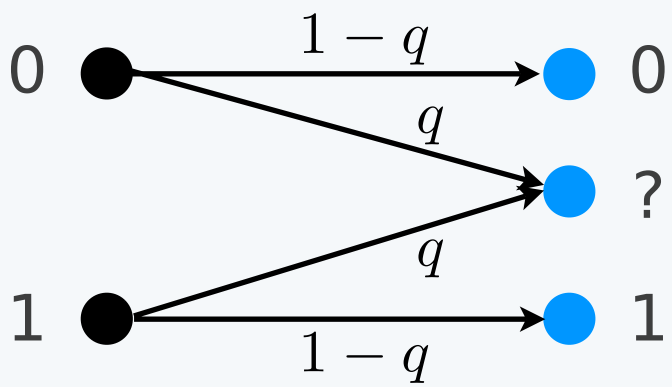
- <u>Binary Symmetric Channel (BSC):</u>



$$P_e(N,R) = \Theta\left(N^{-0.5(1+|E'(R)|)}\right) 2^{-NE(R)}$$

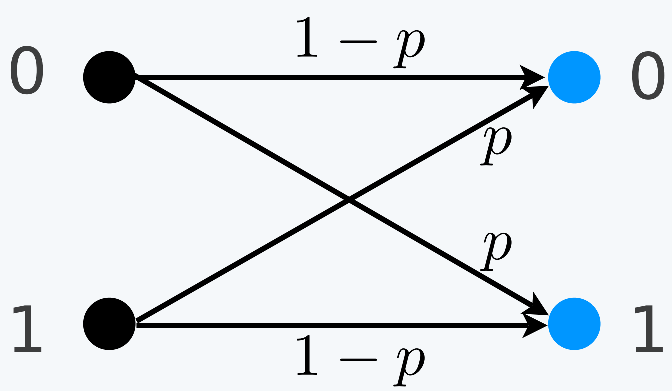
Elias '55

- Binary Erasure Channel (BEC):



$$P_e(N,R) = \Theta\left(N^{-0.5}\right) 2^{-NE(R)}$$

- Binary Symmetric Channel (BSC):



$$P_e(N,R) = \Theta\left(N^{-0.5(1+|E'(R)|)}\right) 2^{-NE(R)}$$

Is it a "zoo?"

Symmetric Channels

Theorem [Altuğ-Wagner '12]:

Symmetric Channels

Theorem [Altuğ-Wagner ’12]:

- If the channel is *regular*, then

$$\mathrm{P_e}(N,R) = \Theta\left(N^{-0.5(1+|\mathrm{E}'(R)|)}\right) 2^{-N\mathrm{E}(R)}$$

Symmetric Channels

Theorem [Altuğ-Wagner ’12]:

- If the channel is *regular*, then

$$\mathrm{P_e}(N,R) = \Theta\left(N^{-0.5(1+|\mathrm{E}'(R)|)}\right) 2^{-N\mathrm{E}(R)}$$

- If the channel is *irregular*, then

$$\mathrm{P_e}(N,R) = \Theta\left(N^{-0.5}\right) 2^{-N\mathrm{E}(R)}$$

General Channels

Theorem [Altuğ-Wagner '12]: For any $\epsilon > 0$,

General Channels

Theorem [Altuğ-Wagner '12]: For any $\epsilon > 0$,

- If the channel is *regular*, then

$$\mathrm{P_e}(N,R) \leq \frac{K_1(R,\epsilon)}{N^{0.5(1+|\mathrm{E}'(R-)|-\epsilon)}} 2^{-N\mathrm{E}(R)}$$

General Channels

Theorem [Altuğ-Wagner '12]: For any $\epsilon > 0$,

- If the channel is *regular*, then

$$\mathrm{P_e}(N, R) \leq \frac{K_1(R, \epsilon)}{N^{0.5(1+|\mathrm{E}'(R-)|-\epsilon)}} 2^{-N\mathrm{E}(R)}$$

- If the channel is *irregular*, then

$$\mathrm{P_e}(N, R) \leq \frac{K_2(R)}{N^{0.5}} 2^{-N\mathrm{E}(R)}$$

General Channels

Theorem [Altuğ-Wagner '12]:
For any *(N,R)* constant composition code and any $\epsilon > 0$,

$$\mathrm{P}_e \geq \frac{K_3(R,\epsilon)}{N^{0.5(1+|\mathrm{E}'(R-)|+\epsilon)}} 2^{-N\mathrm{E}(R)}$$

Summary of Results

Pre-factor Bounds	Symmetric	Asymmetric
Regular	$O(N^{0.5(1+\lvert\mathrm{E}'(R)\rvert)})$ $\Omega(N^{0.5(1+\lvert\mathrm{E}'(R)\rvert)})$	$O(N^{0.5(1+\lvert\mathrm{E}'(R-)\rvert)})$ $\Omega(N^{0.5(1+\lvert\mathrm{E}'(R-)\rvert)})$
Irregular	$O(N^{0.5})$ $\Omega(N^{0.5})$	$O(N^{0.5})$ $\Omega(N^{0.5(1+\lvert\mathrm{E}'(R-)\rvert)})$

Red: constant composition codes only

Summary of Results

What is symmetric?

Pre-factor Bounds	Symmetric	Asymmetric
Regular	$O(N^{0.5(1+\lvert E'(R)\rvert)})$ $\Omega(N^{0.5(1+\lvert E'(R)\rvert)})$	$O(N^{0.5(1+\lvert E'(R-)\rvert)})$ $\Omega(N^{0.5(1+\lvert E'(R-)\rvert)})$
Irregular	$O(N^{0.5})$ $\Omega(N^{0.5})$	$O(N^{0.5})$ $\Omega(N^{0.5(1+\lvert E'(R-)\rvert)})$

Red: constant composition codes only

Summary of Results

Pre-factor Bounds	Symmetric	Asymmetric
Regular	$O(N^{0.5(1+\lvert E'(R)\rvert)})$ $\Omega(N^{0.5(1+\lvert E'(R)\rvert)})$	$O(N^{0.5(1+\lvert E'(R-)\rvert)})$ $\Omega(N^{0.5(1+\lvert E'(R-)\rvert)})$
Irregular	$O(N^{0.5})$ $\Omega(N^{0.5})$	$O(N^{0.5})$ $\Omega(N^{0.5(1+\lvert E'(R-)\rvert)})$

What

is regular?

Red: constant composition codes only

Summary of Results

Pre-factor Bounds	Symmetric	ric
Regular	$O(N^{0.5(1+\lvert E'(R)\rvert)})$ $\Omega(N^{0.5(1+\lvert E'(R)\rvert)})$	$O(N^{0.5(1+\lvert E'(R-)\rvert)})$ $\Omega(N^{0.5(1+\lvert E'(R-)\rvert)})$
Irregular	$O(N^{0.5})$ $\Omega(N^{0.5})$	$O(N^{0.5})$ $\Omega(N^{0.5(1+\lvert E'(R-)\rvert)})$

Why this?

Red: constant composition codes only

Summary of Results

Pre-factor Bounds	Symmetric	Asymmetric
Regular	$O(N^{0.5(1+\|E'(R)\|)})$ $\Omega(N^{0.5(1+\|E'(R)\|)})$	$O(N^{0.5(1+\|E'(R-)\|)})$ $\Omega(N^{0.5(1+\|E'(R-)\|)})$
Irregular	$O(N^{0.5})$ $\Omega(N^{0.5})$	$O(N^{0.5})$ $\Omega(N^{0.5(1+\|E'(R-)\|)})$

Red: constant composition codes only

Symmetric Channels

Definition (Gallager): A channel (stochastic matrix) W is *symmetric* if its columns can be partitioned so that, within each partition, the columns are permutations of each other, as are the rows.

Symmetric Channels

Definition (Gallager): A channel (stochastic matrix) W is *symmetric* if its columns can be partitioned so that, within each partition, the columns are permutations of each other, as are the rows.

$$\begin{bmatrix} 1-\epsilon & 0 & \epsilon \\ 0 & 1-\epsilon & \epsilon \end{bmatrix}$$

Symmetric Channels

Definition (Gallager): A channel (stochastic matrix) W is *symmetric* if its columns can be partitioned so that, within each partition, the columns are permutations of each other, as are the rows.

$$\left[\begin{array}{cc|c} 1-\epsilon & 0 & \epsilon \\ 0 & 1-\epsilon & \epsilon \end{array}\right]$$

Symmetric Channels

Definition (Gallager): A channel (stochastic matrix) W is *symmetric* if its columns can be partitioned so that, within each partition, the columns are permutations of each other, as are the rows.

$$\left[\begin{array}{cc|c} 1-\epsilon & 0 & \epsilon \\ 0 & 1-\epsilon & \epsilon \end{array}\right]$$

Symmetric

Symmetric Channels

Definition (Gallager): A channel (stochastic matrix) W is *symmetric* if its columns can be partitioned so that, within each partition, the columns are permutations of each other, as are the rows.

$$\left[\begin{array}{cc|c} 1-\epsilon & 0 & \epsilon \\ 0 & 1-\epsilon & \epsilon \end{array}\right]$$

Symmetric

$$\left[\begin{array}{cc} 3/4 & 1/4 \\ 1/3 & 2/3 \end{array}\right]$$

Symmetric Channels

Definition (Gallager): A channel (stochastic matrix) W is *symmetric* if its columns can be partitioned so that, within each partition, the columns are permutations of each other, as are the rows.

$$\left[\begin{array}{cc|c} 1-\epsilon & 0 & \epsilon \\ 0 & 1-\epsilon & \epsilon \end{array}\right]$$

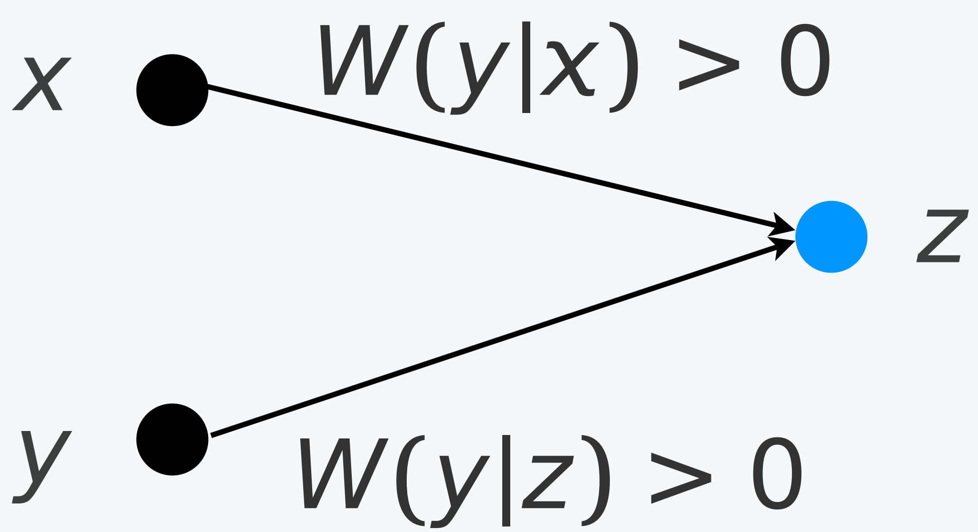
Symmetric

$$\begin{bmatrix} 3/4 & 1/4 \\ 1/3 & 2/3 \end{bmatrix}$$

Not symmetric

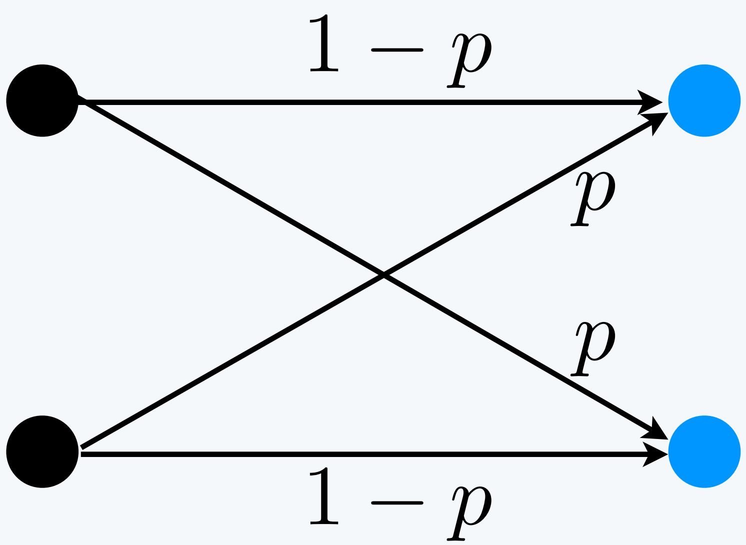
Regular Channels

Definition: A symmetric channel W is *regular* if for some x, y, and z:



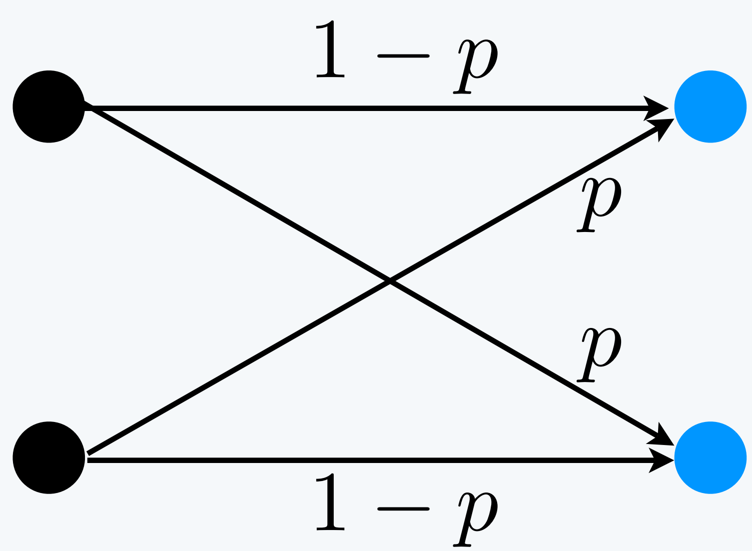
$$W(y|x) \neq W(y|z)$$

Regular Channels

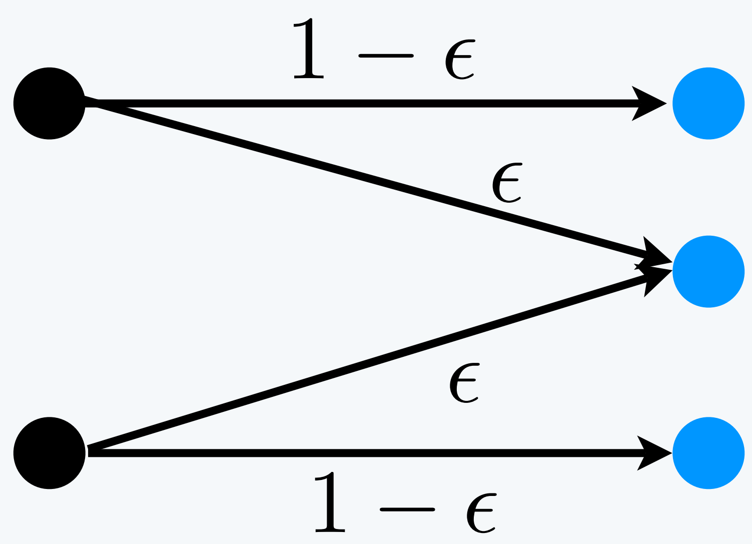


Regular

Regular Channels



Regular



Irregular

Summary of Results

Pre-factor Bounds	Symmetric	Asymmetric
Regular	$O(N^{0.5(1+\|E'(R)\|)})$ $\Omega(N^{0.5(1+\|E'(R)\|)})$	$O(N^{0.5(1+\|E'(R-)\|)})$ $\Omega(N^{0.5(1+\|E'(R-)\|)})$
Irregular	$O(N^{0.5})$ $\Omega(N^{0.5})$	$O(N^{0.5})$ $\Omega(N^{0.5(1+\|E'(R-)\|)})$

Red: constant composition codes only

Summary of Results

Pre-factor Bounds	Symmetric	…ric
Regular	$O(N^{0.5(1+\|E'(R)\|)})$ $\Omega(N^{0.5(1+\|E'(R)\|)})$	$O(N^{0.5(1+\|E'(R-)\|)})$ $\Omega(N^{0.5(1+\|E'(R-)\|)})$
Irregular	$O(N^{0.5})$ $\Omega(N^{0.5})$	$O(N^{0.5})$ $\Omega(N^{0.5(1+\|E'(R-)\|)})$

Why this?

Red: constant composition codes only

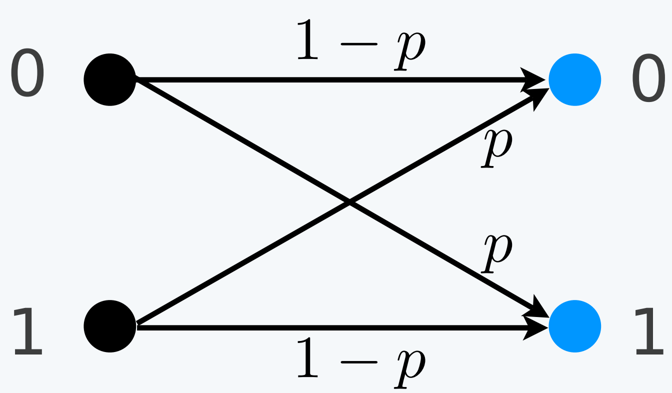
Summary of Results

Pre-factor Bounds	Symmetric	Asymmetric
Regular	$O(N^{0.5(1+\lvert E'(R)\rvert)})$ $\Omega(N^{0.5(1+\lvert E'(R)\rvert)})$	$O(N^{0.5(1+\lvert E'(R-)\rvert)})$ $\Omega(N^{0.5(1+\lvert E'(R-)\rvert)})$
Irregular	$O(N^{0.5})$ $\Omega(N^{0.5})$	$O(N^{0.5})$ $\Omega(N^{0.5(1+\lvert E'(R-)\rvert)})$

Red: constant composition codes only

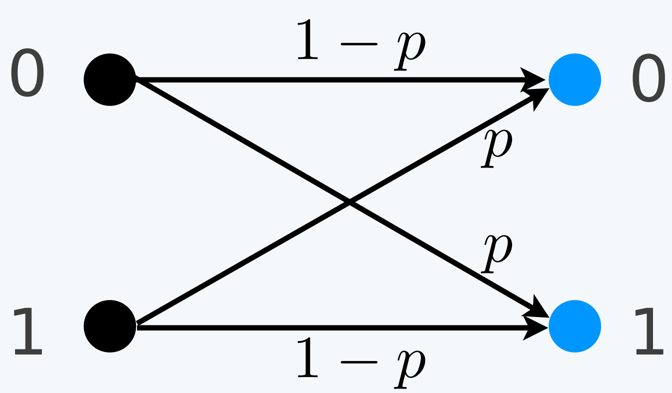
Why the Slope Thingy?

Binary Symmetric Channel (BSC):



Why the Slope Thingy?

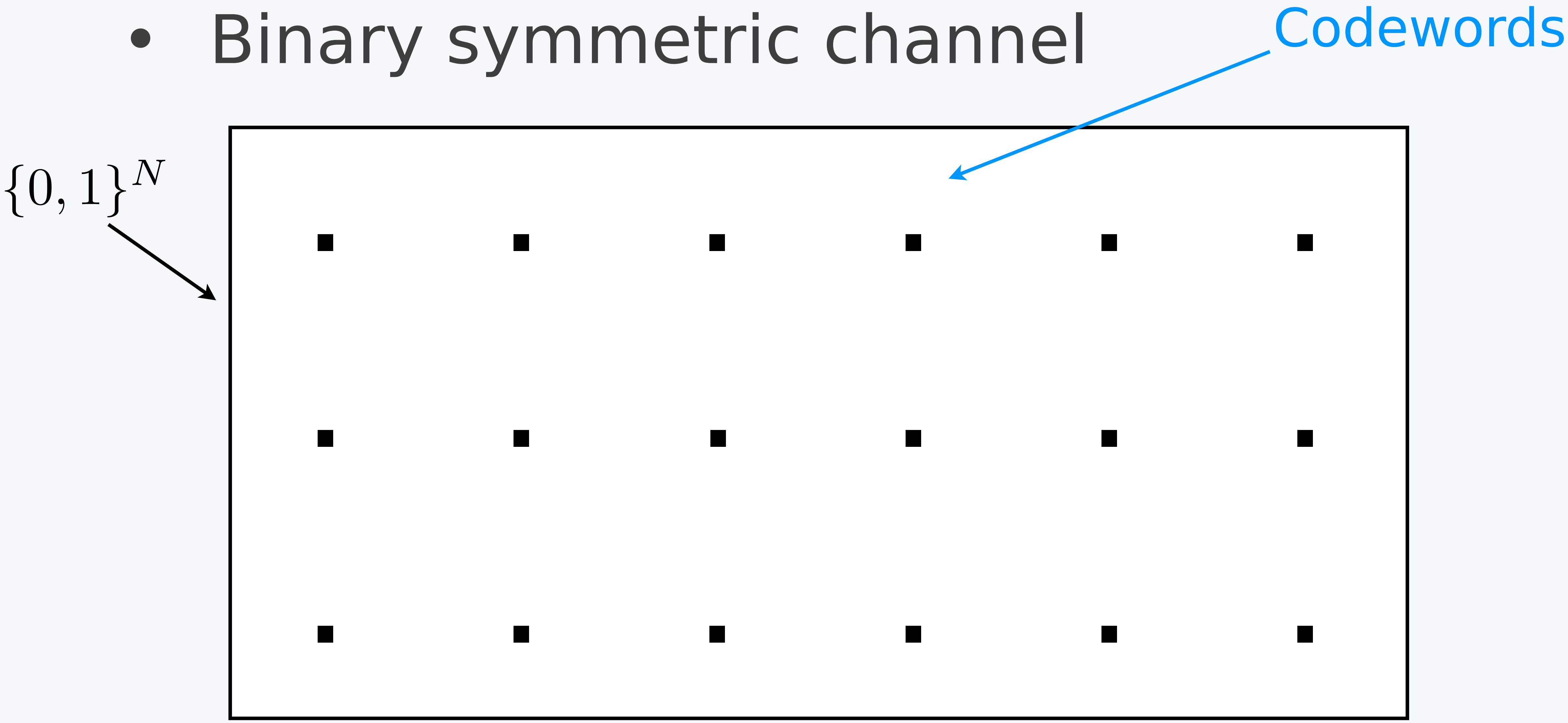
Binary Symmetric Channel (BSC):



$$\mathrm{P_e}(N,R) = \Theta\left(N^{-0.5(1+|\mathrm{E}'(R)|)}\right) 2^{-N\mathrm{E}(R)}$$

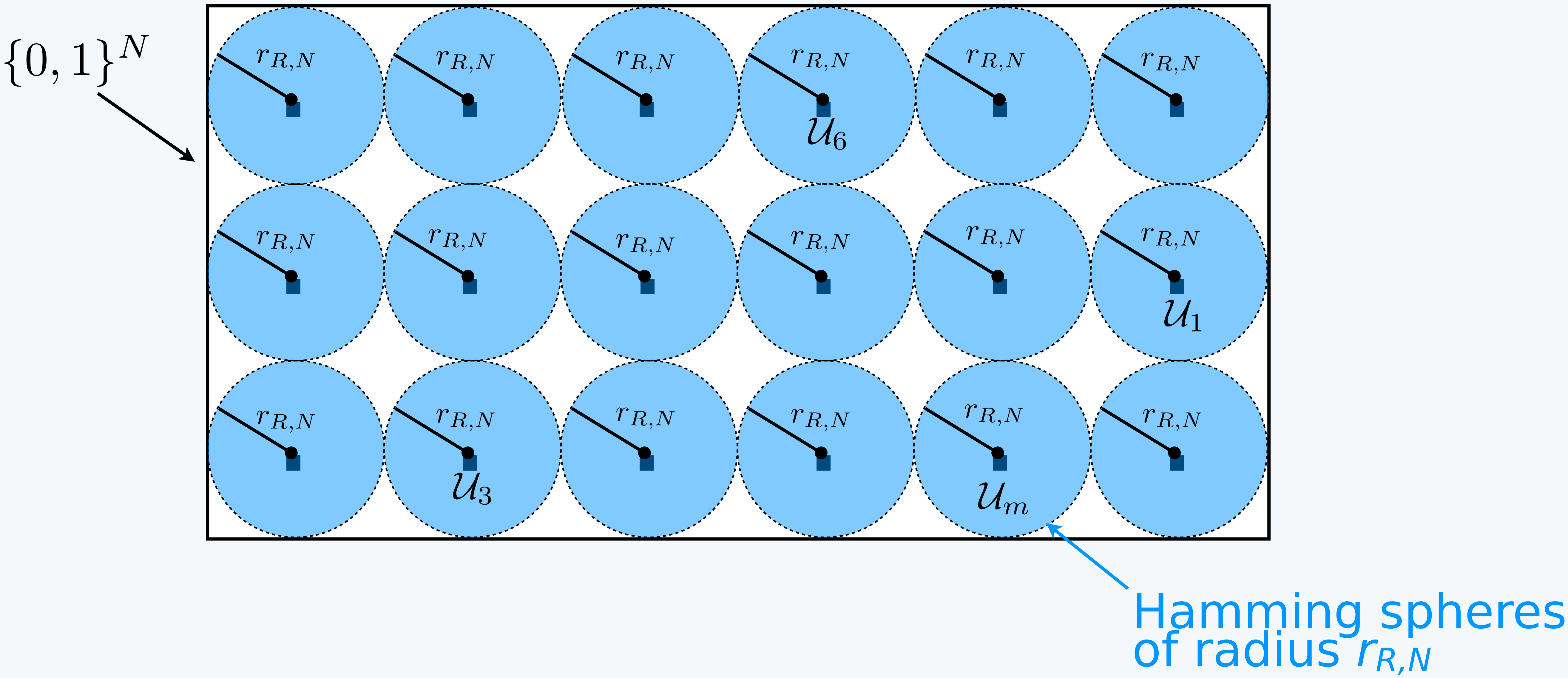
Packing Spheres

- Binary symmetric channel



Packing Spheres

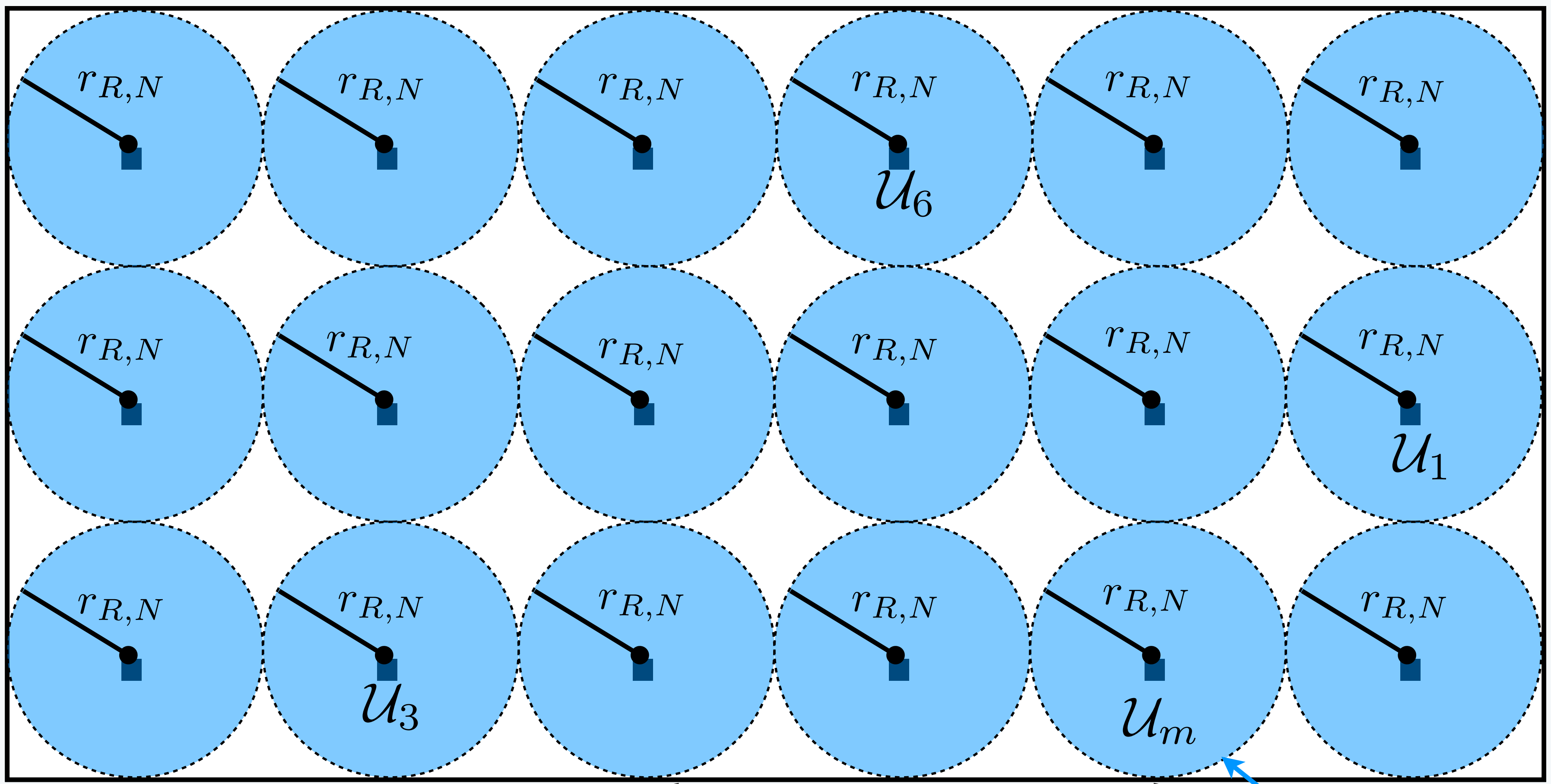
- Binary symmetric channel



Packing Spheres

- Binary symmetric channel

$\{0,1\}^N$



Hamming spheres of radius $r_{R,N}$

$$P_e(N,R) \approx \Pr\left(\sum_{i=1}^{N} T_i > r_{R,N}\right)$$

i.i.d. channel bit flips

$$\approx \frac{c}{\sqrt{N}} 2^{-N\cdot E(R)}?$$

How to Compute $r_{R,N}$?

Vol. of a Hamming sphere with radius $r_{R,N}$?

How to Compute $r_{R,N}$?

Vol. of a Hamming sphere with radius $r_{R,N}$?

$$\frac{1}{\sqrt{N}} 2^{N \mathrm{h}\left(\frac{r_{R,N}}{N}\right)}$$

How to Compute $r_{R,N}$?

Vol. of a Hamming sphere with radius $r_{R,N}$?

$$\frac{1}{\sqrt{N}} 2^{N \mathrm{h}\left(\frac{r_{R,N}}{N}\right)}$$

$$h(\theta) = -\theta \log \theta - (1-\theta) \log(1-\theta)$$

How to Compute $r_{R,N}$?

Vol. of a Hamming sphere with radius $r_{R,N}$?

$$\frac{1}{\sqrt{N}} 2^{N\mathrm{h}\left(\frac{r_{R,N}}{N}\right)} \qquad h(\theta) = -\theta \log \theta - (1-\theta)\log(1-\theta)$$

Vol. of each decoding region?

How to Compute $r_{R,N}$?

Vol. of a Hamming sphere with radius $r_{R,N}$?

$$\frac{1}{\sqrt{N}} 2^{N\mathrm{h}\left(\frac{r_{R,N}}{N}\right)} \qquad h(\theta) = -\theta \log \theta - (1-\theta) \log(1-\theta)$$

Vol. of each decoding region? $\dfrac{2^N}{2^{NR}}$

How to Compute $r_{R,N}$?

Vol. of a Hamming sphere with radius $r_{R,N}$?

$$\frac{1}{\sqrt{N}} 2^{N\mathrm{h}\left(\frac{r_{R,N}}{N}\right)} \qquad h(\theta) = -\theta \log \theta - (1-\theta)\log(1-\theta)$$

Vol. of each decoding region? $\dfrac{2^N}{2^{NR}}$

So $\quad r_{R,N} \approx N \cdot h^{-1}\left(1 - R + \frac{1}{N}\log\sqrt{N}\right)$

How to Compute $r_{R,N}$?

Vol. of a Hamming sphere with radius $r_{R,N}$?

$$\frac{1}{\sqrt{N}} 2^{N\mathrm{h}\left(\frac{r_{R,N}}{N}\right)} \qquad h(\theta) = -\theta \log \theta - (1-\theta)\log(1-\theta)$$

Vol. of each decoding region? $\dfrac{2^N}{2^{NR}}$

$$\text{So} \quad r_{R,N} \approx N \cdot h^{-1}\left(1 - R + \frac{1}{N}\log\sqrt{N}\right)$$

$$\mathrm{P_e}(N,R) \approx \Pr\left(\sum_{i=1}^{N} T_i > r_{R,N}\right) \approx \frac{1}{\sqrt{N}} 2^{-N\cdot E(R-(\log\sqrt{N})/N)}$$

How to Compute $r_{R,N}$?

Vol. of a Hamming sphere with radius $r_{R,N}$?

$$\frac{1}{\boxed{\sqrt{N}}} 2^{N\mathrm{h}\left(\frac{r_{R,N}}{N}\right)}$$

$$h(\theta) = -\theta \log \theta - (1-\theta)\log(1-\theta)$$

Vol. of each decoding region? $\dfrac{2^N}{2^{NR}}$

So $$r_{R,N} \approx N \cdot h^{-1}\left(1 - R + \boxed{\frac{1}{N}\log\sqrt{N}}\right)$$

$$\mathrm{P_e}(N,R) \approx \Pr\left(\sum_{i=1}^{N} T_i > r_{R,N}\right) \approx \frac{1}{\sqrt{N}} 2^{-N\cdot E(R - \boxed{(\log\sqrt{N})/N})}$$

Summary of Results

Pre-factor Bounds	Symmetric	Asymmetric
Regular	$O(N^{0.5(1+\lvert \mathrm{E}'(R)\rvert)})$ $\Omega(N^{0.5(1+\lvert \mathrm{E}'(R)\rvert)})$	$O(N^{0.5(1+\lvert \mathrm{E}'(R-)\rvert)})$ $\Omega(N^{0.5(1+\lvert \mathrm{E}'(R-)\rvert)})$
Irregular	$O(N^{0.5})$ $\Omega(N^{0.5})$	$O(N^{0.5})$ $\Omega(N^{0.5(1+\lvert \mathrm{E}'(R-)\rvert)})$

Red: constant composition codes only

Proof Technique

Proof Technique

- Overall approach:

Proof Technique

- Overall approach:
  - Mimic existing proofs for exponents …

Proof Technique

- Overall approach:
  - Mimic existing proofs for exponents ...
  - ... but use exact asymptotics instead of large deviations

Proof Technique

- Overall approach:
  - Mimic existing proofs for exponents ...
  - ... but use exact asymptotics instead of large deviations
- Upper bound:
  - Csiszar *et al.* ‘77
  - Fano ‘61
  - Gallager ‘65

Proof Technique

- Overall approach:
  - Mimic existing proofs for exponents ...
  - ... but use exact asymptotics instead of large deviations
- Upper bound:
  - Csiszar *et al.* ‘77
  - Fano ‘61
  - Gallager ‘65

Proof Technique

- Overall approach:
  - Mimic existing proofs for exponents ...
  - ... but use exact asymptotics instead of large deviations

- Upper bound:
  - Csiszar *et al.* ‘77
  - Fano ‘61
  - Gallager ‘65

- Lower bound:
  - Shannon *et al.* ‘67
  - Haroutunian ’68
  - Blahut ‘74

Proof Technique

- Overall approach:
  - Mimic existing proofs for exponents ...
  - ... but use exact asymptotics instead of large deviations

- Upper bound:
  - Csiszar *et al.* ‘77
  - Fano ‘61
  - Gallager ‘65

- Lower bound:
  - Shannon *et al.* ‘67
  - Haroutunian ’68
  - Blahut ‘74

Summary of Results

Pre-factor Bounds	Symmetric	Asymmetric
Regular	$O(N^{0.5(1+\lvert E'(R)\rvert)})$ $\Omega(N^{0.5(1+\lvert E'(R)\rvert)})$	$O(N^{0.5(1+\lvert E'(R-)\rvert)})$ $\Omega(N^{0.5(1+\lvert E'(R-)\rvert)})$
Irregular	$O(N^{0.5})$ $\Omega(N^{0.5})$	$O(N^{0.5})$ $\Omega(N^{0.5(1+\lvert E'(R-)\rvert)})$

Red: constant composition codes only

Summary of Results

Pre-factor Bounds	Symmetric	Asymmetric
Regular	$O(N^{0.5(1+\lvert E'(R)\rvert)})$ $\Omega(N^{0.5(1+\lvert E'(R)\rvert)})$	$O(N^{0.5(1+\lvert E'(R-)\rvert)})$ $\Omega(N^{0.5(1+\lvert E'(R-)\rvert)})$
Irregular	$O(N^{0.5})$ $\Omega(N^{0.5})$	$O(N^{0.5})$ $\Omega(N^{0.5(1+\lvert E'(R-)\rvert)})$

What about constants/lower order behavior?

Red: constant composition codes only

A Criticism of Error Exponents

> The venerable large-deviations information theory school seems to have made little headway in convincing engineers to transmit at, say, ... half of capacity for the sake of astronomically low block error probability in the asymptotic regime.

A Criticism of Error Exponents

The venerable large-deviations information theory school seems to have made little headway in convincing engineers to transmit at, say, ... half of capacity for the sake of astronomically low block error probability in the asymptotic regime.

Sergio Verdú (2007 Shannon lecture)

Normal Approximation

- $R^*(N,\epsilon) = \max R$ such that $P_e(N,R) \leq \epsilon$

Normal Approximation

- $R^*(N,\epsilon)$ = max R such that $P_e(N,R) \leq \epsilon$
- Fix an $0 < \epsilon < 0.5$

Normal Approximation

- $R^*(N,\epsilon)$ = max R such that $P_e(N,R) \leq \epsilon$
- Fix an $0 < \epsilon < 0.5$

- **Theorem:** [Strassen '62, Hayashi '09, Polyanskiy, Poor, and Verdú '10]
  As $N \to \infty$

$$R^*(N,\epsilon) = C - \sqrt{\frac{V}{N}}Q^{-1}(\epsilon) + O\left(\frac{\log N}{N}\right)$$

Normal Approximation

- $R^*(N,\epsilon)$ = max R such that $P_e(N,R) \le \epsilon$
- Fix an $0 < \epsilon < 0.5$

- **Theorem:** [Strassen '62, Hayashi '09, Polyanskiy, Poor, and Verdú '10]
  As $N \to \infty$

$$R^*(N,\epsilon) = C - \sqrt{\frac{\boxed{V}}{N}}Q^{-1}(\epsilon) + O\left(\frac{\log N}{N}\right)$$

Channel dispersion

Normal Approximation

- $R^*(N,\epsilon)$ = max R such that $P_e(N,R) \leq \epsilon$
- Fix an $0 < \epsilon < 0.5$

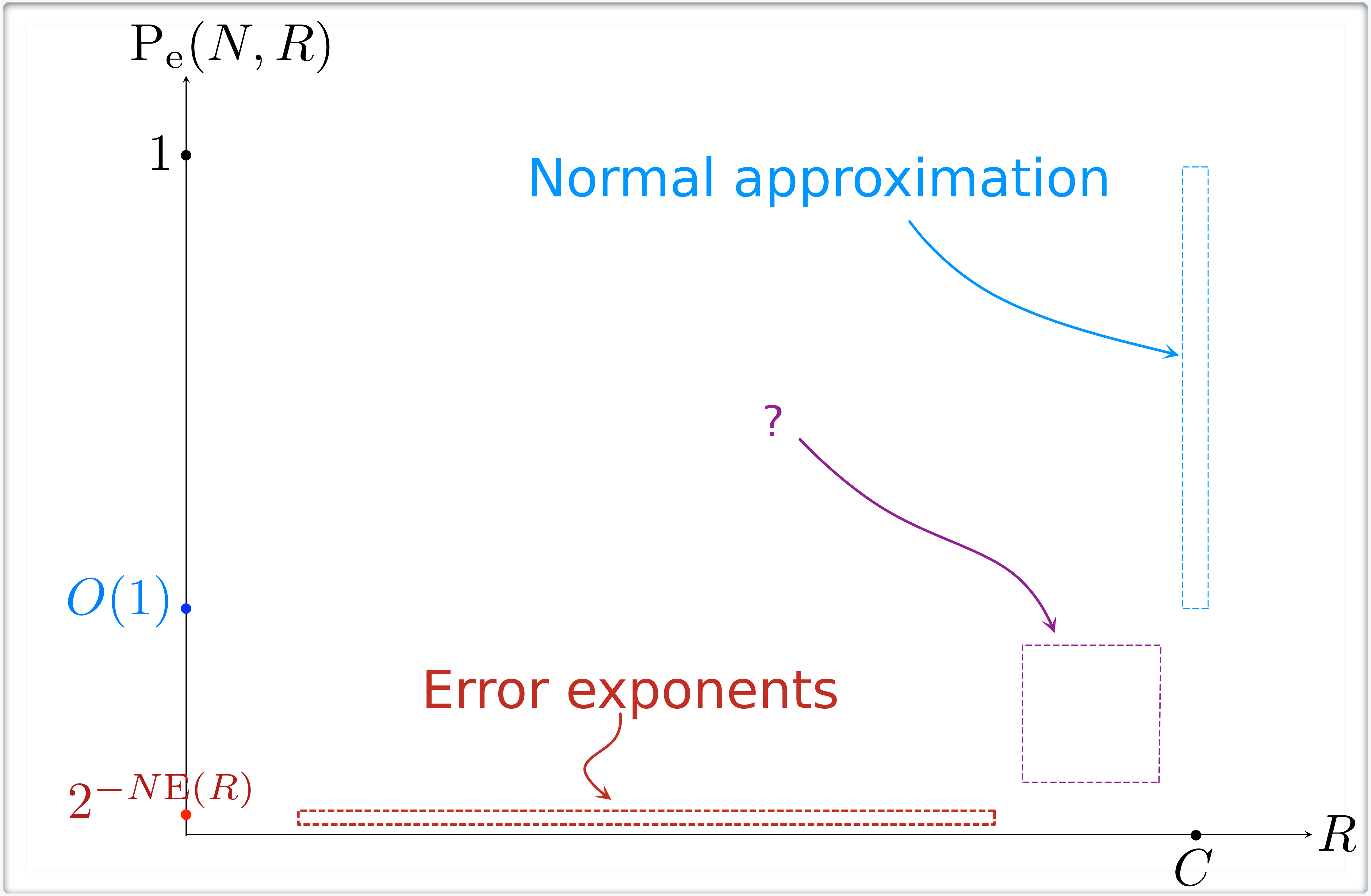
- **Theorem:** [Strassen '62, Hayashi '09, Polyanskiy, Poor, and Verdú '10]
As $N \to \infty$

$$R^*(N,\epsilon) = C - \sqrt{\frac{\boxed{V}}{N}}\,Q^{-1}(\epsilon) + O\left(\frac{\log N}{N}\right)$$

Channel dispersion

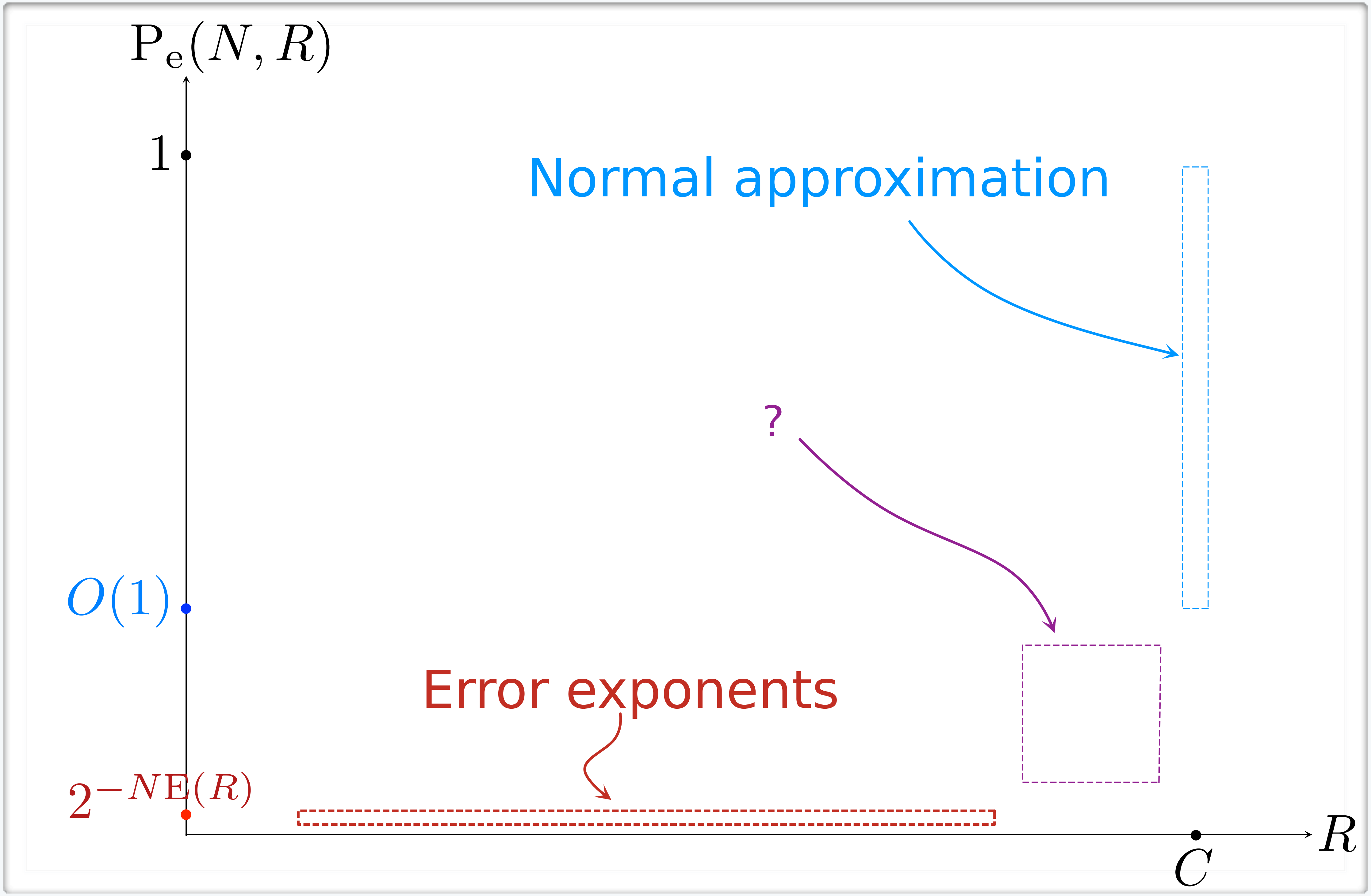
Analogous to CLT. Is it a better approach?

Fix some large *N*.



Moderate Deviations

Fix some large *N*.



Moderate Deviations

Moderate Deviations

- Error exponents:

$$\lim_{N\to\infty} -\frac{\log \mathrm{P_e}(R,N)}{N} = \mathrm{E}(R)$$

Moderate Deviations

- Error exponents:

$$\lim_{N\to\infty} -\frac{\log \mathrm{P_e}(R,N)}{N} = \mathrm{E}(R)$$

- Normal approximation:

$$\mathrm{P_e}(R,N) \approx \epsilon$$

Moderate Deviations

- Error exponents:

$$\lim_{N\to\infty} -\frac{\log \mathrm{P_e}(R,N)}{N} = \mathrm{E}(R)$$

- Normal approximation:

$$\mathrm{P_e}(R,N) \approx \epsilon$$

- What is in between?

Moderate Deviations

Theorem: [Altuğ-Wagner ’10]

- Let $R_N = C$ - ϵ_N and assume

$$\lim_{N\to\infty} \epsilon_N = 0 \qquad \lim_{N\to\infty} \epsilon_N \sqrt{N} = \infty$$

- Then

$$\lim_{N\to\infty} -\frac{\log \mathrm{P_e}(N, R_N)}{\epsilon_N^2 N} = \frac{1}{2V}$$

$$\left(\mathrm{P_e}(N, R_N) \approx 2^{-\epsilon_N^2 \cdot N/(2V)}\right)$$

Proof Comments

First idea: Take existing finite-N bounds, and simply apply the new asymptotic

Proof Comments

First idea: Take existing finite-N bounds, and simply apply the new asymptotic

- Achievability ✓

Proof Comments

First idea: Take existing finite-N bounds, and simply apply the new asymptotic

- Achievability ✓
- Converse ✗

Proof Comments

First idea: Take existing finite-N bounds, and simply apply the new asymptotic

- Achievability ✓
- Converse ✗

$$\mathrm{P_e}(N, R) \geq \frac{K(R)}{N^{0.5(1+|\mathrm{E}'(R)|)}} 2^{-N\mathrm{E}(R)}$$

Proof Comments

First idea: Take existing finite-N bounds, and simply apply the new asymptotic

- Achievability ✓
- Converse ✗

$$\mathrm{P_e}(N,R) \geq \frac{K(R)}{N^{0.5(1+|\mathrm{E}'(R)|)}} 2^{-N\mathrm{E}(R)}$$

appears to blow up as R approaches C

Proof Comments

First idea: Take existing finite-N bounds, and simply apply the new asymptotic

- Achievability ✓
- Converse ✗

$$P_e(N,R) \geq \frac{\boxed{K(R)}}{N^{0.5(1+|E'(R)|)}} 2^{-NE(R)}$$

appears to blow up as R approaches C

Characterizing the lower-order factors would unify error exponents and moderate deviations

Analogies

I.I.D. Sums		**Channel Coding**
1. WLLN	$\longleftrightarrow$	1. Channel Coding Theorem
2. LDP	$\longleftrightarrow$	2. Error Exponents
3. CLT	$\longleftrightarrow$	3. Normal Approximation
4. EA	$\longleftrightarrow$	4. EA in channel coding
5. MDP	$\longleftrightarrow$	5. MDP in channel coding

Analogies

I.I.D. Sums		**Channel Coding**
1. WLLN	$\longleftrightarrow$	1. Channel Coding Theorem
2. LDP	$\longleftrightarrow$	2. Error Exponents
3. CLT	$\longleftrightarrow$	3. Normal Approximation
4. EA	$\longleftrightarrow$	4. EA in channel coding
5. MDP	$\longleftrightarrow$	5. MDP in channel coding

Too many results!

Can we unify them?

Conclusion

Pre-factor Bounds	Symmetric	Asymmetric
Regular	$O(N^{0.5(1+\lvert E'(R)\rvert)})$ $\Omega(N^{0.5(1+\lvert E'(R)\rvert)})$	$O(N^{0.5(1+\lvert E'(R-)\rvert)})$ $\Omega(N^{0.5(1+\lvert E'(R-)\rvert)})$
Irregular	$O(N^{0.5})$ $\Omega(N^{0.5})$	$O(N^{0.5})$ $\Omega(N^{0.5(1+\lvert E'(R-)\rvert)})$

Conclusion

Pre-factor Bounds	Symmetric	Asymmetric
Regular	$O(N^{0.5(1+\lvert E'(R)\rvert)})$ $\Omega(N^{0.5(1+\lvert E'(R)\rvert)})$	$O(N^{0.5(1+\lvert E'(R-)\rvert)})$ $\Omega(N^{0.5(1+\lvert E'(R-)\rvert)})$
Irregular	$O(N^{0.5})$ $\Omega(N^{0.5})$	$O(N^{0.5})$ $\Omega(N^{0.5(1+\lvert E'(R-)\rvert)})$

- Normal approximation in channel coding

Conclusion

Pre-factor Bounds	Symmetric	Asymmetric
Regular	$O(N^{0.5(1+\lvert E'(R)\rvert)})$ $\Omega(N^{0.5(1+\lvert E'(R)\rvert)})$	$O(N^{0.5(1+\lvert E'(R-)\rvert)})$ $\Omega(N^{0.5(1+\lvert E'(R-)\rvert)})$
Irregular	$O(N^{0.5})$ $\Omega(N^{0.5})$	$O(N^{0.5})$ $\Omega(N^{0.5(1+\lvert E'(R-)\rvert)})$

- Normal approximation in channel coding
- Moderate deviations in channel coding

Conclusion

Pre-factor Bounds	Symmetric	Asymmetric
Regular	$O(N^{0.5(1+\lvert E'(R)\rvert)})$ $\Omega(N^{0.5(1+\lvert E'(R)\rvert)})$	$O(N^{0.5(1+\lvert E'(R-)\rvert)})$ $\Omega(N^{0.5(1+\lvert E'(R-)\rvert)})$
Irregular	$O(N^{0.5})$ $\Omega(N^{0.5})$	$O(N^{0.5})$ $\Omega(N^{0.5(1+\lvert E'(R-)\rvert)})$

- Normal approximation in channel coding
- Moderate deviations in channel coding
- Lower-order results could unify regimes

I.I.D. Sums

I.I.D. Sums

- Exact Asymptotics [Bahadur-Rao ’60]:

$$\lim_{N \to 0} \frac{\Pr\left(\sum_{i=1}^{N} X_i > N\epsilon\right)}{\frac{c}{\sqrt{N}} 2^{-N\Lambda^*(\epsilon)}} = 1$$

I.I.D. Sums

- Exact Asymptotics [Bahadur-Rao ’60]:

$$\lim_{N \to 0} \frac{\Pr\left(\sum_{i=1}^{N} X_i > N\epsilon\right)}{\frac{c}{\sqrt{N}} 2^{-N\Lambda^*(\epsilon)}} = 1$$

- Moderate Deviations: if β is in (1/2, 1):

$$\lim_{N \to 0} \frac{1}{N^{2\beta-1}} \log \Pr\left(\sum_{i=1}^{N} X_i > \epsilon N^{\beta}\right) = \Lambda^*_{\mathcal{N}}(\epsilon)$$