# **Entropy and Shannon's Theorem**

Lecture 24 November 18, 2014

## Part I

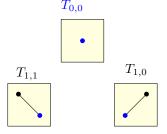
# Entropy

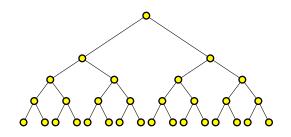
#### Part II

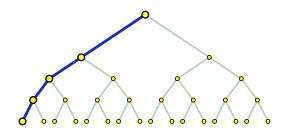
## Extracting randomness

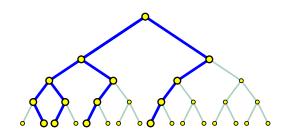
#### Storing all strings of length n and j bits on

- **1**  $S_{n,j}$ : set of all strings of length n with j ones in them.
- $oldsymbol{O}$   $T_{n,j}$ : prefix tree storing all  $S_{n,j}$ .

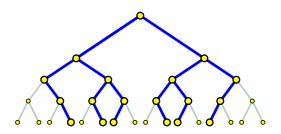








```
 S_{4,1} = \{0001, 0010, 0100, 1000\} 
 \Longrightarrow \#(0001) = 0. 
 \#(0010) = 1. 
 \#(0100) = 2. 
 \#(1000) = 3.
```



$$S_{4,2} = \{0011, 0101, 0110, 1001, 1010, 1100\}$$

$$\Rightarrow$$

$$\#(0011) = 0.$$

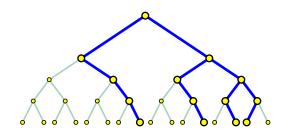
$$\#(0101) = 1.$$

$$\#(0110) = 2.$$

$$\#(1001) = 3.$$

$$\#(1010) = 4.$$

$$\#(1100) = 5.$$



$$S_{4,3} = \{0111, 1011, 1101, 1110\}$$

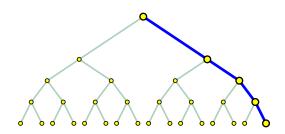
$$\Longrightarrow$$

$$\#(0111) = 0.$$

$$\#(1011) = 1.$$

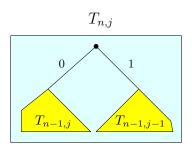
$$\#(1101) = 2.$$

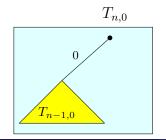
$$\#(1110) = 3.$$

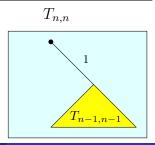


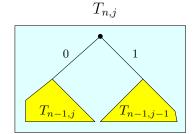
$$\begin{array}{ccc} \bullet & S_{4,4} = \{1111\} \\ & \Longrightarrow \\ & \#(1111) = 0. \end{array}$$

2

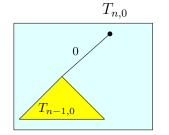


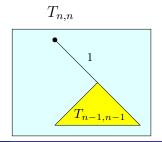


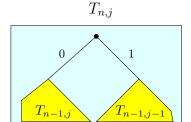




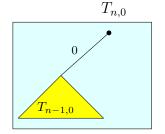
$$\#$$
 of leafs:  $|T_{n,j}|=|T_{n-1,j}|+|T_{n-1,j-1}|$ 

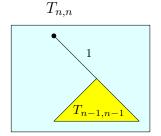


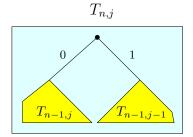




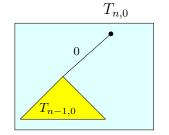
$$\#$$
 of leafs:  $|T_{n,j}|=|T_{n-1,j}|+|T_{n-1,j-1}|$   ${n\choose j}={n-1\choose j}+{n-1\choose j-1}$ 

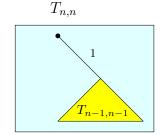






$$\#$$
 of leafs:  $|T_{n,j}| = |T_{n-1,j}| + |T_{n-1,j-1}|$   $\binom{n}{j} = \binom{n-1}{j} + \binom{n-1}{j-1}$   $\Longrightarrow |T_{n,j}| = \binom{n}{j}.$ 

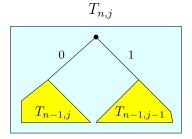




- **1**  $T_{n,j}$  leafs corresponds to strings of  $S_{n,j}$ .
- ② Order all strings of  $S_{n,j}$  order in lexicographical ordering
- $\odot$   $\equiv$  ordering leafs of  $T_{n,j}$  from left to right.
- EncodeBinomCoeff(s) denote this polytime procedure.

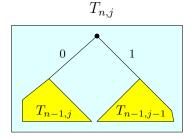
- **1**  $T_{n,j}$  leafs corresponds to strings of  $S_{n,j}$ .
- ② Order all strings of  $S_{n,j}$  order in lexicographical ordering
- $\odot$  = ordering leafs of  $T_{n,j}$  from left to right.
- lacksquare Input:  $s \in S_{n,j}$ : compute index of s in sorted set  $S_{n,j}$ .
- ullet EncodeBinomCoeff(s) denote this polytime procedure.

- **1**  $T_{n,j}$  leafs corresponds to strings of  $S_{n,j}$ .
- ② Order all strings of  $S_{n,j}$  order in lexicographical ordering
- ullet ordering leafs of  $T_{n,j}$  from left to right.



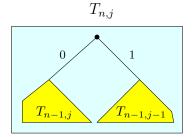
- ullet Input:  $s \in S_{n,j}$ : compute index of s in sorted set  $S_{n,j}$ .
- EncodeBinomCoeff(s) denote this polytime procedure.

- **1**  $T_{n,j}$  leafs corresponds to strings of  $S_{n,j}$ .
- ② Order all strings of  $S_{n,j}$  order in lexicographical ordering
- ullet ordering leafs of  $T_{n,j}$  from left to right.



- **1** Input:  $s \in S_{n,j}$ : compute index of s in sorted set  $S_{n,j}$ .
- EncodeBinomCoeff(s) denote this polytime procedure.

- **1**  $T_{n,j}$  leafs corresponds to strings of  $S_{n,j}$ .
- ② Order all strings of  $S_{n,j}$  order in lexicographical ordering
- ullet ordering leafs of  $T_{n,j}$  from left to right.

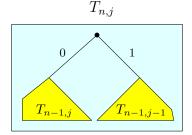


- **1** Input:  $s \in S_{n,j}$ : compute index of s in sorted set  $S_{n,j}$ .
- **Solution EncodeBinomCoeff**(s) denote this polytime procedure.

- **1**  $T_{n,j}$  leafs corresponds to strings of  $S_{n,j}$ .
- $oldsymbol{\circ}$  Order all strings of  $S_{n,j}$  order in lexicographical ordering
- $\odot$   $\equiv$  ordering leafs of  $T_{n,j}$  from left to right.
- $lack x \in \{1,\ldots, {n \choose j}\}$ : compute xth string in  $S_{n,j}$  in polytime.
- **5 DecodeBinomCoeff** (x) denote this procedure.

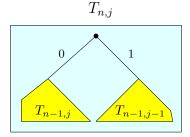
- **1**  $T_{n,j}$  leafs corresponds to strings of  $S_{n,j}$ .
- $oldsymbol{\circ}$  Order all strings of  $S_{n,j}$  order in lexicographical ordering
- ullet ordering leafs of  $T_{n,j}$  from left to right.
- $x \in \{1, \ldots, \binom{n}{i}\}$ : compute xth string in  $S_{n,j}$  in polytime.
- **5 DecodeBinomCoeff** (x) denote this procedure.

- **1**  $T_{n,j}$  leafs corresponds to strings of  $S_{n,j}$ .
- ② Order all strings of  $S_{n,j}$  order in lexicographical ordering
- lacktriangledown lacktriangledown ordering leafs of  $T_{n,j}$  from left to right.



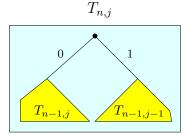
- $x \in \{1, \ldots, \binom{n}{i}\}$ : compute xth string in  $S_{n,j}$  in polytime.
- **5 DecodeBinomCoeff** (x) denote this procedure.

- **1**  $T_{n,j}$  leafs corresponds to strings of  $S_{n,j}$ .
- ② Order all strings of  $S_{n,j}$  order in lexicographical ordering
- ullet ordering leafs of  $T_{n,j}$  from left to right.



- $x \in \{1, \ldots, \binom{n}{i}\}$ : compute xth string in  $S_{n,j}$  in polytime.
- DecodeBinomCoeff (x) denote this procedure.

- **1**  $T_{n,j}$  leafs corresponds to strings of  $S_{n,j}$ .
- ② Order all strings of  $S_{n,j}$  order in lexicographical ordering
- ullet ordering leafs of  $T_{n,j}$  from left to right.



- $x \in \{1, \ldots, \binom{n}{i}\}$ : compute xth string in  $S_{n,j}$  in polytime.
- DecodeBinomCoeff (x) denote this procedure.

## Encoding/decoding strings of $S_{n,j}$

#### Lemma

 $S_{n,j}$ : Set of binary strings of length n with j ones, sorted lexicographically.

- **1** EncodeBinomCoeff( $\alpha$ ): Input is string  $\alpha \in S_{n,j}$ , compute index x of  $\alpha$  in  $S_{n,j}$  in polynomial time in n.
- ② DecodeBinomCoeff(x): Input index  $x \in \{1, \ldots, \binom{n}{j}\}$ . Output xth string  $\alpha$  in  $S_{n,j}$ , in time  $O(\operatorname{polylog} n + n)$ .

#### Extracting randomness

#### **Theorem**

Consider a coin that comes up heads with probability p>1/2. For any constant  $\delta>0$  and for n sufficiently large:

- (A) One can extract, from an input of a sequence of n flips, an output sequence of  $(1-\delta)n\mathbb{H}(p)$  (unbiased) independent random bits.
- (B) One can not extract more than  $n\mathbb{H}(p)$  bits from such a sequence.

- There are  $\binom{n}{j}$  input strings with exactly j heads.
- $ext{@}$  each has probability  $p^j(1-p)^{n-j}$ .
- $oldsymbol{\circ}$  map string s like that to index number in the set  $S_j = \left\{1,\ldots, {n \choose j}
  ight\}.$
- ① Given that input string s has j ones (out of n bits) defines a uniform distribution on  $S_{n,j}$ .
- $lacksquare{0}$  x uniform distributed in  $\{1,\ldots,N\}$ ,  $N=inom{n}{j}$ .
- Seen in previous lecture...
- ... extract in expectation,  $\lfloor \lg N \rfloor 1$  bits from uniform random variable in the range  $1, \ldots, N$ .
- Extract bits using ExtractRandomness(x, N):.

- There are  $\binom{n}{j}$  input strings with exactly j heads.
- $oldsymbol{2}$  each has probability  $p^j(1-p)^{n-j}$ .
- $oldsymbol{\circ}$  map string s like that to index number in the set  $S_j = ig\{1,\ldots,inom{n}{j}ig\}.$
- ① Given that input string s has j ones (out of n bits) defines a uniform distribution on  $S_{n,j}$ .
- $lacksquare{0}$  x uniform distributed in  $\{1,\ldots,N\}$ ,  $N=inom{n}{j}$ .
- Seen in previous lecture...
- ... extract in expectation,  $\lfloor \lg N \rfloor 1$  bits from uniform random variable in the range  $1, \ldots, N$ .
- ① Extract bits using **ExtractRandomness**(x, N):.

- There are  $\binom{n}{j}$  input strings with exactly j heads.
- $oldsymbol{2}$  each has probability  $p^j(1-p)^{n-j}$ .
- $oldsymbol{\circ}$  map string s like that to index number in the set  $S_j = \left\{1,\ldots, {n \choose j}
  ight\}.$
- ① Given that input string s has j ones (out of n bits) defines a uniform distribution on  $S_{n,j}$ .
- $lacksquare{0}$  x uniform distributed in  $\{1,\ldots,N\}$ ,  $N=inom{n}{j}$ .
- Seen in previous lecture...
- @ ... extract in expectation,  $\lfloor \lg N \rfloor 1$  bits from uniform random variable in the range  $1, \ldots, N$ .
- ① Extract bits using **ExtractRandomness**(x, N):.

- There are  $\binom{n}{j}$  input strings with exactly j heads.
- $oldsymbol{2}$  each has probability  $p^j(1-p)^{n-j}$ .
- lacktriangledown map string s like that to index number in the set  $S_j = \left\{1,\ldots, {n \choose j}
  ight\}.$
- **4** Given that input string s has j ones (out of n bits) defines a uniform distribution on  $S_{n,j}$ .
- $lacksquare{0}$  x uniform distributed in  $\{1,\ldots,N\}$ ,  $N=inom{n}{j}$ .
- Seen in previous lecture...
- ... extract in expectation,  $\lfloor \lg N \rfloor 1$  bits from uniform random variable in the range  $1, \ldots, N$ .
- ① Extract bits using **ExtractRandomness**(x, N):.

- There are  $\binom{n}{j}$  input strings with exactly j heads.
- $oldsymbol{2}$  each has probability  $p^j(1-p)^{n-j}$ .
- lacktriangledown map string s like that to index number in the set  $S_j = \left\{1,\ldots, {n \choose j}
  ight\}.$
- **4** Given that input string s has j ones (out of n bits) defines a uniform distribution on  $S_{n,j}$ .
- **5**  $x \leftarrow \mathsf{EncodeBinomCoeff}(s)$
- $lacksquare{0}$  x uniform distributed in  $\{1,\ldots,N\}$ ,  $N=inom{n}{j}$ .
- Seen in previous lecture...
- ... extract in expectation,  $\lfloor \lg N \rfloor 1$  bits from uniform random variable in the range  $1, \ldots, N$ .
- ① Extract bits using **ExtractRandomness**(x, N):.

- There are  $\binom{n}{j}$  input strings with exactly j heads.
- $oldsymbol{2}$  each has probability  $p^j(1-p)^{n-j}$ .
- lacktriangledown map string s like that to index number in the set  $S_j = \left\{1,\ldots, {n \choose j}
  ight\}.$
- **4** Given that input string s has j ones (out of n bits) defines a uniform distribution on  $S_{n,j}$ .
- **5**  $x \leftarrow \mathsf{EncodeBinomCoeff}(s)$
- **6** x uniform distributed in  $\{1,\ldots,N\}$ ,  $N={n\choose j}$ .
- Seen in previous lecture...
- ... extract in expectation,  $\lfloor \lg N \rfloor 1$  bits from uniform random variable in the range  $1, \ldots, N$ .
- ① Extract bits using **ExtractRandomness**(x, N):.

- There are  $\binom{n}{j}$  input strings with exactly j heads.
- $oldsymbol{2}$  each has probability  $p^j(1-p)^{n-j}$ .
- $oldsymbol{\circ}$  map string s like that to index number in the set  $S_j = \left\{1,\ldots, {n \choose j}
  ight\}.$
- **3** Given that input string s has j ones (out of n bits) defines a uniform distribution on  $S_{n,j}$ .
- **5**  $x \leftarrow \mathsf{EncodeBinomCoeff}(s)$
- $oldsymbol{0}$  x uniform distributed in  $\{1,\ldots,N\}$ ,  $N=inom{n}{j}$ .
- Seen in previous lecture...
- ... extract in expectation,  $\lfloor \lg N \rfloor 1$  bits from uniform random variable in the range  $1, \ldots, N$ .
- ① Extract bits using **ExtractRandomness**(x, N):.

- There are  $\binom{n}{j}$  input strings with exactly j heads.
- 2 each has probability  $p^j(1-p)^{n-j}$ .
- lacktriangledown map string s like that to index number in the set  $S_j = \left\{1,\ldots, {n \choose j}
  ight\}.$
- **4** Given that input string s has j ones (out of n bits) defines a uniform distribution on  $S_{n,j}$ .
- **5**  $x \leftarrow \mathsf{EncodeBinomCoeff}(s)$
- $oldsymbol{0}$  x uniform distributed in  $\{1,\ldots,N\}$ ,  $N=inom{n}{j}$ .
- Seen in previous lecture...
- ... extract in expectation,  $\lfloor \lg N \rfloor 1$  bits from uniform random variable in the range  $1, \ldots, N$ .
- **①** Extract bits using **ExtractRandomness**(x, N):.

- There are  $\binom{n}{j}$  input strings with exactly j heads.
- $oldsymbol{2}$  each has probability  $p^j(1-p)^{n-j}$ .
- lacktriangledown map string s like that to index number in the set  $S_j = \left\{1,\ldots, {n \choose j}
  ight\}.$
- **3** Given that input string s has j ones (out of n bits) defines a uniform distribution on  $S_{n,j}$ .
- **5**  $x \leftarrow \mathsf{EncodeBinomCoeff}(s)$
- $oldsymbol{0}$  x uniform distributed in  $\{1,\ldots,N\}$ ,  $N=inom{n}{j}$ .
- Seen in previous lecture...
- ... extract in expectation,  $\lfloor \lg N \rfloor 1$  bits from uniform random variable in the range  $1, \ldots, N$ .
- **1** Extract bits using **ExtractRandomness**(x, N):.

#### Exciting proof continued...

- ① Z: random variable: number of heads in input string s.
- ② B: number of random bits extracted.

$$\operatorname{E}igl[Bigr] = \sum_{k=0}^{n} \operatorname{Pr}[Z=k] \operatorname{E}igl[B \ igr| Z=kigr],$$

- $lacksquare{1}{3}$  Know:  $\mathrm{E}ig[B \ ig| Z = kig] \geq \left\lfloor \lg ig( n top k ig) 
  ight
  floor 1.$

$$egin{pmatrix} n \ k \end{pmatrix} \geq egin{pmatrix} n \ \lfloor n(p+arepsilon) 
floor \geq rac{2^{n\mathbb{H}(p+arepsilon)}}{n+1},$$

... since  $2^{n\mathbb{H}(p)}$  is a good approximation to  $\binom{n}{np}$  as proved in previous lecture.

- lacksquare Z: random variable: number of heads in input string s.
- 2 B: number of random bits extracted

$$\operatorname{E}igl[Bigr] = \sum_{k=0}^n \operatorname{Pr}[Z=k] \operatorname{E}igl[B \ igr| Z=kigr],$$

- $lacksquare{1}{3}$  Know:  $\mathrm{E}ig[B \ \Big| \ Z = kig] \geq \left\lfloor \lg igg( n \ k ig) 
  ight
  floor 1.$

$$egin{pmatrix} n \ k \end{pmatrix} \geq egin{pmatrix} n \ \lfloor n(p+arepsilon) 
floor \geq rac{2^{n\mathbb{H}(p+arepsilon)}}{n+1},$$

... since  $2^{n\mathbb{H}(p)}$  is a good approximation to  $\binom{n}{np}$  as proved in previous lecture.

- lacksquare Z: random variable: number of heads in input string s.
- 2 B: number of random bits extracted.

$$\mathrm{E}ig[Big] = \sum_{k=0}^n \Pr[Z=k] \, \mathrm{E}ig[B \ ig| \, Z=kig] \, ,$$

- $lacksquare{1}{3}$  Know:  $\mathrm{E}ig[B \ \Big| \ Z = k ig] \geq \left\lfloor \lg ig( n top k ig) 
  ight
  floor 1.$

$$egin{pmatrix} n \ k \end{pmatrix} \geq egin{pmatrix} n \ \lfloor n(p+arepsilon) 
floor \geq rac{2^{n\mathbb{H}(p+arepsilon)}}{n+1},$$

① ... since  $2^{n\mathbb{H}(p)}$  is a good approximation to  $\binom{n}{np}$  as proved in previous lecture.

- **1** Z: random variable: number of heads in input string s.
- 2 B: number of random bits extracted.

$$\mathrm{E}ig[Big] = \sum_{k=0}^n \mathrm{Pr}[Z=k] \, \mathrm{E}ig[B \ ig| \, Z=kig] \, ,$$

- $lacksquare{1}{3}$  Know:  $\mathrm{E}ig[B \ \Big| \ Z = kig] \geq \left\lfloor \lg ig( egin{matrix} n \ k \end{matrix} 
  ight) 
  ight
  floor 1.$

$$egin{pmatrix} n \ k \end{pmatrix} \geq egin{pmatrix} n \ \lfloor n(p+arepsilon) 
floor \geq rac{2^{n\mathbb{H}(p+arepsilon)}}{n+1},$$

... since  $2^{n\mathbb{H}(p)}$  is a good approximation to  $\binom{n}{np}$  as proved in previous lecture.

- **1** Z: random variable: number of heads in input string s.
- 2 B: number of random bits extracted.

$$\mathrm{E}igl[Bigr] = \sum_{k=0}^n \mathrm{Pr}[Z=k] \, \mathrm{E}igl[B \, igr| \, Z=kigr] \, ,$$

- $lacksquare{1}{3}$  Know:  $\mathrm{E}ig[B \ \Big| \ Z = k ig] \geq \left[ \mathrm{lg} ig( egin{matrix} n \ k \end{matrix} ig) ig] 1.$

$$egin{pmatrix} n \ k \end{pmatrix} \geq egin{pmatrix} n \ \lfloor n(p+arepsilon) 
floor \geq rac{2^{n\mathbb{H}(p+arepsilon)}}{n+1},$$

... since  $2^{n\mathbb{H}(p)}$  is a good approximation to  $\binom{n}{np}$  as proved in previous lecture.

- **1** Z: random variable: number of heads in input string s.
- 2 B: number of random bits extracted.

$$\mathrm{E}igl[Bigr] = \sum_{k=0}^n \mathrm{Pr}[Z=k]\,\mathrm{E}igl[B\,\,igr|\, Z=kigr]\,,$$

- $lacksquare{1}{3}$  Know:  $\mathrm{E}ig[B \ \Big| \ Z = kig] \geq \left\lfloor \lg ig( n top k ig) 
  ight
  floor 1.$
- $\bullet$   $\varepsilon < p-1/2$ : sufficiently small constant.

$$egin{pmatrix} n \ k \end{pmatrix} \geq egin{pmatrix} n \ \lfloor n(p+arepsilon) 
floor \geq rac{2^{n\mathbb{H}(p+arepsilon)}}{n+1}, \end{cases}$$

① ... since  $2^{n\mathbb{H}(p)}$  is a good approximation to  $\binom{n}{np}$  as proved in previous lecture.

- **1** Z: random variable: number of heads in input string s.
- 2 B: number of random bits extracted.

$$\mathrm{E}igl[Bigr] = \sum_{k=0}^n \mathrm{Pr}[Z=k]\,\mathrm{E}igl[B\,\,igr|\, Z=kigr]\,,$$

- $lacksquare{1}{3}$  Know:  $\mathrm{E}ig[B \ \Big| \ Z = k ig] \geq \left\lfloor \lg ig( n top k ig) 
  ight
  floor 1.$
- $\bullet$   $\varepsilon < p-1/2$ : sufficiently small constant.

$$egin{pmatrix} n \ k \end{pmatrix} \geq egin{pmatrix} n \ \lfloor n(p+arepsilon) 
floor \geq rac{2^{n\mathbb{H}(p+arepsilon)}}{n+1}, \end{cases}$$

**1** ... since  $2^{n\mathbb{H}(p)}$  is a good approximation to  $\binom{n}{np}$  as proved in previous lecture.

$$\begin{split} \mathbf{E} \Big[ B \Big] &= \sum_{k=0}^{n} \Pr[Z=k] \, \mathbf{E} \Big[ B \, \Big| \, Z=k \Big] \, . \\ \mathbf{E} \Big[ B \Big] &\geq \sum_{k=\lfloor n(p-\varepsilon) \rfloor}^{\lceil n(p+\varepsilon) \rceil} \Pr \Big[ Z=k \Big] \, \mathbf{E} \Big[ B \, \Big| \, Z=k \Big] \\ &\geq \sum_{k=\lfloor n(p-\varepsilon) \rfloor}^{\lceil n(p+\varepsilon) \rceil} \Pr \Big[ Z=k \Big] \left( \left\lfloor \lg \left( \frac{n}{k} \right) \right\rfloor -1 \right) \\ &\geq \sum_{k=\lfloor n(p-\varepsilon) \rfloor}^{\lceil n(p+\varepsilon) \rceil} \Pr \Big[ Z=k \Big] \left( \lg \frac{2^{n\mathbb{H}(p+\varepsilon)}}{n+1} -2 \right) \\ &= \left( n\mathbb{H}(p+\varepsilon) - \lg(n+1) -2 \right) \Pr[|Z-np| \leq \varepsilon \, n] \\ &\geq \left( n\mathbb{H}(p+\varepsilon) - \lg(n+1) -2 \right) \left( 1 - 2 \exp \left( -\frac{n\varepsilon^2}{4p} \right) \right), \\ \text{since } \mu &= \mathbf{E}[Z] = np \text{ and} \\ &\Pr \Big[ |Z-np| \geq \frac{\varepsilon}{p} pn \Big] \leq 2 \exp \left( -\frac{np}{4} \left( \frac{\varepsilon}{p} \right)^2 \right) = 2 \exp \left( -\frac{n\varepsilon^2}{4p} \right), \text{ by the Chernoff inequality.} \end{split}$$

$$\begin{split} \mathbf{E} \Big[ \boldsymbol{B} \Big] &= \sum_{k=0}^n \Pr[\boldsymbol{Z} = k] \, \mathbf{E} \Big[ \boldsymbol{B} \, \Big| \, \boldsymbol{Z} = k \Big] \, . \\ \mathbf{E} \Big[ \boldsymbol{B} \Big] &\geq \sum_{k=\lfloor n(p-\varepsilon) \rfloor}^{\lceil n(p+\varepsilon) \rceil} \Pr[\boldsymbol{Z} = k] \, \mathbf{E} \Big[ \boldsymbol{B} \, \Big| \, \boldsymbol{Z} = k \Big] \\ &\geq \sum_{k=\lfloor n(p-\varepsilon) \rfloor}^{\lceil n(p+\varepsilon) \rceil} \Pr[\boldsymbol{Z} = k] \, \bigg( \left| \lg \left( \frac{n}{k} \right) \right| - 1 \bigg) \\ &\geq \sum_{k=\lfloor n(p-\varepsilon) \rfloor}^{\lceil n(p+\varepsilon) \rceil} \Pr[\boldsymbol{Z} = k] \, \bigg( \lg \frac{2^{n\mathbb{H}(p+\varepsilon)}}{n+1} - 2 \bigg) \\ &= \Big( n\mathbb{H}(p+\varepsilon) - \lg(n+1) - 2 \Big) \Pr[|\boldsymbol{Z} - np| \leq \varepsilon n] \\ &\geq \Big( n\mathbb{H}(p+\varepsilon) - \lg(n+1) - 2 \Big) \bigg( 1 - 2 \exp \bigg( - \frac{n\varepsilon^2}{4p} \bigg) \bigg), \\ \text{since } \boldsymbol{\mu} = \mathbf{E}[\boldsymbol{Z}] = np \text{ and} \\ &\Pr[|\boldsymbol{Z} - np| \geq \frac{\varepsilon}{p} pn] \leq 2 \exp \bigg( - \frac{np}{4} \Big( \frac{\varepsilon}{p} \Big)^2 \bigg) = 2 \exp \Big( - \frac{n\varepsilon^2}{4p} \Big), \text{ by the Chernoff inequality.} \end{split}$$

$$\begin{split} &\mathbf{E} \Big[ B \Big] = \sum_{k=0}^{n} \Pr[Z=k] \, \mathbf{E} \Big[ B \, \Big| \, Z=k \Big]. \\ &\mathbf{E} \Big[ B \Big] \geq \sum_{k=\lfloor n(p-\varepsilon) \rfloor}^{\lceil n(p+\varepsilon) \rceil} \Pr \Big[ Z=k \Big] \, \mathbf{E} \Big[ B \, \Big| \, Z=k \Big] \\ &\geq \sum_{k=\lfloor n(p-\varepsilon) \rfloor}^{\lceil n(p+\varepsilon) \rceil} \Pr \Big[ Z=k \Big] \, \left( \left\lfloor \lg \binom{n}{k} \right\rfloor -1 \right) \\ &\geq \sum_{k=\lfloor n(p-\varepsilon) \rfloor}^{\lceil n(p+\varepsilon) \rceil} \Pr \Big[ Z=k \Big] \left( \lg \frac{2^{n\mathbb{H}(p+\varepsilon)}}{n+1} -2 \right) \\ &= \left( n\mathbb{H}(p+\varepsilon) - \lg(n+1) - 2 \right) \Pr[|Z-np| \leq \varepsilon n] \\ &\geq \left( n\mathbb{H}(p+\varepsilon) - \lg(n+1) - 2 \right) \left( 1 - 2 \exp \left( -\frac{n\varepsilon^2}{4p} \right) \right), \\ \text{since } \mu = \mathbf{E}[Z] = np \text{ and} \\ &\mathbf{Pr} \Big[ |Z-np| \geq \frac{\varepsilon}{p} pn \Big] \leq 2 \exp \left( -\frac{np}{4} \left( \frac{\varepsilon}{p} \right)^2 \right) = 2 \exp \left( -\frac{n\varepsilon^2}{4p} \right), \text{ by the Chernoff inequality.} \end{split}$$

$$\begin{split} \mathbf{E} \Big[ B \Big] &= \sum_{k=0}^n \Pr[Z=k] \, \mathbf{E} \Big[ B \, \Big| \, Z=k \Big]. \\ \mathbf{E} \Big[ B \Big] &\geq \sum_{k=\lfloor n(p-\varepsilon)\rfloor}^{\lceil n(p+\varepsilon)\rceil} \Pr \Big[ Z=k \Big] \, \mathbf{E} \Big[ B \, \Big| \, Z=k \Big] \\ &\geq \sum_{k=\lfloor n(p-\varepsilon)\rfloor}^{\lceil n(p+\varepsilon)\rceil} \Pr \Big[ Z=k \Big] \left( \left\lfloor \lg \binom{n}{k} \right\rfloor -1 \right) \\ &\geq \sum_{k=\lfloor n(p-\varepsilon)\rfloor}^{\lceil n(p+\varepsilon)\rceil} \Pr \Big[ Z=k \Big] \left( \lg \frac{2^{n\mathbb{H}(p+\varepsilon)}}{n+1} -2 \right) \\ &= \left( n\mathbb{H}(p+\varepsilon) - \lg(n+1) - 2 \right) \Pr[|Z-np| \leq \varepsilon n] \\ &\geq \left( n\mathbb{H}(p+\varepsilon) - \lg(n+1) - 2 \right) \left( 1 - 2 \exp \left( -\frac{n\varepsilon^2}{4p} \right) \right), \\ \text{since } \mu &= \mathbf{E}[Z] = np \text{ and} \\ &\Pr[|Z-np| \geq \frac{\varepsilon}{p} pn] \leq 2 \exp \left( -\frac{np}{4} \left( \frac{\varepsilon}{p} \right)^2 \right) = 2 \exp \left( -\frac{n\varepsilon^2}{4p} \right), \text{ by the Chernoff inequality.} \end{split}$$

$$\begin{split} &\mathbf{E} \Big[ B \Big] = \sum_{k=0}^{n} \Pr[Z=k] \, \mathbf{E} \Big[ B \, \Big| \, Z=k \Big]. \\ &\mathbf{E} \Big[ B \Big] \geq \sum_{k=\lfloor n(p-\varepsilon) \rfloor}^{\lceil n(p+\varepsilon) \rceil} \Pr \Big[ Z=k \Big] \, \mathbf{E} \Big[ B \, \Big| \, Z=k \Big] \\ &\geq \sum_{k=\lfloor n(p-\varepsilon) \rfloor}^{\lceil n(p+\varepsilon) \rceil} \Pr \Big[ Z=k \Big] \, \bigg( \Big\lfloor \lg \binom{n}{k} \Big\rfloor -1 \bigg) \\ &\geq \sum_{k=\lfloor n(p-\varepsilon) \rfloor}^{\lceil n(p+\varepsilon) \rceil} \Pr \Big[ Z=k \Big] \bigg( \lg \frac{2^{n\mathbb{H}(p+\varepsilon)}}{n+1} -2 \bigg) \\ &= \Big( n\mathbb{H}(p+\varepsilon) - \lg(n+1) - 2 \Big) \Pr[|Z-np| \leq \varepsilon n] \\ &\geq \Big( n\mathbb{H}(p+\varepsilon) - \lg(n+1) - 2 \Big) \bigg( 1 - 2 \exp \bigg( -\frac{n\varepsilon^2}{4p} \bigg) \bigg), \\ \text{since } \mu = \mathbf{E}[Z] = np \text{ and} \\ &\mathbf{Pr} \Big[ |Z-np| \geq \frac{\varepsilon}{p} pn \Big] \leq 2 \exp \bigg( -\frac{np}{4} \Big( \frac{\varepsilon}{p} \Big)^2 \bigg) = 2 \exp \Big( -\frac{n\varepsilon^2}{4p} \Big), \text{ by the Chernoff inequality.} \end{split}$$

$$\begin{split} &\mathbf{E} \Big[ B \Big] = \sum_{k=0}^{n} \Pr[Z=k] \, \mathbf{E} \Big[ B \, \Big| \, Z=k \Big]. \\ &\mathbf{E} \Big[ B \Big] \geq \sum_{k=\lfloor n(p-\varepsilon)\rfloor}^{\lceil n(p+\varepsilon) \rceil} \Pr \Big[ Z=k \Big] \, \mathbf{E} \Big[ B \, \Big| \, Z=k \Big] \\ &\geq \sum_{k=\lfloor n(p-\varepsilon) \rfloor}^{\lceil n(p+\varepsilon) \rceil} \Pr \Big[ Z=k \Big] \left( \left\lfloor \lg \binom{n}{k} \right\rfloor -1 \right) \\ &\geq \sum_{k=\lfloor n(p-\varepsilon) \rfloor}^{\lceil n(p+\varepsilon) \rceil} \Pr \Big[ Z=k \Big] \left( \lg \frac{2^{n\mathbb{H}(p+\varepsilon)}}{n+1} -2 \right) \\ &= \left( n\mathbb{H}(p+\varepsilon) - \lg(n+1) -2 \right) \Pr[|Z-np| \leq \varepsilon n] \\ &\geq \left( n\mathbb{H}(p+\varepsilon) - \lg(n+1) -2 \right) \left( 1 - 2 \exp \left( -\frac{n\varepsilon^2}{4p} \right) \right), \\ \text{since } \mu = \mathbf{E}[Z] = np \text{ and} \\ &\mathbf{Pr} \Big[ |Z-np| \geq \frac{\varepsilon}{p} pn \Big] \leq 2 \exp \left( -\frac{np}{4} \left( \frac{\varepsilon}{p} \right)^2 \right) = 2 \exp \left( -\frac{n\varepsilon^2}{4p} \right), \text{ by the Chernoff inequality.} \end{split}$$

$$\begin{split} &\mathbf{E} \Big[ B \Big] = \sum_{k=0}^n \Pr[Z=k] \, \mathbf{E} \Big[ B \, \Big| \, Z=k \Big]. \\ &\mathbf{E} \Big[ B \Big] \geq \sum_{k=\lfloor n(p-\varepsilon) \rfloor}^{\lceil n(p+\varepsilon) \rceil} \Pr \Big[ Z=k \Big] \, \mathbf{E} \Big[ B \, \Big| \, Z=k \Big] \\ &\geq \sum_{k=\lfloor n(p-\varepsilon) \rfloor}^{\lceil n(p+\varepsilon) \rceil} \Pr \Big[ Z=k \Big] \, \bigg( \Big| \lg \left( \frac{n}{k} \right) \Big| -1 \bigg) \\ &\geq \sum_{k=\lfloor n(p-\varepsilon) \rfloor}^{\lceil n(p+\varepsilon) \rceil} \Pr \Big[ Z=k \Big] \bigg( \lg \frac{2^{n\mathbb{H}(p+\varepsilon)}}{n+1} - 2 \bigg) \\ &= \Big( n\mathbb{H}(p+\varepsilon) - \lg(n+1) - 2 \Big) \Pr[|Z-np| \leq \varepsilon n] \\ &\geq \Big( n\mathbb{H}(p+\varepsilon) - \lg(n+1) - 2 \Big) \bigg( 1 - 2 \exp \bigg( -\frac{n\varepsilon^2}{4p} \bigg) \bigg), \\ \text{since } \mu = \mathbf{E}[Z] = np \text{ and} \\ &\mathbf{Pr} \Big[ |Z-np| \geq \frac{\varepsilon}{p} pn \Big] \leq 2 \exp \bigg( -\frac{np}{4} \Big( \frac{\varepsilon}{p} \Big)^2 \bigg) = 2 \exp \Big( -\frac{n\varepsilon^2}{4p} \Big), \text{ by the Chernoff inequality.} \end{split}$$

- lacksquare Fix arepsilon>0, such that  $\mathbb{H}(p+arepsilon)>(1-\delta/4)\mathbb{H}(p)$ , p is fixed.
- ① For n sufficiently large:  $-\lg(n+1) \geq -\frac{\delta}{10} n \mathbb{H}(p)$ .

$$egin{aligned} \mathbf{E}[B] &\geq \left(1 - rac{\delta}{4} - rac{\delta}{10}
ight) n \mathbb{H}(p) \left(1 - rac{\delta}{10}
ight) \ &\geq (1 - \delta) \ n \mathbb{H}(p) \ , \end{aligned}$$

- lacksquare Fix arepsilon>0, such that  $\mathbb{H}(p+arepsilon)>(1-\delta/4)\mathbb{H}(p)$ , p is fixed.
- $oxed{3}$  For n sufficiently large:  $-\lg(n+1) \geq -rac{\delta}{10} n \mathbb{H}(p)$ .

$$egin{aligned} \mathbf{E}[B] &\geq \left(1 - rac{\delta}{4} - rac{\delta}{10}
ight) n \mathbb{H}(p) \left(1 - rac{\delta}{10}
ight) \ &\geq (1 - \delta) \ n \mathbb{H}(p) \ , \end{aligned}$$

- $\P \text{ Fix } \varepsilon > 0 \text{, such that } \mathbb{H}(p+\varepsilon) > (1-\delta/4)\mathbb{H}(p) \text{, } p \text{ is fixed.}$
- **3** For n sufficiently large:  $-\lg(n+1) \geq -\frac{\delta}{10} n \mathbb{H}(p)$ .

$$egin{aligned} \mathbf{E}[B] &\geq \left(1 - rac{\delta}{4} - rac{\delta}{10}
ight) n \mathbb{H}(p) \left(1 - rac{\delta}{10}
ight) \ &\geq (1 - \delta) \ n \mathbb{H}(p) \ , \end{aligned}$$

- ① Fix  $\varepsilon>0$ , such that  $\mathbb{H}(p+\varepsilon)>(1-\delta/4)\mathbb{H}(p)$ , p is fixed.
- **3** For n sufficiently large:  $-\lg(n+1) \geq -\frac{\delta}{10} n \mathbb{H}(p)$ .
- $\bullet$  ... also  $2\exp\left(-rac{narepsilon^2}{4p}
  ight) \leq rac{\delta}{10}$ .

$$egin{aligned} \mathbf{E}[B] &\geq \left(1 - rac{\delta}{4} - rac{\delta}{10}
ight) n \mathbb{H}(p) \left(1 - rac{\delta}{10}
ight) \ &\geq (1 - \delta) \ n \mathbb{H}(p) \ , \end{aligned}$$

- lacksquare Fix arepsilon>0, such that  $\mathbb{H}(p+arepsilon)>(1-\delta/4)\mathbb{H}(p)$ , p is fixed.
- **3** For n sufficiently large:  $-\lg(n+1) \ge -\frac{\delta}{10} n \mathbb{H}(p)$ .
- $\bullet$  ... also  $2\exp\left(-rac{narepsilon^2}{4p}
  ight) \leq rac{\delta}{10}$ .
- $\bullet$  For n large enough;

$$egin{aligned} \mathrm{E}[B] &\geq \left(1 - rac{\delta}{4} - rac{\delta}{10}
ight) n \mathbb{H}(p) igg(1 - rac{\delta}{10}igg) \ &\geq (1 - \delta) \ n \mathbb{H}(p) \,, \end{aligned}$$

- Need to prove upper bound.
- ① If input sequence x has probability  $\Pr[X=x]$ , then  $y=\mathsf{Ext}(x)$  has probability to be generated  $\geq \Pr[X=x]$ .
- $\odot$  All sequences of length |y| have equal probability to be generated (by definition).
- $\textcircled{0} \ \ 2^{|\mathsf{Ext}(x)|} \Pr[X = x] \leq 2^{|\mathsf{Ext}(x)|} \Pr[y = \mathsf{Ext}(x)] \leq 1.$
- $) \implies |\mathsf{Ext}(x)| \le \lg(1/\Pr[X=x])$
- $egin{aligned} egin{aligned} egin{aligned\\ egin{aligned} egi$

- Need to prove upper bound.
- ② If input sequence x has probability  $\Pr[X=x]$ , then  $y = \mathsf{Ext}(x)$  has probability to be generated  $> \Pr[X = x]$ .
- **a**  $2^{|\mathsf{Ext}(x)|} \Pr[X = x] < 2^{|\mathsf{Ext}(x)|} \Pr[y = \mathsf{Ext}(x)] < 1.$
- $\Longrightarrow |\mathsf{Ext}(x)| \leq \lg(1/\Pr[X=x])$
- $\bullet$   $\to$   $|B| = \sum_x \Pr[X = x] |\mathsf{Ext}(x)|$  $\leq \sum_x \Pr[X=x | \lg \frac{1}{\Pr[X=x]} = \mathbb{H}(X)$ .

Sariel (UIUC) CS573 Fall 2014 15 / 25

- Need to prove upper bound.
- ② If input sequence x has probability  $\Pr[X = x]$ , then  $y = \mathsf{Ext}(x)$  has probability to be generated  $\geq \Pr[X = x]$ .
- 3 All sequences of length |y| have equal probability to be generated (by definition).
- $\textcircled{0} \ \ 2^{|\mathsf{Ext}(x)|} \Pr[X = x] \leq 2^{|\mathsf{Ext}(x)|} \Pr[y = \mathsf{Ext}(x)] \leq 1.$
- $) \implies |\mathsf{Ext}(x)| \le \lg(1/\Pr[X=x])$
- $\begin{array}{l} \bullet \ \ \mathrm{E}\big[B\big] = \sum_x \Pr\big[X = x\big] \, |\mathsf{Ext}(x)| \\ \leq \sum_x \Pr\big[X = x\big] \lg \tfrac{1}{\Pr[X = x]} = \mathbb{H}(X) \,. \end{array}$

- Need to prove upper bound.
- ② If input sequence x has probability  $\Pr[X = x]$ , then  $y = \mathsf{Ext}(x)$  has probability to be generated  $\geq \Pr[X = x]$ .
- 3 All sequences of length |y| have equal probability to be generated (by definition).
- $) \implies |\mathsf{Ext}(x)| \le \lg(1/\Pr[X=x])$
- $\begin{array}{l} \bullet \ \ \mathrm{E}\big[B\big] = \sum_x \Pr\big[X = x\big] \, |\mathsf{Ext}(x)| \\ \leq \sum_x \Pr\big[X = x\big] \lg \tfrac{1}{\Pr[X = x]} = \mathbb{H}(X) \,. \end{array}$

- Need to prove upper bound.
- ② If input sequence x has probability  $\Pr[X = x]$ , then  $y = \mathsf{Ext}(x)$  has probability to be generated  $\geq \Pr[X = x]$ .
- **3** All sequences of length |y| have equal probability to be generated (by definition).

- $\begin{array}{l} \bullet \ \ \mathrm{E}\big[B\big] = \sum_x \Pr\big[X = x\big] \, |\mathsf{Ext}(x)| \\ \leq \sum_x \Pr\big[X = x\big] \lg \tfrac{1}{\Pr[X = x]} = \mathbb{H}(X) \,. \end{array}$

- Need to prove upper bound.
- ② If input sequence x has probability  $\Pr[X = x]$ , then  $y = \mathsf{Ext}(x)$  has probability to be generated  $\geq \Pr[X = x]$ .
- **3** All sequences of length |y| have equal probability to be generated (by definition).

- $\mathbf{0} \ \mathbf{E}[B] = \sum_{x} \Pr[X = x] |\mathsf{Ext}(x)| \\ \leq \sum_{x} \Pr[X = x] \lg \frac{1}{\Pr[X = x]} = \mathbb{H}(X).$

- Need to prove upper bound.
- ② If input sequence x has probability  $\Pr[X = x]$ , then  $y = \mathsf{Ext}(x)$  has probability to be generated  $\geq \Pr[X = x]$ .
- **3** All sequences of length |y| have equal probability to be generated (by definition).

- $\begin{array}{l} \bullet \ \ \mathrm{E}\big[B\big] = \sum_x \Pr\big[X = x\big] \, |\mathsf{Ext}(x)| \\ \leq \sum_x \Pr\big[X = x\big] \lg \frac{1}{\Pr[X = x]} = \mathbb{H}(X) \, . \end{array}$

- Need to prove upper bound.
- ② If input sequence x has probability  $\Pr[X = x]$ , then  $y = \mathsf{Ext}(x)$  has probability to be generated  $\geq \Pr[X = x]$ .
- **3** All sequences of length |y| have equal probability to be generated (by definition).

- $egin{aligned} egin{aligned} egin{aligned\\ egin{aligned} egi$

- Need to prove upper bound.
- ② If input sequence x has probability  $\Pr[X=x]$ , then  $y = \mathsf{Ext}(x)$  has probability to be generated  $> \Pr[X = x]$ .
- 3 All sequences of length |y| have equal probability to be generated (by definition).
- $2^{|\mathsf{Ext}(x)|}\Pr[X=x] < 2^{|\mathsf{Ext}(x)|}\Pr[y=\mathsf{Ext}(x)] < 1.$
- $\bullet \implies |\mathsf{Ext}(x)| < \lg(1/\Pr[X=x])$
- $lackbox{f 0} \; \mathrm{E}igl[Bigr] = \sum_x \Prigl[X=xigr] |\mathsf{Ext}(x)|$  $\leq \sum_{x} \Pr |X = x| \lg \frac{1}{\Pr[X = x]} = \mathbb{H}(X)$ .

Sariel (UIUC) CS573 Fall 2014 15 / 25

#### Part III

Coding: Shannon's Theorem

- binary symmetric channel with parameter p
- 2 sequence of bits  $x_1, x_2, \ldots$ , an
- output:  $y_1, y_2, \ldots,$  a sequence of bits such that...
- ${f O} \ \Pr[x_i=y_i]=1-p$  independently for each i.

- binary symmetric channel with parameter p
- $oldsymbol{2}$  sequence of bits  $x_1, x_2, \ldots$ , an
- $oldsymbol{0}$  output:  $y_1, y_2, \ldots,$  a sequence of bits such that...
- ①  $\Pr[x_i = y_i] = 1 p$  independently for each i.

- **1 binary symmetric channel** with parameter **p**
- 2 sequence of bits  $x_1, x_2, \ldots$ , an
- $oldsymbol{0}$  output:  $y_1, y_2, \ldots,$  a sequence of bits such that...
- ${f 0} \ \Pr[x_i=y_i]=1-p$  independently for each i.

- binary symmetric channel with parameter p
- 2 sequence of bits  $x_1, x_2, \ldots$ , an
- ullet output:  $y_1, y_2, \ldots,$  a sequence of bits such that...
- $\Pr[x_i = y_i] = 1 p$  independently for each i.

## Encoding/decoding with noise

- **1** (k, n) encoding function Enc:  $\{0, 1\}^k \to \{0, 1\}^n$  takes as input a sequence of k bits and outputs a sequence of n bits.
- $m{m{a}}(k,n)$  decoding function  $m{m{Dec}}: \{0,1\}^n o \{0,1\}^k$  takes as input a sequence of  $m{n}$  bits and outputs a sequence of  $m{k}$  bits.

## Encoding/decoding with noise

#### **Definition**

- **1** (k, n) encoding function Enc:  $\{0, 1\}^k \to \{0, 1\}^n$  takes as input a sequence of k bits and outputs a sequence of n bits.
- ② (k,n) decoding function  $Dec: \{0,1\}^n \to \{0,1\}^k$  takes as input a sequence of n bits and outputs a sequence of k bits.

#### Claude Elwood Shannon

Claude Elwood Shannon (April 30, 1916 - February 24, 2001), an American electrical engineer and mathematician, has been called "the father of information theory".

His master thesis was how to building boolean circuits for any boolean function.

## Shannon's theorem (1948)

#### Theorem (Shannon's theorem)

For a binary symmetric channel with parameter p < 1/2 and for any constants  $\delta, \gamma > 0$ , where n is sufficiently large, the following holds:

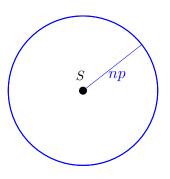
- (i) For an  $k \leq n(1 \mathbb{H}(p) \delta)$  there exists (k, n) encoding and decoding functions such that the probability the receiver fails to obtain the correct message is at most  $\gamma$  for every possible k-bit input messages.
- (ii) There are no (k, n) encoding and decoding functions with  $k \geq n(1 \mathbb{H}(p) + \delta)$  such that the probability of decoding correctly is at least  $\gamma$  for a k-bit input message chosen uniformly at random.

Sariel (UIUC) CS573 20 Fall 2014 20 / 25

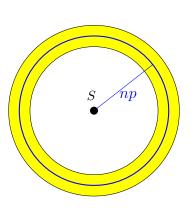
$$S = s_1 s_2 \dots s_n$$

S

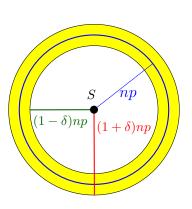
$$S = s_1 s_2 \dots s_n$$



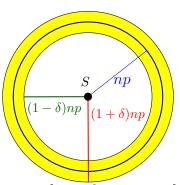
$$S = s_1 s_2 \dots s_n$$



$$S = s_1 s_2 \dots s_n$$



$$S = s_1 s_2 \dots s_n$$



One ring to rule them all!

- senders sent string  $S = s_1 s_2 \dots s_n$ .
- ② receiver got string  $T = t_1 t_2 \dots t_n$ .
- $oldsymbol{0}$  U: Hamming distance between S and T:  $U = \sum_i igl[ s_i 
  eq t_i igr]$  .
- $lacksquare{1}{2}$  By assumption:  $\mathbf{E}[U]=pn$ , and U is a binomial variable.
- ① By Chernoff inequality:  $U\in \left[(1-\delta)np,(1+\delta)np\right]$  with high probability, where  $\delta$  is tiny constant.
- $m{O}$   $m{T}$  is in a ring  $m{R}$  centered at  $m{S}$ , with inner radius  $(1-\delta)nm{p}$  and outer radius  $(1+\delta)nm{p}$ .
- This ring has

$$\sum_{i=(1-\delta)np}^{(1+\delta)np} inom{n}{i} \leq 2inom{n}{(1+\delta)np} \leq lpha = 2\cdot 2^{n\mathbb{H}((1+\delta)p)}.$$

- lacksquare senders sent string  $S = s_1 s_2 \dots s_n$ .
- 2 receiver got string  $T = t_1 t_2 \dots t_n$ .
- ① U: Hamming distance between S and T:  $U = \sum_i [s_i 
  eq t_i]$ .
- $lacksquare{1}{2}$  By assumption:  $\mathbf{E}[U]=pn$ , and U is a binomial variable.
- ① By Chernoff inequality:  $U \in [(1-\delta)np, (1+\delta)np]$  with high probability, where  $\delta$  is tiny constant.
- ② T is in a ring R centered at S, with inner radius  $(1 \delta)np$  and outer radius  $(1 + \delta)np$ .
- This ring has

$$\sum_{i=(1-\delta)np}^{(1+\delta)np} inom{n}{i} \leq 2inom{n}{(1+\delta)np} \leq lpha = 2\cdot 2^{n\mathbb{H}((1+\delta)p)}.$$

- lacksquare senders sent string  $S = s_1 s_2 \dots s_n$ .
- 2 receiver got string  $T = t_1 t_2 \dots t_n$ .
- $oldsymbol{0}$  U: Hamming distance between S and T:  $U=\sum_i igl|s_i
  eq t_iigr|$  .
- $lacksquare{1}{2}$  By assumption:  $\mathbf{E}[U]=pn$ , and U is a binomial variable.
- ① By Chernoff inequality:  $U\in \left[(1-\delta)np,(1+\delta)np\right]$  with high probability, where  $\delta$  is tiny constant.
- ② T is in a ring R centered at S, with inner radius  $(1 \delta)np$  and outer radius  $(1 + \delta)np$ .
- This ring has

$$\sum_{i=(1-\delta)np}^{(1+\delta)np} \binom{n}{i} \leq 2 \binom{n}{(1+\delta)np} \leq \alpha = 2 \cdot 2^{n\mathbb{H}((1+\delta)p)}.$$

- lacksquare senders sent string  $S = s_1 s_2 \dots s_n$ .
- 2 receiver got string  $T = t_1 t_2 \dots t_n$ .
- $lacksquare{0}\ U$ : Hamming distance between S and T:  $U=\sum_i \!\! \left[s_i 
  eq t_i
  ight].$
- $lacksquare{1}{2}$  By assumption:  $\mathbf{E}[U]=pn$ , and U is a binomial variable.
- ① By Chernoff inequality:  $U\in \left[(1-\delta)np,(1+\delta)np\right]$  with high probability, where  $\delta$  is tiny constant.
- $m{O}$   $m{T}$  is in a ring  $m{R}$  centered at  $m{S}$ , with inner radius  $(1-\delta)nm{p}$  and outer radius  $(1+\delta)nm{p}$ .
- This ring has

$$\sum_{i=(1-\delta)np}^{(1+\delta)np} \binom{n}{i} \leq 2 \binom{n}{(1+\delta)np} \leq \alpha = 2 \cdot 2^{n\mathbb{H}((1+\delta)p)}.$$

- lacksquare senders sent string  $S = s_1 s_2 \dots s_n$ .
- 2 receiver got string  $T = t_1 t_2 \dots t_n$ .
- $lacksquare{0}\ U$ : Hamming distance between S and T:  $U=\sum_i \! \left[s_i 
  eq t_i
  ight].$
- **3** By assumption:  $\mathbf{E}[U] = pn$ , and U is a binomial variable.
- ullet By Chernoff inequality:  $U\in \left[(1-\delta)np,(1+\delta)np
  ight]$  with high probability, where  $\delta$  is tiny constant.
- **1** T is in a ring R centered at S, with inner radius  $(1 \delta)np$  and outer radius  $(1 + \delta)np$ .
- This ring has

$$\sum_{i=(1-\delta)np}^{(1+\delta)np} \binom{n}{i} \leq 2 \binom{n}{(1+\delta)np} \leq \alpha = 2 \cdot 2^{n\mathbb{H}((1+\delta)p)}.$$

- lacksquare senders sent string  $S = s_1 s_2 \dots s_n$ .
- 2 receiver got string  $T = t_1 t_2 \dots t_n$ .
- **1** U: Hamming distance between S and T:  $U = \sum_i [s_i 
  eq t_i]$ .
- **5** By assumption:  $\mathbf{E}[U] = pn$ , and U is a binomial variable.
- **1** By Chernoff inequality:  $U \in \left[ (1 \delta) np, (1 + \delta) np \right]$  with high probability, where  $\delta$  is tiny constant.
- **1** T is in a ring R centered at S, with inner radius  $(1 \delta)np$  and outer radius  $(1 + \delta)np$ .
- This ring has

$$\sum_{i=(1-\delta)np}^{(1+\delta)np} inom{n}{i} \leq 2inom{n}{(1+\delta)np} \leq lpha = 2\cdot 2^{n\mathbb{H}((1+\delta)p)}.$$

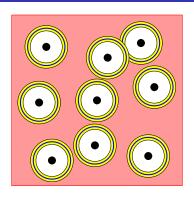
- lacksquare senders sent string  $S = s_1 s_2 \dots s_n$ .
- $oldsymbol{2}$  receiver got string  $T=t_1t_2\ldots t_n$ .
- $oldsymbol{0}$  U: Hamming distance between S and T:  $U = \sum_i [s_i 
  eq t_i]$ .
- **5** By assumption:  $\mathbf{E}[U] = pn$ , and U is a binomial variable.
- **1** By Chernoff inequality:  $U \in \left[ (1 \delta) np, (1 + \delta) np \right]$  with high probability, where  $\delta$  is tiny constant.
- **1** T is in a ring R centered at S, with inner radius  $(1 \delta)np$  and outer radius  $(1 + \delta)np$ .
- This ring has

$$\sum_{i=(1-\delta)np}^{(1+\delta)np} inom{n}{i} \leq 2inom{n}{(1+\delta)np} \leq lpha = 2\cdot 2^{n\mathbb{H}((1+\delta)p)}.$$

- lacksquare senders sent string  $S = s_1 s_2 \dots s_n$ .
- ② receiver got string  $T=t_1t_2\dots t_n$ .
- $oldsymbol{0}$  U: Hamming distance between S and T:  $U = \sum_i [s_i 
  eq t_i]$ .
- **5** By assumption:  $\mathbf{E}[U] = pn$ , and U is a binomial variable.
- **1** By Chernoff inequality:  $U \in \left[ (1 \delta) np, (1 + \delta) np \right]$  with high probability, where  $\delta$  is tiny constant.
- **1** T is in a ring R centered at S, with inner radius  $(1 \delta)np$  and outer radius  $(1 + \delta)np$ .
- This ring has

$$\sum_{i=(1-\delta)np}^{(1+\delta)np} \binom{n}{i} \leq 2 \binom{n}{(1+\delta)np} \leq \alpha = 2 \cdot 2^{n\mathbb{H}((1+\delta)p)}.$$

# Many rings for many codewords...



- ① Pick as many disjoint rings as possible:  $R_1, \ldots, R_{\kappa}$ .
- If every word in the hypercube would be covered...
- ullet ... use  $\mathbf{2}^n$  codewords  $\implies \kappa \geq$

$$\kappa \geq rac{2^n}{|R|} \geq rac{2^n}{2 \cdot 2^{n\mathbb{H}((1+\delta)p)}} pprox 2^{n(1-\mathbb{H}((1+\delta)p))}.$$

- ① Consider all possible strings of length k such that  $2^k \leq \kappa$ .
- ullet Map ith string in  $\{0,1\}^k$  to the center  $C_i$  of the ith ring  $R_i$ .
- $lack {f 0}$  If send  $C_i \Longrightarrow$  receiver gets a string in  $R_i$ .
- O Decoding is easy find the ring  $R_i$  containing the received string, take its center string  $C_i$ , and output the original string it was mapped to.
- $egin{aligned} egin{aligned} egin{aligned} & ext{How many bits?} \ & k = \lfloor \log \kappa 
  floor & = n ig(1 \mathbb{H}\left((1 + \delta)p
  ight)ig) pprox n(1 \mathbb{H}\left(p
  ight)), \end{aligned}$

- **①** Pick as many disjoint rings as possible:  $R_1, \ldots, R_{\kappa}$ .
- ② If every word in the hypercube would be covered...
- ${ ext{ odewords }} \implies \kappa \geq$

$$\kappa \geq rac{2^n}{|R|} \geq rac{2^n}{2 \cdot 2^{n \mathbb{H}((1+\delta)p)}} pprox 2^{n(1-\mathbb{H}((1+\delta)p))}.$$

- ① Consider all possible strings of length k such that  $2^k \leq \kappa$ .
- ullet Map ith string in  $\{0,1\}^k$  to the center  $C_i$  of the ith ring  $R_i$ .
- ullet If send  $C_i \Longrightarrow$  receiver gets a string in  $R_i$ .
- ① Decoding is easy find the ring  $R_i$  containing the received string, take its center string  $C_i$ , and output the original string it was mapped to.
- $egin{aligned} egin{aligned} egin{aligned\\ egin{aligned} egi$

- **①** Pick as many disjoint rings as possible:  $R_1, \ldots, R_{\kappa}$ .
- If every word in the hypercube would be covered...
- $oxed{solution} : ... ext{ use } 2^n ext{ codewords } \Longrightarrow \kappa \geq \ \kappa \geq rac{2^n}{|R|} \geq rac{2^n}{2 \cdot 2^{n \mathbb{H}((1+\delta)p)}} pprox 2^{n(1-\mathbb{H}((1+\delta)p))}.$
- $ext{ }$  Consider all possible strings of length k such that  $2^k \leq \kappa$ .
- ullet Map ith string in  $\left\{0,1
  ight\}^k$  to the center  $C_i$  of the ith ring  $R_i$
- ullet If send  $C_i \Longrightarrow$  receiver gets a string in  $R_i$ .
- $\odot$  Decoding is easy find the ring  $R_i$  containing the received string, take its center string  $C_i$ , and output the original string it was mapped to.
- $egin{aligned} egin{aligned} egin{aligned\\ egin{aligned} egi$

- **①** Pick as many disjoint rings as possible:  $R_1, \ldots, R_{\kappa}$ .
- If every word in the hypercube would be covered...
- ullet ... use  $\mathbf{2}^n$  codewords  $\implies \kappa \geq$

$$\kappa \geq rac{2^n}{|R|} \geq rac{2^n}{2 \cdot 2^{n\mathbb{H}((1+\delta)p)}} pprox 2^{n(1-\mathbb{H}((1+\delta)p))}.$$

- Consider all possible strings of length k such that  $2^k \le \kappa$ .
- ullet Map ith string in  $\{0,1\}^k$  to the center  $C_i$  of the ith ring  $R_i$ .
- O Decoding is easy find the ring  $R_i$  containing the received string, take its center string  $C_i$ , and output the original string it was mapped to.
- $egin{aligned} egin{aligned} egin{aligned} & ext{How many bits?} \ & k = \lfloor \log \kappa 
  floor & = n ig(1 \mathbb{H}\left((1 + \delta)p
  ight)ig) pprox n(1 \mathbb{H}\left(p
  ight)), \end{aligned}$

- **①** Pick as many disjoint rings as possible:  $R_1, \ldots, R_{\kappa}$ .
- If every word in the hypercube would be covered...
- ullet ... use  $\mathbf{2}^n$  codewords  $\implies \kappa \geq$

$$\kappa \geq rac{2^n}{|R|} \geq rac{2^n}{2 \cdot 2^{n\mathbb{H}((1+\delta)p)}} pprox 2^{n(1-\mathbb{H}((1+\delta)p))}.$$

- **©** Consider all possible strings of length k such that  $2^k \le \kappa$ .
- ullet Map ith string in  $\{0,1\}^k$  to the center  $C_i$  of the ith ring  $R_i$ .
- ullet If send  $C_i \Longrightarrow$  receiver gets a string in  $R_i$ .
- ① Decoding is easy find the ring  $R_i$  containing the received string, take its center string  $C_i$ , and output the original string it was mapped to.
- $egin{aligned} egin{aligned} egin{aligned\\ egin{aligned} egi$

- **①** Pick as many disjoint rings as possible:  $R_1, \ldots, R_{\kappa}$ .
- If every word in the hypercube would be covered...
- ullet ... use  $\mathbf{2}^n$  codewords  $\implies \kappa \geq$

$$\kappa \geq rac{2^n}{|R|} \geq rac{2^n}{2 \cdot 2^{n\mathbb{H}((1+\delta)p)}} pprox 2^{n(1-\mathbb{H}((1+\delta)p))}.$$

- **©** Consider all possible strings of length k such that  $2^k \le \kappa$ .
- **1** Map ith string in  $\{0,1\}^k$  to the center  $C_i$  of the ith ring  $R_i$ .
- ullet If send  $C_i \Longrightarrow$  receiver gets a string in  $R_i$ .
- ① Decoding is easy find the ring  $R_i$  containing the received string, take its center string  $C_i$ , and output the original string it was mapped to.
- $egin{aligned} egin{aligned} egin{aligned\\ egin{aligned} egi$

- **①** Pick as many disjoint rings as possible:  $R_1, \ldots, R_{\kappa}$ .
- If every word in the hypercube would be covered...
- ullet ... use  $\mathbf{2}^n$  codewords  $\implies \kappa \geq$

$$\kappa \geq rac{2^n}{|R|} \geq rac{2^n}{2 \cdot 2^{n\mathbb{H}((1+\delta)p)}} pprox 2^{n(1-\mathbb{H}((1+\delta)p))}.$$

- **©** Consider all possible strings of length k such that  $2^k \le \kappa$ .
- **1** Map ith string in  $\{0,1\}^k$  to the center  $C_i$  of the ith ring  $R_i$ .
- $lackbox{0}$  If send  $C_i \implies$  receiver gets a string in  $R_i$ .
- ① Decoding is easy find the ring  $R_i$  containing the received string, take its center string  $C_i$ , and output the original string it was mapped to.
- $egin{aligned} egin{aligned} egin{aligned\\ egin{aligned} egi$

- **①** Pick as many disjoint rings as possible:  $R_1, \ldots, R_{\kappa}$ .
- If every word in the hypercube would be covered...
- ullet ... use  $2^n$  codewords  $\implies \kappa \geq$

$$\kappa \geq rac{2^n}{|R|} \geq rac{2^n}{2 \cdot 2^{n\mathbb{H}((1+\delta)p)}} pprox 2^{n(1-\mathbb{H}((1+\delta)p))}.$$

- **©** Consider all possible strings of length k such that  $2^k \le \kappa$ .
- **1** Map ith string in  $\{0,1\}^k$  to the center  $C_i$  of the ith ring  $R_i$ .
- lacktriangledown If send  $C_i \implies$  receiver gets a string in  $R_i$ .
- lacktriangledown Decoding is easy find the ring  $R_i$  containing the received string, take its center string  $C_i$ , and output the original string it was mapped to.
- $egin{aligned} egin{aligned} egin{aligned} & ext{How many bits?} \ & k = \lfloor \log \kappa 
  floor & = n ig(1 \mathbb{H}\left((1 + \delta)p
  ight)ig) pprox n(1 \mathbb{H}\left(p
  ight)), \end{aligned}$

- **①** Pick as many disjoint rings as possible:  $R_1, \ldots, R_{\kappa}$ .
- If every word in the hypercube would be covered...
- lacksquare ... use  $\mathbf{2}^n$  codewords  $\implies \kappa \geq$

$$\kappa \geq rac{2^n}{|R|} \geq rac{2^n}{2 \cdot 2^{n\mathbb{H}((1+\delta)p)}} pprox 2^{n(1-\mathbb{H}((1+\delta)p))}.$$

- **©** Consider all possible strings of length k such that  $2^k \le \kappa$ .
- $lacksquare{0}$  Map ith string in  $\{0,1\}^k$  to the center  $C_i$  of the ith ring  $R_i$ .
- lacktriangledown If send  $C_i \implies$  receiver gets a string in  $R_i$ .
- lacktriangledown Decoding is easy find the ring  $R_i$  containing the received string, take its center string  $C_i$ , and output the original string it was mapped to.
- $egin{aligned} egin{aligned} egin{aligned\\ egin{aligned} egi$

- Can not find such a large set of disjoint rings.
- Reason is that when you pack rings (or balls) you are going to have wasted spaces around.
- Overcome this: allow rings to overlap somewhat.
- Makes things considerably more involved.
- Details in class notes.

- Can not find such a large set of disjoint rings.
- Reason is that when you pack rings (or balls) you are going to have wasted spaces around.
- Overcome this: allow rings to overlap somewhat.
- Makes things considerably more involved.
- Details in class notes.

- Can not find such a large set of disjoint rings.
- Reason is that when you pack rings (or balls) you are going to have wasted spaces around.
- Overcome this: allow rings to overlap somewhat.
- Makes things considerably more involved.
- Details in class notes.

- Can not find such a large set of disjoint rings.
- Reason is that when you pack rings (or balls) you are going to have wasted spaces around.
- Overcome this: allow rings to overlap somewhat.
- Makes things considerably more involved.
- Details in class notes.

- Can not find such a large set of disjoint rings.
- Reason is that when you pack rings (or balls) you are going to have wasted spaces around.
- Overcome this: allow rings to overlap somewhat.
- Makes things considerably more involved.
- Details in class notes.