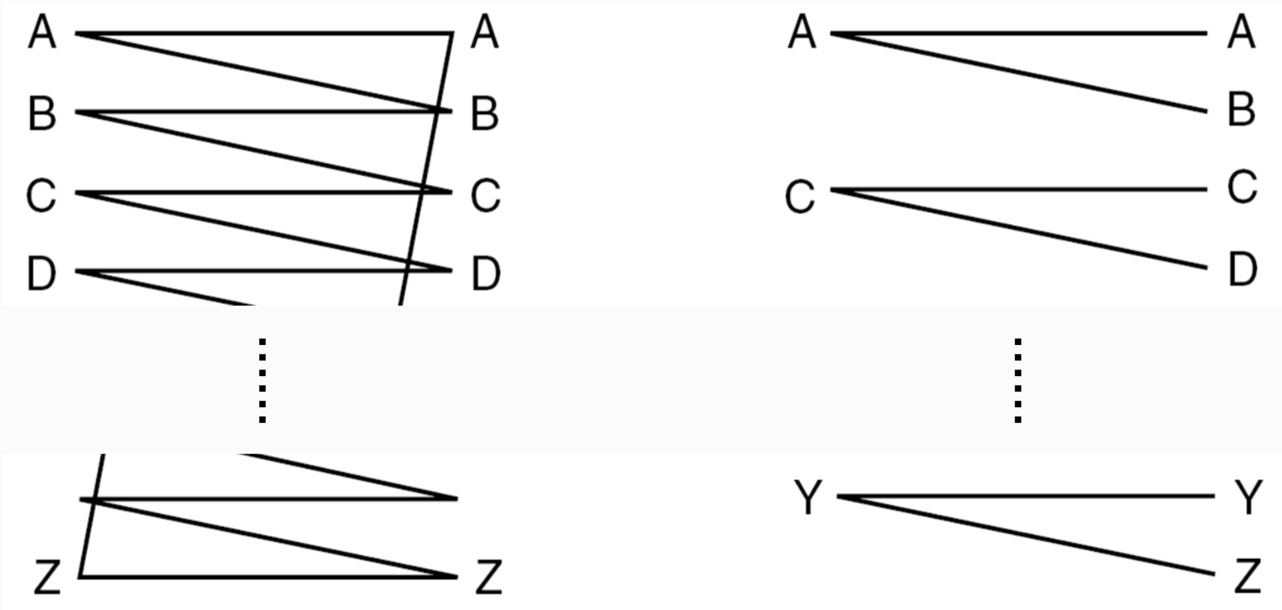
信息论导论

第7讲 信道编码的基本概念，联合典型集与联合渐近均分性

[信息论教材中页码范围] 信道编码的基本概念：p191~p195，
联合典型集与联合渐进均分性：p195~p198

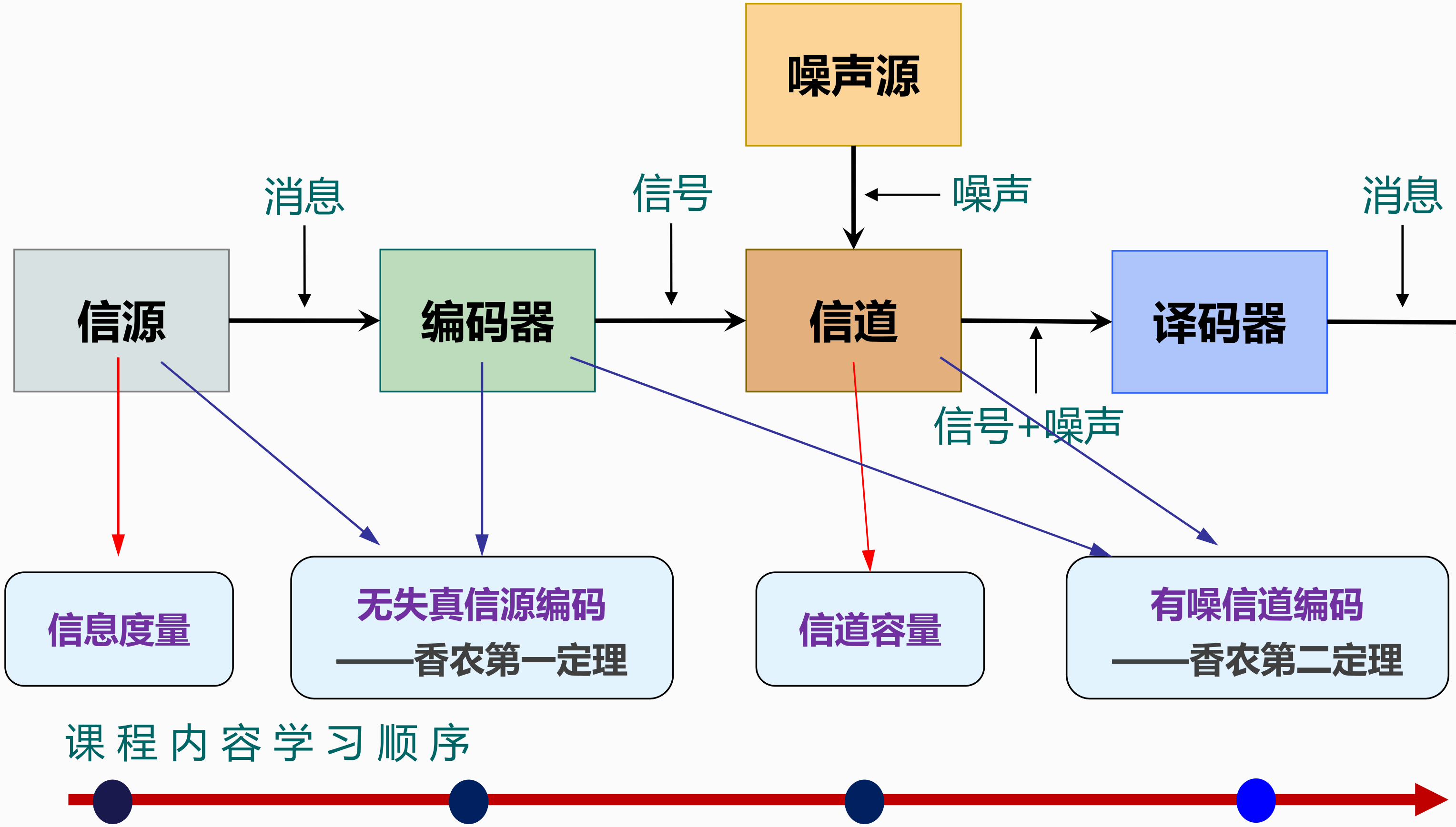
信息学部-信息科学与技术学院 吴绍华
hitwush@hit.edu.cn

思考：有噪声打字机信道



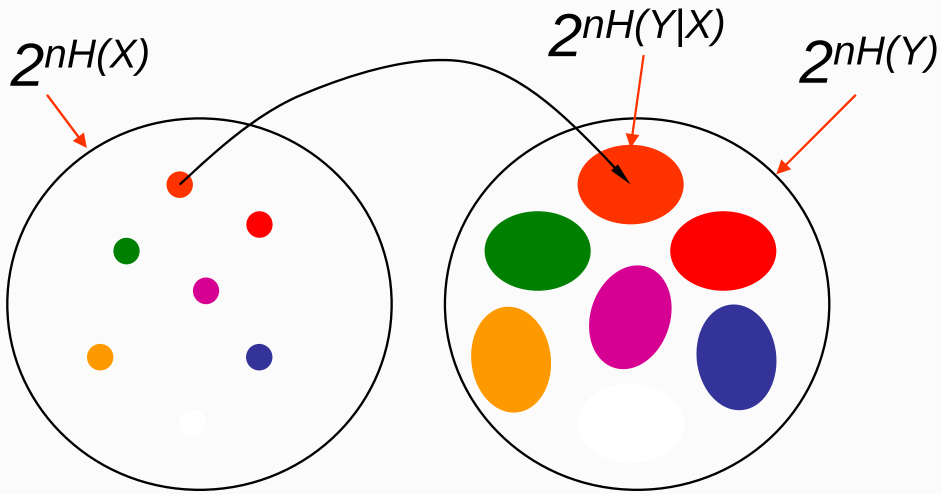
$$C = \max I(X;Y) = \max\left(H(Y) - H(Y|X)\right)$$
$$= \max H(Y) - 1 = \log 26 - 1 = \log 13$$

课程内容进度安排



作为导论课，本课程只讨论“离散”信源、信道

通过信道编码实现近容量传输的整体认识

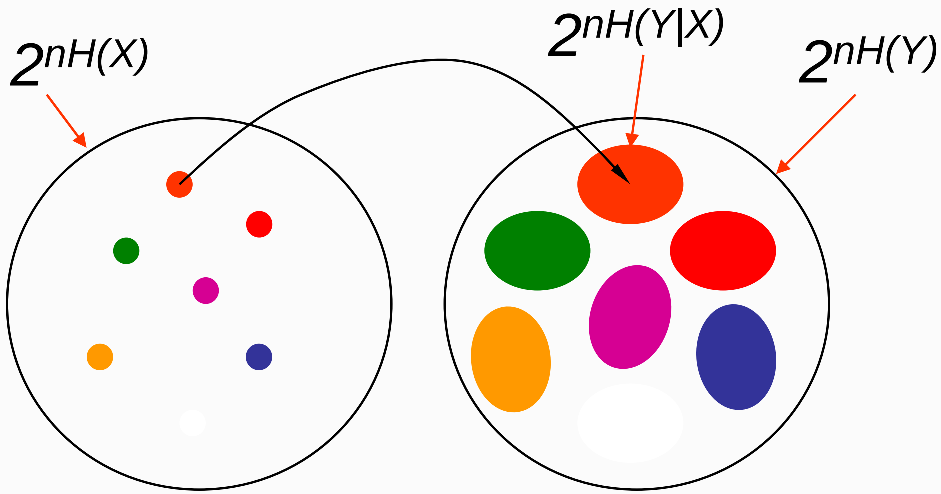


= n次使用信道 = 码长为 n的码字

考虑将长度为 n 的符号分组送入信道（n 次使用信道，每个分组即为一个信道编码的码字）:

- 每个输入典型序列 $X_{1:n}$，对应的输出典型序列约有 $2^{nH(Y|X)}$ 个
- 所有的输出典型序列的总数约为 $2^{nH(Y)}$
- 为了实现近乎无差错传输，挑选出那些输出典型集几乎没有重叠的输入序列作为传输码字即可（根据输出可近乎无误差估计输入）。**思考**：可最多挑选出多少个输入序列？

通过信道编码实现近容量传输的整体认识

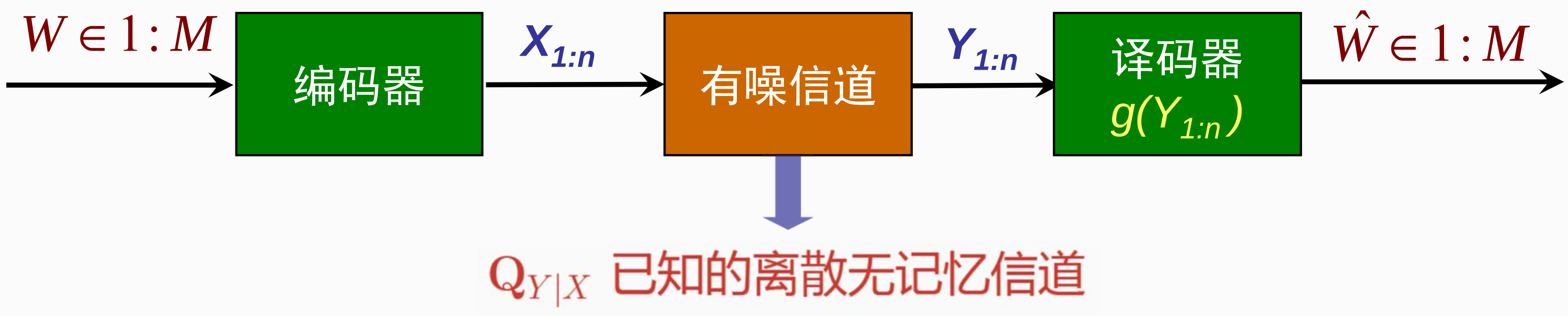


- 可挑选出的输出典型集几乎不重叠的输入序列最大数量为：

$$\frac{2^{nH(Y)}}{2^{nH(Y|X)}} = 2^{n(H(Y)-H(Y|X))} = 2^{nI(X;Y)}$$

“信道编码定理”提前概览：只要码字长度 n 足够大，对任意小于 C 的码率均可实现误差概率近似为零的可靠传输。

信道编码



定义

信道编码是从待编码序列（也可泛称为消息）到信道编码码字的映射，消息总数为 M、码字长度为 n 的信道编码记为(M, n) 码：

- 消息的序号为$\{1, 2, \ldots, M\}$
- 编码函数记为$X_{1:n} : \{1, 2, \ldots, M\} \to \mathscr{X}^n$，生成的码字记为$X_{1:n}(1), X_{1:n}(2), \ldots, X_{1:n}(M)$，这些码字的集合被称为码簿。
- 译码器记为：$g(Y_{1:n}) \in 1 : M$

误差概率

定义

- 码字误差概率:

$$\lambda_w = \Pr\left(g(Y_{1:n}) \neq w \middle| X_{1:n} = X_{1:n}(w)\right) = \sum_{y_{1:n} \in \mathscr{Y}^n} p\left(y_{1:n} \middle| x_{1:n}(w)\right) \delta_{g(y_{1:n}) \neq w}$$

- 最大误差概率:

$$\lambda^{(n)} = \max_{1 \leq w \leq M} \lambda_w$$

- 平均误差概率:

$$P_e^{(n)} = \frac{1}{M} \sum_{w=1}^{M} \lambda_w$$

码率，可达码率

定义（码率）

(M, n) 码的码率： $R = (\log M)/n$，是平均每个码符号传输携带的信息

定义（可达码率）

如果存在一个 $(\lceil 2^{nR} \rceil, n)$ 码，随着$n \to \infty$，码字的最大误差概率 $\lambda^{(n)}$ 趋于 0，则称 “码率R 可达”。

注释：为了简化书写，通常把 $(\lceil 2^{nR} \rceil, n)$ 码用 $(2^{nR}, n)$ 码来表示。

定义

离散无记忆信道（DMC）的容量是其所有可达速率的上确界。

联合典型集

定义（联合典型集）

联合典型集 $J_\varepsilon^{(n)}$ 是经验熵与真实熵相差 ε 以内的长度为 n 的序列对的集合，即：

$$J_\varepsilon^{(n)} = \left\{ (x_{1:n}, y_{1:n}) \in \mathscr{X}^n \times \mathscr{Y}^n : \begin{array}{l} \left| -\frac{1}{n} \log p(x_{1:n}) - H(X) \right| < \varepsilon, \\ \left| -\frac{1}{n} \log p(y_{1:n}) - H(Y) \right| < \varepsilon, \\ \left| -\frac{1}{n} \log p(x_{1:n}, y_{1:n}) - H(X,Y) \right| < \varepsilon \end{array} \right\}$$

其中 $p(x_{1:n}, y_{1:n}) = \prod_{i=1}^{n} p(x_i, y_i)$。

联合典型 - 举例

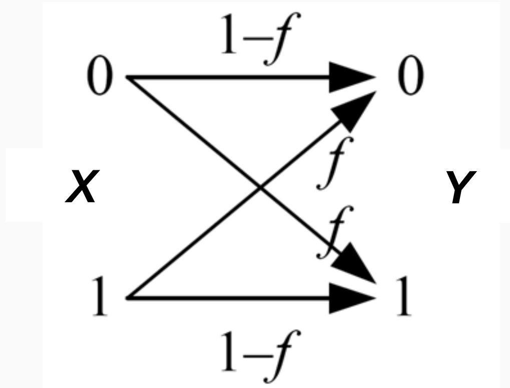
- 二元对称信道

  $f = 0.2$

  $\mathbf{p}_X = [0.75; 0.25]$

  $\mathbf{p}_Y = [0.65; 0.35]$

  $$\mathbf{P}_{XY} = \begin{bmatrix} 0.6 & 0.15 \\ 0.05 & 0.2 \end{bmatrix}$$



- 联合典型示例（对任意 ε ）：

$$X_{1:20} = 11111000000000000000$$

$$Y_{1:20} = 11110111000000000000$$

可按联合典型集定义中的三个条件，判断出这两个序列是联合典型的。

联合典型集的性质

哈爾濱工業大學(深圳) HARBIN INSTITUTE OF TECHNOLOGY, SHENZHEN

联合典型集的性质

1. 个体概率： $(x_{1:n}, y_{1:n}) \in J_\varepsilon^{(n)} \Rightarrow$

$$H(X,Y) - \varepsilon \le -\frac{1}{n}\log p(x_{1:n}, y_{1:n}) \le H(X,Y) + \varepsilon$$

2. 整体概率：

$$\Pr\{J_\varepsilon^{(n)}\} = p((x_{1:n}, y_{1:n}) \in J_\varepsilon^{(n)}) > 1 - \varepsilon \quad \text{for } n > N_\varepsilon$$

3. 元素个数：

$$(1-\varepsilon)2^{n(H(X,Y)-\varepsilon)} \le^{\text{for } n > N_\varepsilon} |J_\varepsilon^{(n)}| \le 2^{n(H(X,Y)+\varepsilon)}$$

联合典型集的性质 – 证明

性质 (1) 和 (2) 的证明

- 性质 (1) 的证明可直接由联合典型集 $J_\varepsilon^{(n)}$ 的定义得出。
- 根据弱大数定律：

$$-\frac{1}{n}\log p(X_{1:n}) \xrightarrow[n\to\infty]{\text{prob}} -E\log p(X) = H(X)$$

因此，存在 N_1，使得对于所有 $n > N_1$，

$$p\left(\left|-\frac{1}{n}\log p(x_{1:n}) - H(X)\right| > \varepsilon\right) < \frac{\varepsilon}{3}$$

类似地，对应另外两个条件存在 N_2、N_3。通过选取 $N_\varepsilon = \max\{N_1, N_2, N_3\}$，三个集合的并集的概率必定小于 ε 。因此，对于 $n > N_\varepsilon$，有 $\Pr\{J_\varepsilon^{(n)}\} > 1-\varepsilon$ 。

联合典型集的性质 – 证明

性质 (3) 的证明

对于 $n > N_\varepsilon$ ，有：

$$1-\varepsilon < p((x_{1:n}, y_{1:n}) \in J_\varepsilon^{(n)}) = \sum_{(x_{1:n}, y_{1:n}) \in J_\varepsilon^{(n)}} p(x_{1:n}, y_{1:n})$$

$$\le |J_\varepsilon^{(n)}| \max_{(x_{1:n}, y_{1:n}) \in J_\varepsilon^{(n)}} p(x_{1:n}, y_{1:n}) = |J_\varepsilon^{(n)}| 2^{-n(H(X,Y)-\varepsilon)} \quad \Rightarrow \quad (3) - \text{下界}$$

$$1 = \sum_{(x_{1:n}, y_{1:n}) \in \mathscr{X}^n \times \mathscr{Y}^n} p(x_{1:n}, y_{1:n}) \ge \sum_{(x_{1:n}, y_{1:n}) \in J_\varepsilon^{(n)}} p(x_{1:n}, y_{1:n})$$

$$\ge |J_\varepsilon^{(n)}| \min_{(x_{1:n}, y_{1:n}) \in J_\varepsilon^{(n)}} p(x_{1:n}, y_{1:n}) = |T_\varepsilon^{(n)}| 2^{-n(H(X,Y)+\varepsilon)} \quad \Rightarrow \quad (3) - \text{上界}$$

联合渐近均分性

联合渐近均分性（Joint AEP）

若 $\mathbf{p}_{X'} = \mathbf{p}_X$ 且 $\mathbf{p}_{Y'} = \mathbf{p}_Y$ ，但 X' 和 Y' 是相互独立的，则

$$p((x'_{1:n}, y'_{1:n}) \in J_\varepsilon^{(n)}) \leq 2^{-n(I(X;Y)-3\varepsilon)}$$

并且，对于足够大的 n（$n > N_\varepsilon$ ），

$$p((x'_{1:n}, y'_{1:n}) \in J_\varepsilon^{(n)}) \geq (1-\varepsilon)2^{-n(I(X;Y)+3\varepsilon)}$$

- 总体概率 $\leq$ 最大数量 $\times$ 最大个体概率
- 总体概率 $\geq$ 最小数量 $\times$ 最小个体概率

联合渐近均分性 - 证明

哈爾濱工業大學(深圳) HARBIN INSTITUTE OF TECHNOLOGY, SHENZHEN

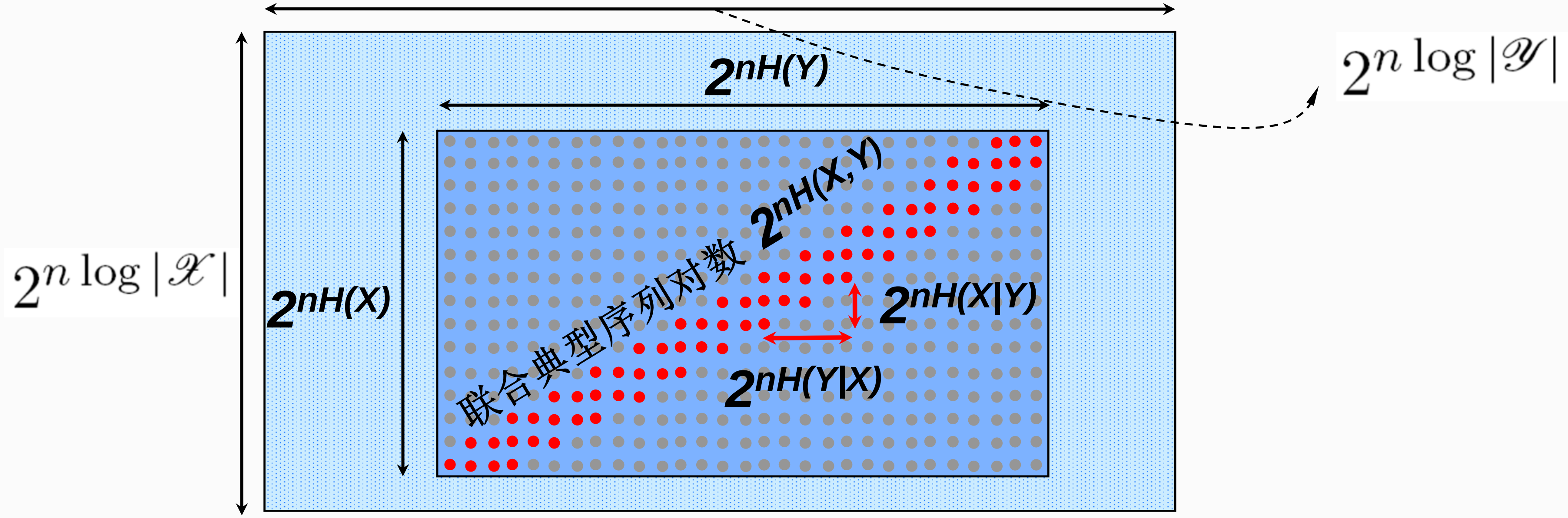
证明

$$p((x'_{1:n}, y'_{1:n}) \in J_\varepsilon^{(n)}) = \sum_{(x'_{1:n}, y'_{1:n}) \in J_\varepsilon^{(n)}} p(x'_{1:n}, y'_{1:n})$$
$$= \sum_{(x'_{1:n}, y'_{1:n}) \in J_\varepsilon^{(n)}} p(x'_{1:n}) p(y'_{1:n})$$

$$p((x'_{1:n}, y'_{1:n}) \in J_\varepsilon^{(n)}) \le |J_\varepsilon^{(n)}| \max_{(x'_{1:n}, y'_{1:n}) \in J_\varepsilon^{(n)}} p(x'_{1:n}, y'_{1:n})$$
$$\le 2^{n(H(X,Y)+\varepsilon)} 2^{-n(H(X)-\varepsilon)} 2^{-n(H(Y)-\varepsilon)} = 2^{-n(I(X,Y)-3\varepsilon)}$$

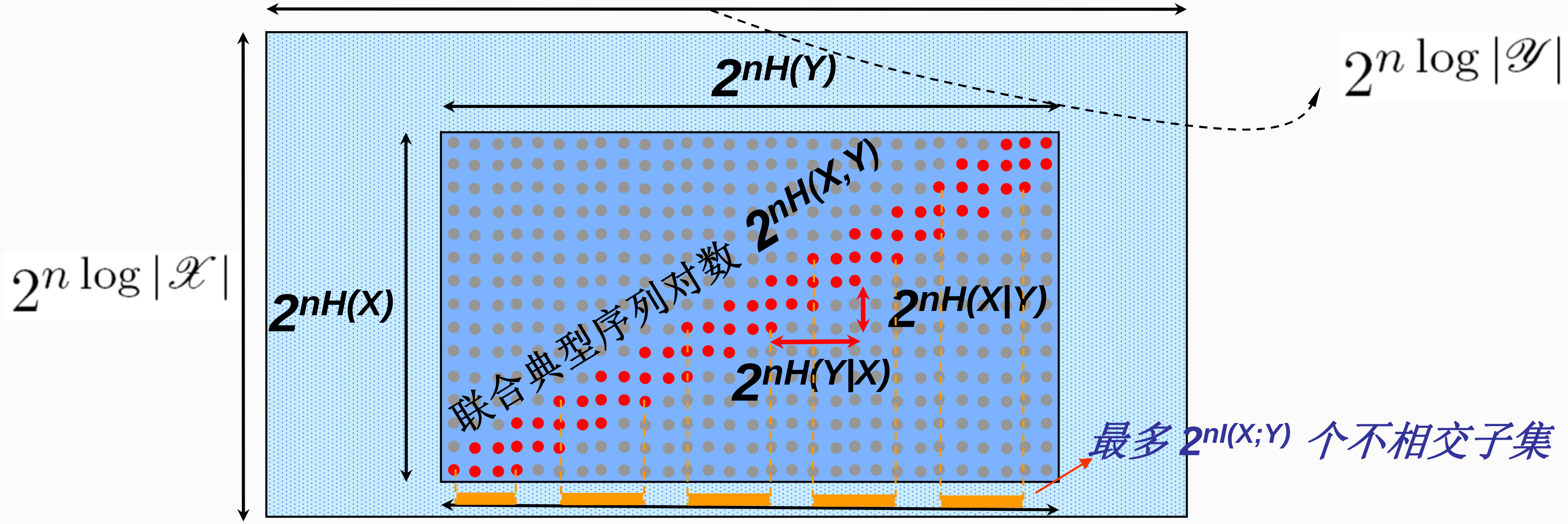
$$p((x'_{1:n}, y'_{1:n}) \in J_\varepsilon^{(n)}) \ge |J_\varepsilon^{(n)}| \min_{(x'_{1:n}, y'_{1:n}) \in J_\varepsilon^{(n)}} p(x'_{1:n}, y'_{1:n})$$
$$\ge (1-\varepsilon) 2^{-n(I(X,Y)+3\varepsilon)} \quad \text{对于 } n > N_\varepsilon$$

联合典型 - 图示理解



- 每个点表示一个序列对 $(x_{1:n}, y_{1:n})$
- 外部矩形表示所有可能的序列对
- 内部矩形表示在 $X_{1:n}$ 典型、或 $Y_{1:n}$ 典型的序列对，但不一定是联合典型的
- 红点表示联合典型的序列对

联合典型集 - 图示理解



- 典型的 $X_{1:n}$ 序列大约共有 $2^{nH(X)}$ 个，每个典型的 $X_{1:n}$ 与大约 $2^{nH(Y|X)}$ 个典型的 $Y_{1:n}$ 构成联合典型序列对
- 联合典型对在内部矩形中所占比例为 $2^{-nI(X;Y)}$
- 无差错信道编码方案：选择其联合典型的 $Y_{1:n}$ 序列集合不重叠的 $X_{1:n}$ 序列进行编码传输，使用联合典型性进行解码

信道编码定理（Shannon第二定理）

哈爾濱工業大學(深圳)
HARBIN INSTITUTE OF TECHNOLOGY, SHENZHEN

信道编码定理

如果码率 $R < C$ ，则该码率是可达的；如果 $R > C$ ，则该码率是不可达的。

- 对任意码率 $R < C$ ，一定存在至少一个 $(2^{nR}, n)$ 码，当 $n \to \infty$ 时，其最大误差概率 $\lambda^{(n)}$ 趋于 0 。
- 任何最大误差概率 $\lambda^{(n)}$ 趋于 0（随 $n \to \infty$ ）的 $(2^{nR}, n)$ 码，一定满足 $R \leq C$ 。

上述定理在提出之时，是非常反直觉的结论：尽管存在信道传输错误，但只要 $R < C$ ，就可以得到任意低的误码率。

———————将在下一次课中证明

结束