

# 信息论导论

## 第8讲 Shannon第二定理

[信息论教材中页码范围] 正定理证明: p198~p205,  
逆定理证明: p206~p208

信息学部-信息科学与技术学院 吴绍华

[hitwush@hit.edu.cn](mailto:hitwush@hit.edu.cn)

## 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  $n \rightarrow \infty$  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  $n \rightarrow \infty$ ) 的  $(2^{nR}, n)$  码，一定满足  $R \leq C$ 。
- 也被称为 Shannon 第二定理。
- 难点：一个码的最大误差概率很难确定。
- Shannon 的思路：
  - 考虑随机生成的码；
  - 计算所有可能的随机生成码的平均误差概率的期望值，并证明其很小；
  - 证明这意味着至少存在一个码，其最大误差概率很小。

# 信道编码定理可达性（正定理）证明步骤

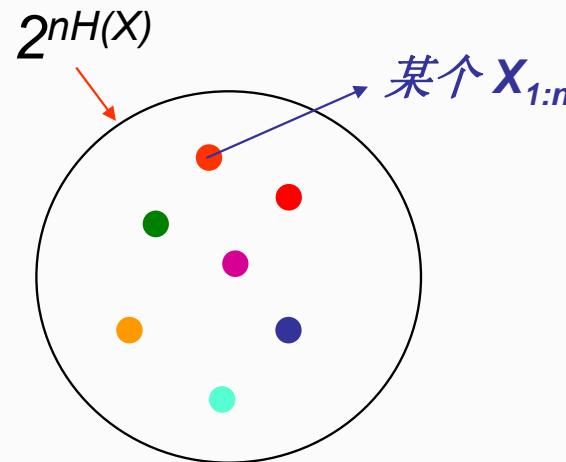


等价于直接随机生成一个码

- ① 产生典型集
- ② 构造码集合
- ③ 从码集合中随机选择一个码用于编码
- ④ 码字传输
- ⑤ 接收端译码

- ⑥ 计算译码误差概率：码集合中所有可能的码的平均错误概率的期望值
- ⑦ 进一步计算译码误差概率：基于联合典型译码规则（使用联合典型集、联合AEP等性质）
- ⑧ 从码集合中挑选出一个“好码”
- ⑨ 对挑选出的好码进行裁剪/删除

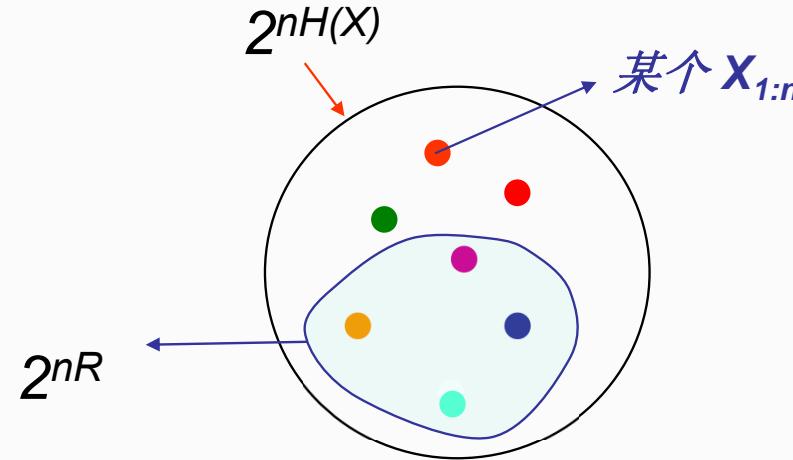
# (1) 产生典型集



按照信道的输入分布  $p_X$ , 产生长度为  $n$  的序列  $X_{1:n}$ ; 由 AEP 可知, 只要  $n$  足够大 ( $n > N_\varepsilon$ ), 产生的几乎都是典型序列, 构成信道的输入典型集; 由典型集性质, 输入典型集中约有  $2^{nH(X)}$  个元素。每个  $X_{1:n}$  的生成概率为

$$p(x_{1:n}) = \prod_{i=1}^n p(x_i)$$

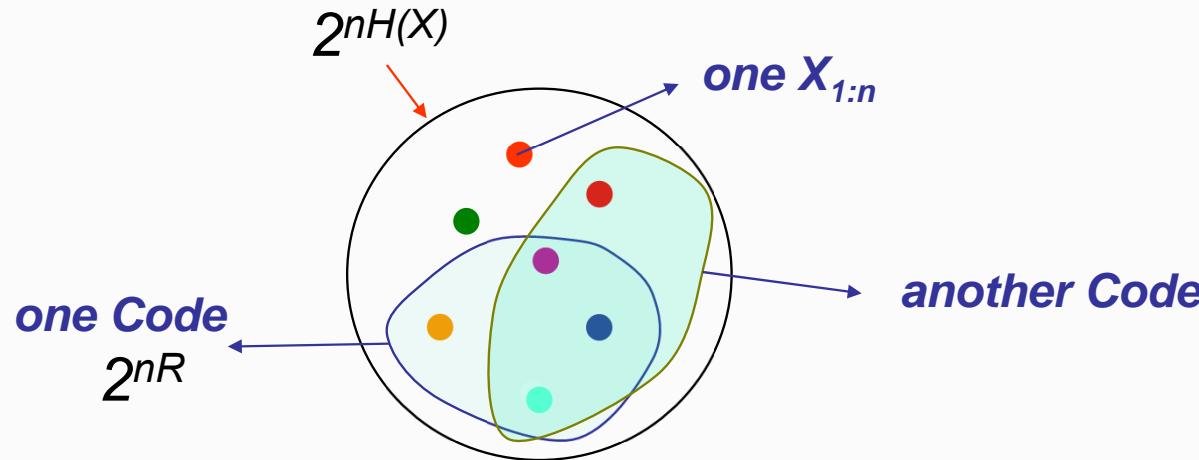
## (2) 构造码集合



每一个  $X_{1:n}$  都可以作为一个码字，从输入典型集中选取  $2^{nR}$  个  $X_{1:n}$  即可构成一个  $(2^{nR}, n)$  码（码簿），可用如下矩阵表示：

$$\mathcal{C} = \begin{bmatrix} x_{1:n}(1) \\ \vdots \\ x_{1:n}(2^{nR}) \end{bmatrix} = \begin{bmatrix} x_1(1) & x_2(1) & \cdots & x_n(1) \\ \vdots & \vdots & \ddots & \vdots \\ x_1(2^{nR}) & x_2(2^{nR}) & \cdots & x_n(2^{nR}) \end{bmatrix}$$

## (2) 构造码集合



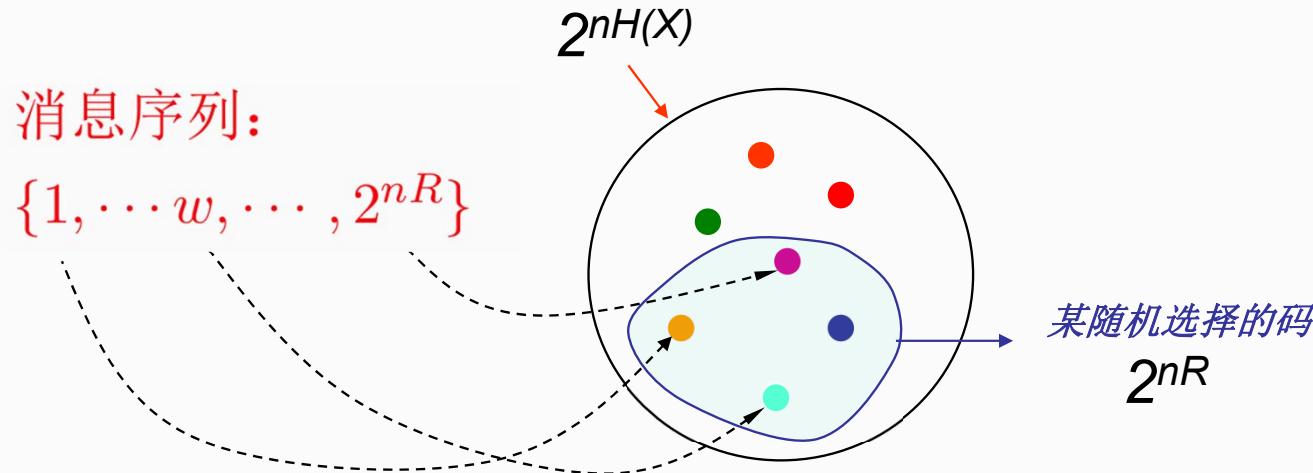
不同选取方式得到不同的码，所有可能的码构成码集合  $\mathcal{C}$ 。码集合  $\mathcal{C}$  的大小约为

$$|\mathcal{C}| = C_{2^{nH(X)}}^{2^{nR}}$$

码集合  $\mathcal{C}$  中某个特定码  $\mathcal{C}$  的生成概率为

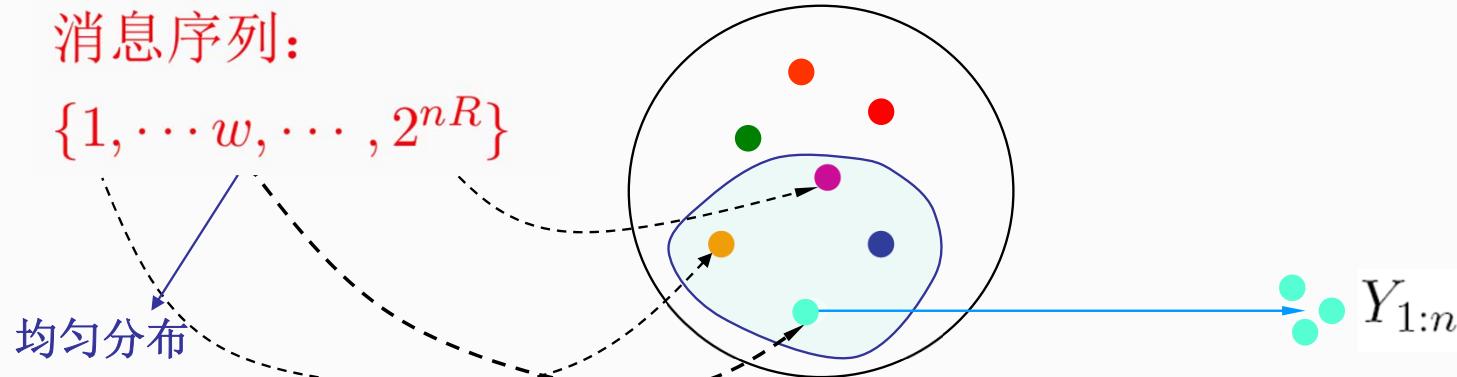
$$\Pr(\mathcal{C}) = \prod_{w=1}^{2^{nR}} \prod_{i=1}^n p(x_i(w))$$

### (3) 从码集中随机选择一个码用于编码



- 从码集  $\mathcal{C}$  中随机选择一个码  $\mathcal{C}$  用于信道编码——将待编码的  $2^{nR}$  个消息一一映射至  $\mathcal{C}$  中的各个码字。例如，第  $w$  个消息映射至  $\mathcal{C}$  的第  $w$  行。
- 接收端使用与发送端相同的码  $\mathcal{C}$  进行后续译码。同时，假设信道的概率转移矩阵  $\mathbf{Q}$  对于接收端也是已知的。

## (4) 码字传输



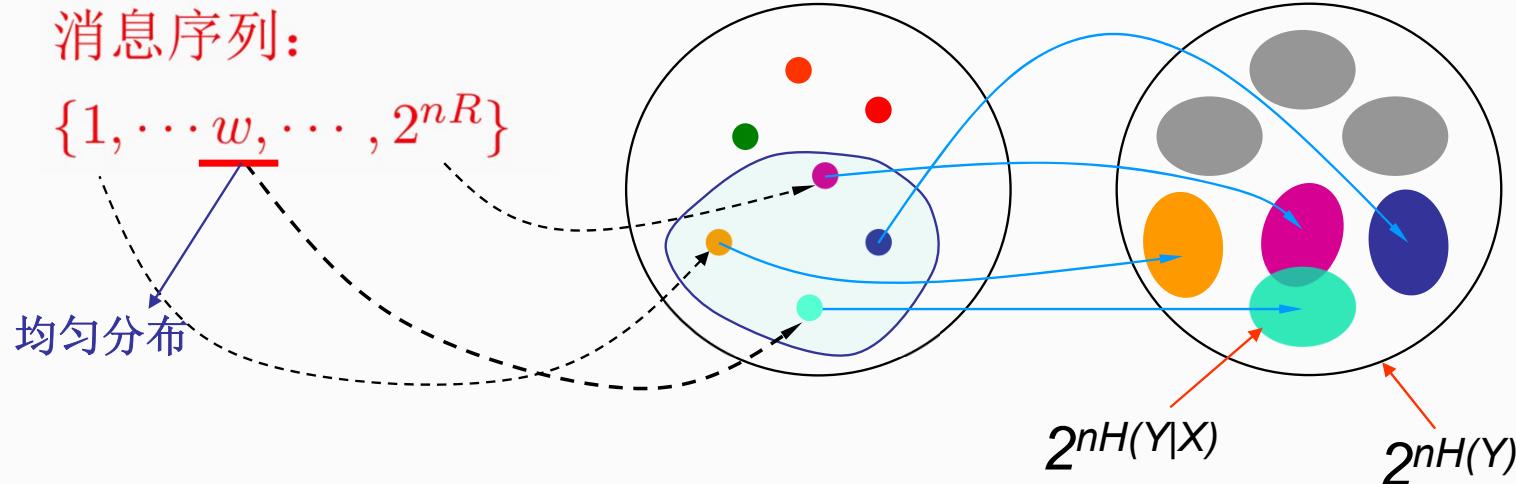
- 各消息服从均匀分布，即

$$\Pr(W = w) = 2^{-nR}, \quad w = 1, 2, \dots, 2^{nR}$$

- 相应的码字经过信道传输，按信道转移概率在接收端得到接收序列  $Y_{1:n}$ :

$$p(y_{1:n}|x_{1:n}(w)) = \prod_{i=1}^n p(y_i|x_i(w))$$

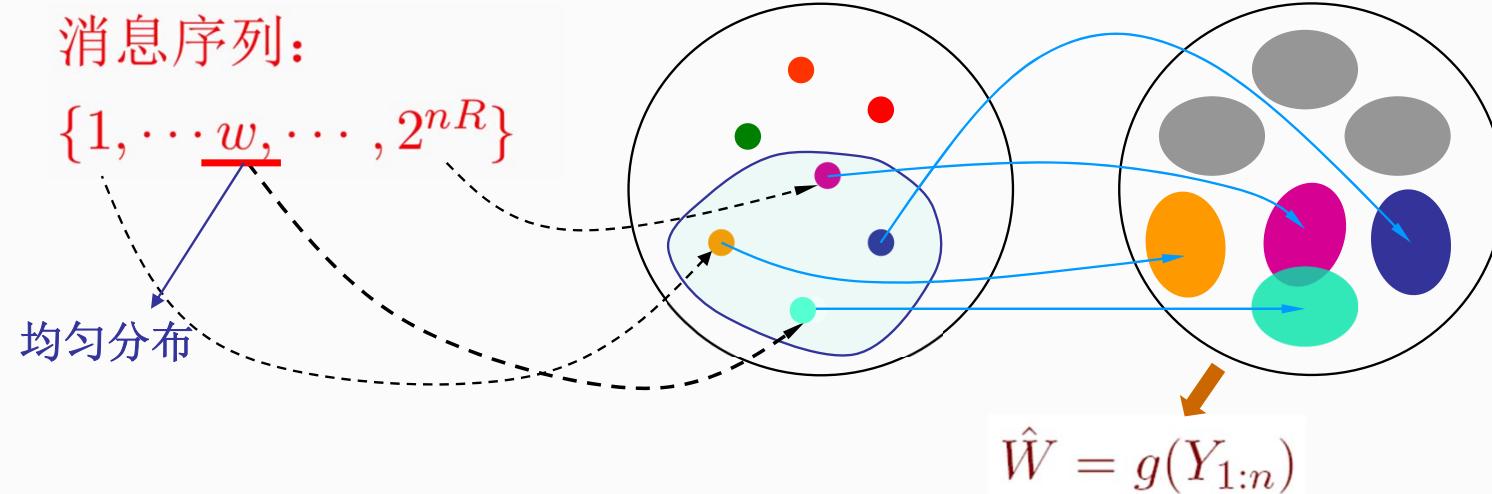
## (4) 码字传输



$n$  足够大的情况下:

- 接收序列  $Y_{1:n}$  几乎都会落在与发送码字  $X_{1:n}(w)$  相对应的输出典型集中；该输出典型集的大小约为  $2^{nH(Y|X)}$ 。
- 并且，不同码字对应的输出典型集之间几乎不交叠（相交的概率很小）。

## (5) 接收端译码



- 最佳译码方式是最大似然译码，但难以分析。此处我们采用联合典型译码，尽管只是次优，但分析简单，而且同样能达到证明目标（所有小于容量的码率都是可达的）。
- 用  $\mathcal{E}$  表示译码出错事件  $\hat{W} \neq W$ ，接下来计算译码误差概率  $\Pr(\mathcal{E})$ ：考虑码集合中所有可能的码，计算所有码的平均错误概率的期望值。

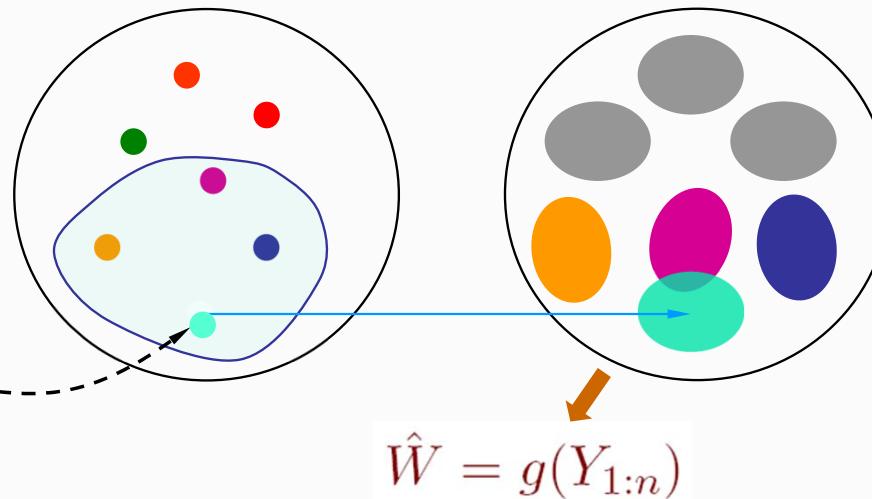
## (6) 计算译码误差概率

$$\begin{aligned}\Pr(\mathcal{E}) &= \sum_{\mathcal{C} \in \mathfrak{C}} p(\mathcal{C}) P_e^{(n)}(\mathcal{C}) \rightarrow \text{某一特定码的平均错误概率} \\ &\quad \text{遍历所有码} \\ &= \sum_{\mathcal{C} \in \mathfrak{C}} p(\mathcal{C}) \frac{1}{2^{nR}} \sum_{w=1}^{2^{nR}} \lambda_w(\mathcal{C}) \\ &= \frac{1}{2^{nR}} \sum_{w=1}^{2^{nR}} \sum_{\mathcal{C} \in \mathfrak{C}} p(\mathcal{C}) \lambda_w(\mathcal{C}) \\ &= \sum_{\mathcal{C} \in \mathfrak{C}} p(\mathcal{C}) \lambda_1(\mathcal{C}) \quad ? \\ &= \Pr(\mathcal{E} | W = 1)\end{aligned}$$

## (7) 进一步计算译码误差概率(联合典型译码)

消息序列:

$$\{1, \dots, w, \dots, 2^{nR}\}$$



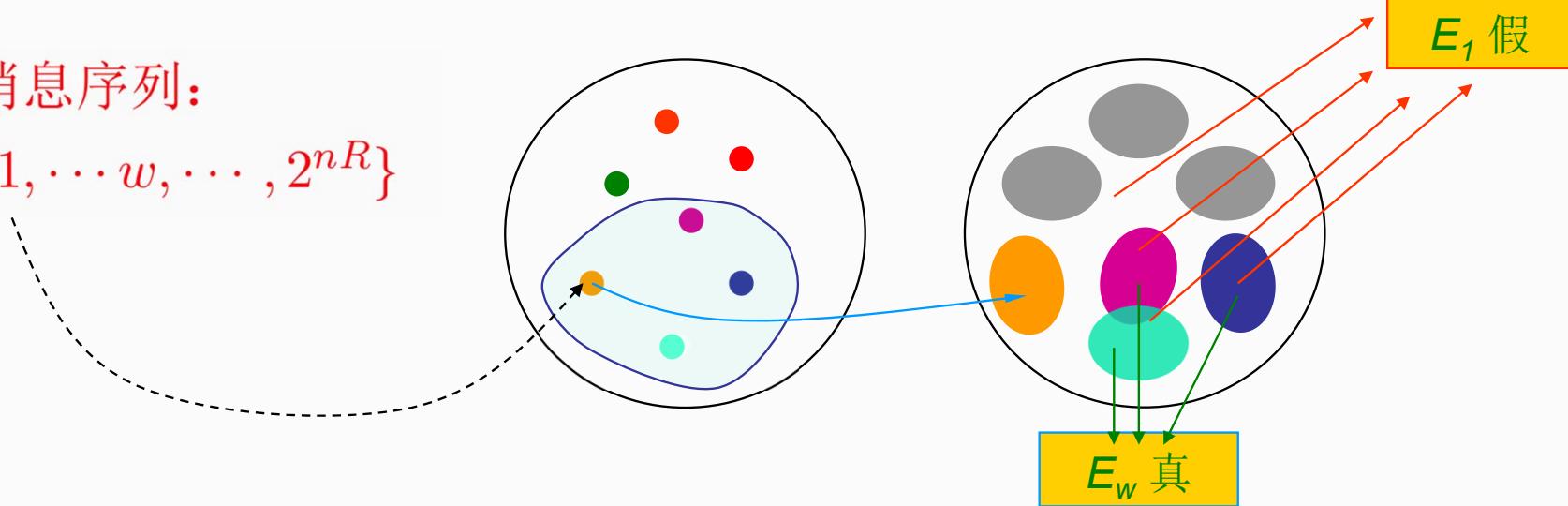
**【联合典型译码规则】**接收到序列  $Y_{1:n}$  后，查看码簿中有多少个码字  $X_{1:n}(W)$  与它联合典型：

- 若  $X_{1:n}(k)$  是唯一与  $Y_{1:n}$  联合典型的码字，则译码结果为  $k$ ；
- 若存在 0 个或大于等于 2 个可能的  $k$ ，则译码出错，可将译码结果赋值为消息集合中任意值，如 1 (或输出一个无效结果，如 0 )。

## (7) 进一步计算译码误差概率(联合典型译码)

消息序列:

$$\{1, \dots, w, \dots, 2^{nR}\}$$

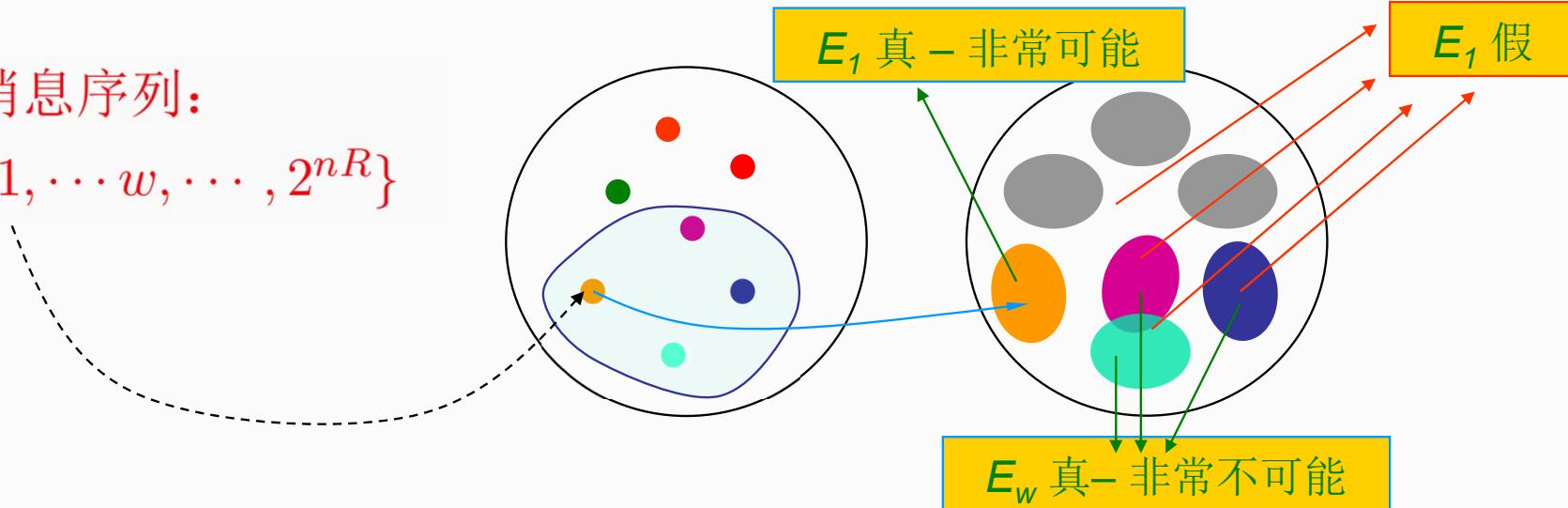


- 由于前面已经分析得到  $\Pr(\mathcal{E}) = \Pr(\mathcal{E}|W = 1)$ ，所以，译码误差概率的进一步分析、计算就可以以发送  $X_{1:n}(1)$  为例。
- 定义事件：
$$E_w = \{(X_{1:n}(w), Y_{1:n}) \in J_\varepsilon^{(n)}\}, \quad w \in 1 : 2^{nR}$$
- 译码出错对应的事件类型包括： $E_1$  为假，或者  $w \geq 2$  对应的某  $E_w$  为真。

## (7) 进一步计算译码误差概率(联合典型译码)

消息序列:

$$\{1, \dots, w, \dots, 2^{nR}\}$$



- 由联合典型集的性质 (2) (第 7 讲, 第 11 页), 对于  $n > N_\varepsilon$ , 有  $p(E_1) = \Pr((X_{1:n}(1), Y_{1:n}) \in J_\varepsilon^{(n)}) > 1 - \varepsilon$ , 所以: 对于  $n > N_\varepsilon$ ,  $p(\overline{E_1}) < \varepsilon$
- $w \neq 1$  所对应的码字  $X_{1:n}(w)$  与  $X_{1:n}(1)$  相互独立, 因此也与  $Y_{1:n}$  相互独立。由联合渐近均分性 (第 7 讲, 第 14 页), 有  $p(E_w) \leq 2^{-n(I(X,Y)-3\varepsilon)}$

## (7) 进一步计算译码误差概率(联合典型译码)

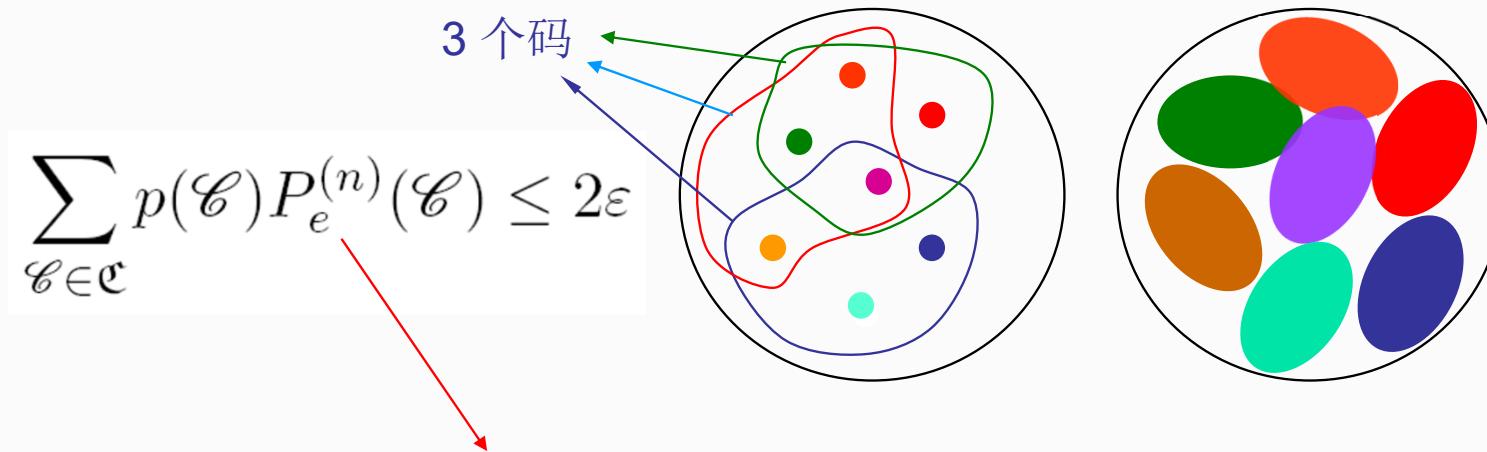


$$\Pr(\mathcal{E}|W=1) = p(\overline{E_1} \cup E_2 \cup E_3 \cup \dots \cup E_{2^{nR}})$$

$$\begin{aligned} p(A \cup B) &\leq p(A) + p(B) \\ &\leq \underbrace{p(\overline{E_1})}_{\text{联合典型性}} + \sum_{w=2}^{2^{nR}} \underbrace{p(E_w)}_{\text{联合渐近均分性}} \\ n > N_\varepsilon &\leq \varepsilon + \sum_{w=2}^{2^{nR}} 2^{-n(I(X,Y)-3\varepsilon)} \\ &= \varepsilon + (2^{nR} - 1) 2^{-n(I(X,Y)-3\varepsilon)} \\ &\leq \varepsilon + \underbrace{2^{-n(I(X,Y)-3\varepsilon-R)}}_{<\varepsilon} \\ &\leq 2\varepsilon \end{aligned}$$

for  $R < I(X, Y) - 3\varepsilon$  and  $n > -\frac{\log \varepsilon}{I(X;Y)-R-3\varepsilon}$

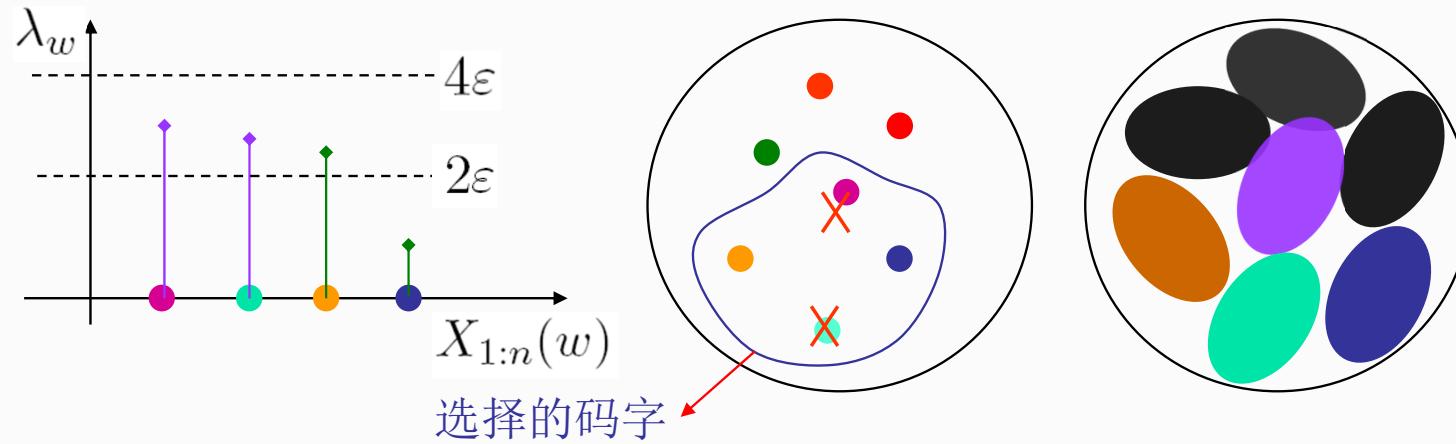
## (8) 从码集中挑选出一个“好码”



由于所有码的  $P_e^{(n)}$  的期望值  $\leq 2\varepsilon$ ，那么必然存在至少一个“好码”，其  $P_e^{(n)}$  确实就是  $\leq 2\varepsilon$ 。比如，所有码中  $P_e^{(n)}$  最小的那个码必然就满足：

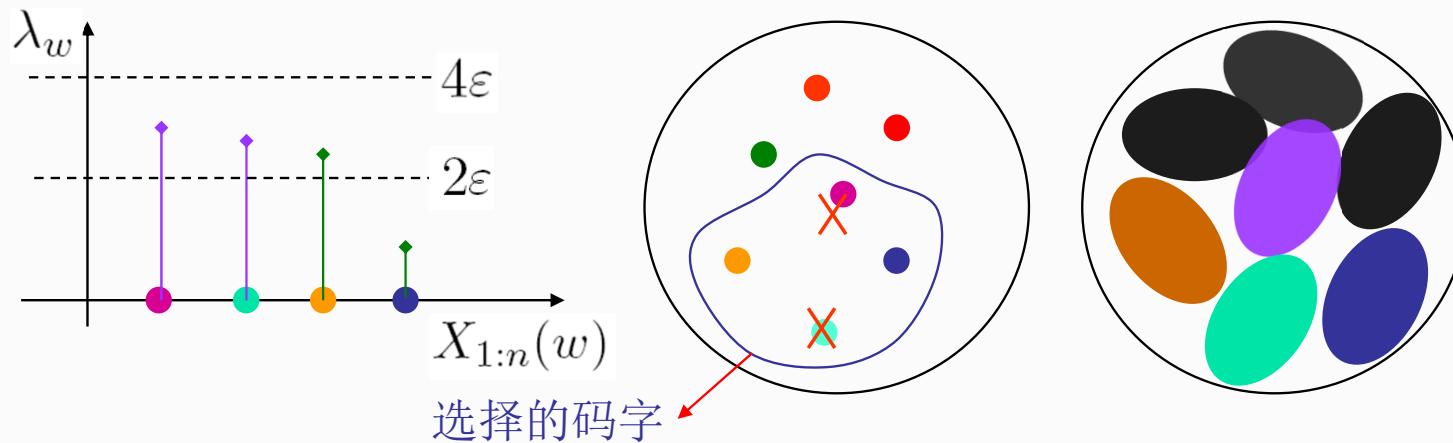
$$2\varepsilon \geq |\mathfrak{C}|^{-1} \sum_{i=1}^{|\mathfrak{C}|} P_{e,i}^{(n)} \geq |\mathfrak{C}|^{-1} \sum_{i=1}^{|\mathfrak{C}|} \min_i (P_{e,i}^{(n)}) = \min_i (P_{e,i}^{(n)})$$

## (9) 对挑选出的好码进行裁剪/删余



- 所选码的  $P_e^{(n)} = \frac{1}{M} \sum_{w=1}^M \lambda_w \leq 2\varepsilon$ ，其中  $M = 2^{nR}$ 。
- **删余**：剔除性能最差的一半码字，剩余的码字必然都满足  $\lambda_w \leq 4\varepsilon$ 。

## (9) 对挑选出的好码进行裁剪/删余



证明：假设  $\lambda_w$  按降序排列

$$2\epsilon \geq M^{-1} \sum_{w=1}^M \lambda_w \geq M^{-1} \sum_{w=1}^{M/2} \lambda_w \geq M^{-1} \sum_{w=1}^{M/2} \lambda_{M/2} \geq \frac{1}{2} \lambda_{M/2}$$
$$\Rightarrow \lambda_{M/2} \leq 4\epsilon \Rightarrow \lambda_w \leq 4\epsilon, \forall w > M/2$$

剩余的（好的那一半码字）所组成的码：

$$M' = \frac{1}{2} \times 2^{nR} \Rightarrow R' = \frac{\log M'}{n} = R - \frac{1}{n} > R - \epsilon, \text{ 对于 } n > \frac{1}{\epsilon}$$

# 可达性证明过程小结

- 信道输入分布  $p_X$  可在一开始就设置为最佳分布，即使得  $I(X; Y) = C$  的分布，前述证明中所有  $I(X; Y)$  就都可以替换为  $C$ 。
- 对于任意  $R < C - 3\varepsilon$ ，令

$$n > \max \left\{ N_\varepsilon, -\frac{\log \varepsilon}{C - R - 3\varepsilon}, \frac{1}{\varepsilon} \right\}$$

- 按信道输入分布  $p_X$  生成序列长度为  $n$  的输入典型集，然后利用典型集中的元素构造一系列  $(2^{nR}, n)$  码。随机选择一个码进行编码，并在接收端使用联合典型译码。

- 由于所有码的  $P_e^{(n)}$  期望值  $\leq 2\varepsilon$ ，那么必然存在至少一个好码满足此条件，理论上通过穷举搜索即可找到。
- 剔除最差的一半码字。此时，剩余的那一半码字的最大误差概率  $\leq 4\varepsilon$ ，码率  $R' = R - \frac{1}{n} > R - \varepsilon$ 。
- 所得的码即为满足证明目标的码：以码率  $R'$  传输，其误差概率可以尽可能小。

## 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  $n \rightarrow \infty$  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  $n \rightarrow \infty$ ）的  $(2^{nR}, n)$  码，一定满足  $R \leq C$ 。



- 首先考虑零误差码：

$$\lambda^{(n)} = 0 \Rightarrow R \leq C$$

- 将证明扩展到任意小误差码：

$$\lambda^{(n)} \rightarrow 0 \Rightarrow R \leq C$$



$$\hat{W} = g(Y_{1:n}) = W$$

## 引理 7.9.2

若离散无记忆信道的容量为  $C$ , 则该信道的  $n$  次使用容量为  $nC$ , 即

$$I(X_{1:n}; Y_{1:n}) \leq nC$$

——证明直接来自于“第 6 讲第 11 页”结果, 即  $I(X_{1:n}; Y_{1:n}) \leq \sum_{i=1}^n I(X_i; Y_i)$



$$\hat{W} = g(Y_{1:n}) = W$$

假设  $W$  均匀分布

$$\begin{aligned}
 nR &= H(W) = \cancel{H(W|Y_{1:n})} + I(W; Y_{1:n}) \\
 &= I(W; Y_{1:n}) \\
 &\leq I(X_{1:n}; Y_{1:n}) \xleftarrow{\text{Markov: } W \rightarrow X_{1:n}(W) \rightarrow Y_{1:n}} \\
 &\leq nC \quad \xleftarrow{\text{引理 7.9.2}}
 \end{aligned}$$

# 任意小误差码



- 将Fano 不等式 ( $H(X|Y) \leq 1 + P_e \log |\mathcal{X}|$ ) 应用于当前情况，可得到

$$H(W|Y_{1:n}) \leq 1 + P_e^{(n)} \log |\mathcal{W}| = 1 + nR P_e^{(n)}$$

- 因此

$$\begin{aligned} nR &= H(W) = H(W|Y_{1:n}) + I(W; Y_{1:n}) \\ &\leq 1 + nR P_e^{(n)} + nC \quad \xrightarrow[n \rightarrow \infty]{} 0 \\ \Rightarrow R &\leq C + \frac{1}{n} + R P_e^{(n)} \quad \xrightarrow[n \rightarrow \infty]{} R \leq C \end{aligned}$$

# 任意小误差码

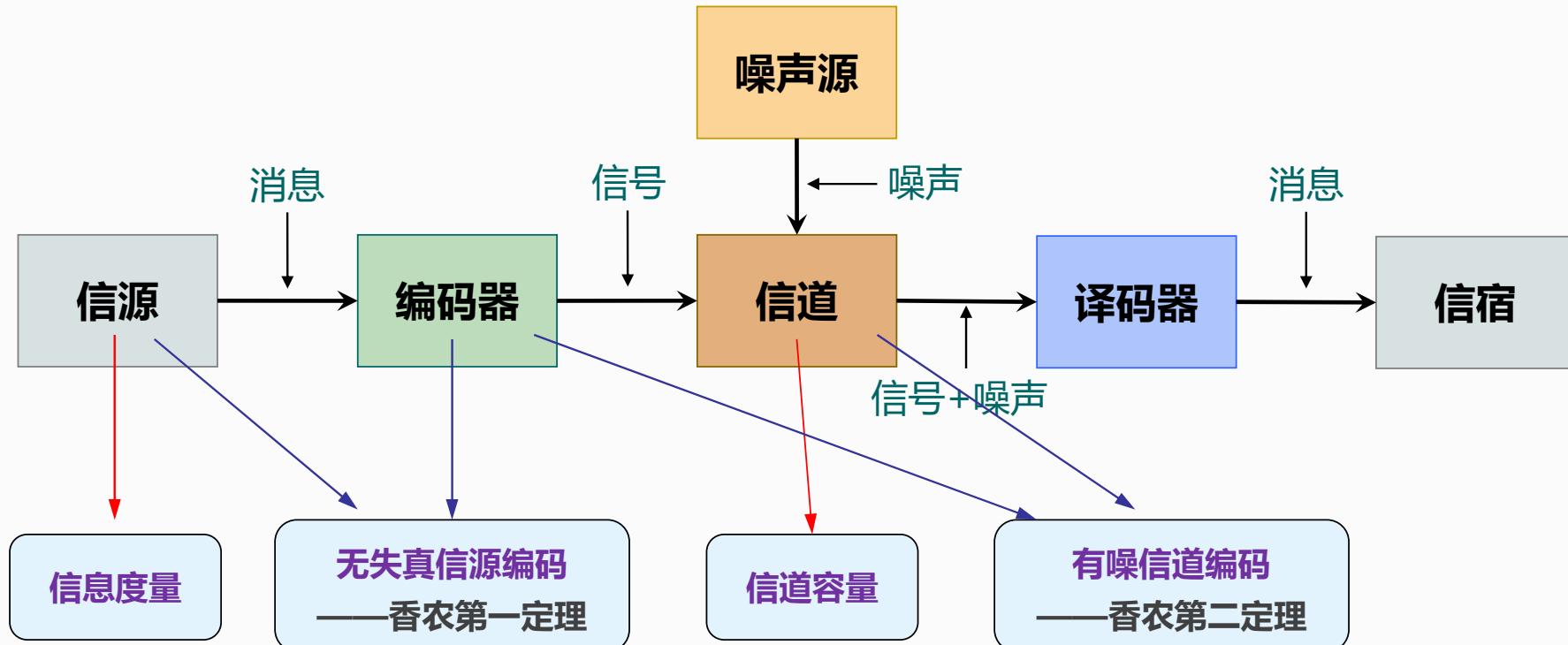


将结果变形为

$$P_e^{(n)} \geq \frac{R - C - n^{-1}}{R} \xrightarrow{n \rightarrow \infty} \frac{R - C}{R}$$

- 若  $R > C$ ，对于足够大的  $n$  (进而对于所有的  $n$ )，误差概率会远离 0。
- 在高于信道容量的码率下，我们无法实现任意低的误差概率。

# 课程内容进度安排

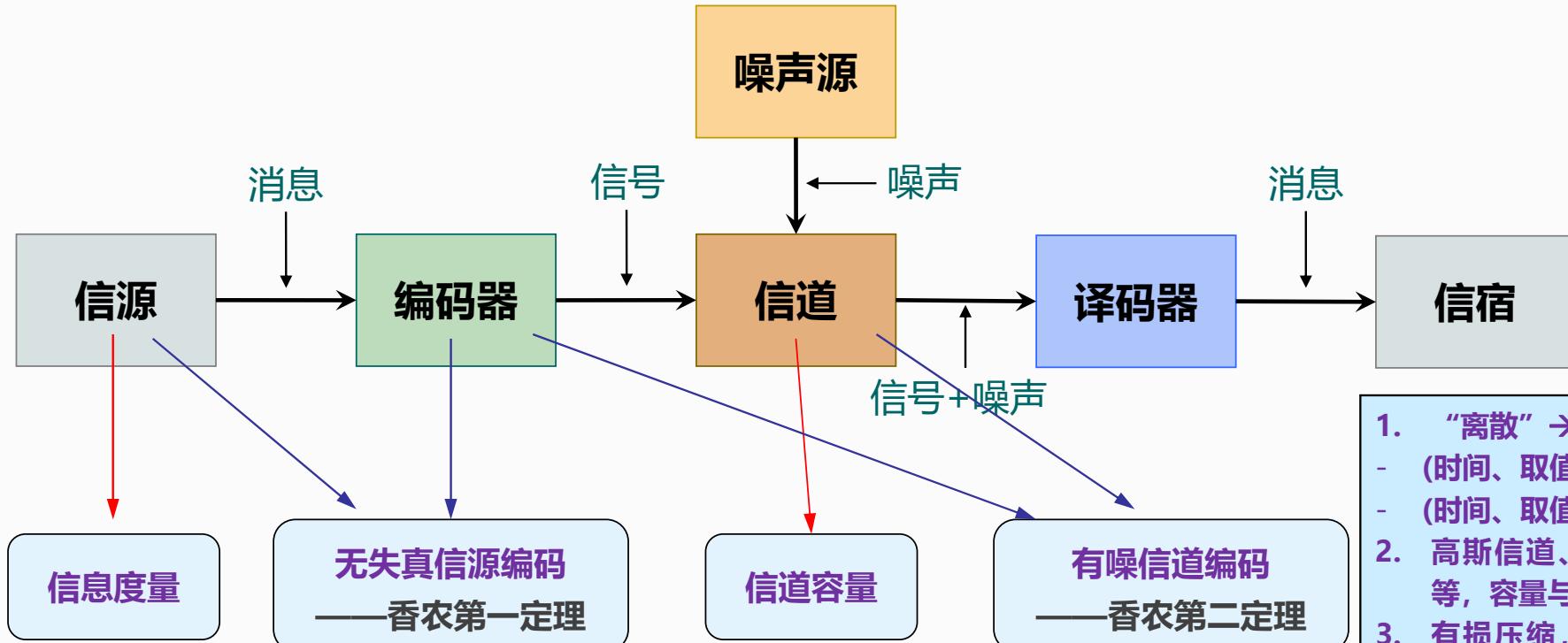


课程内容学习顺序



作为导论课，本课程只讨论“离散”信源、信道

# 以导论课为起点，进一步延伸学习



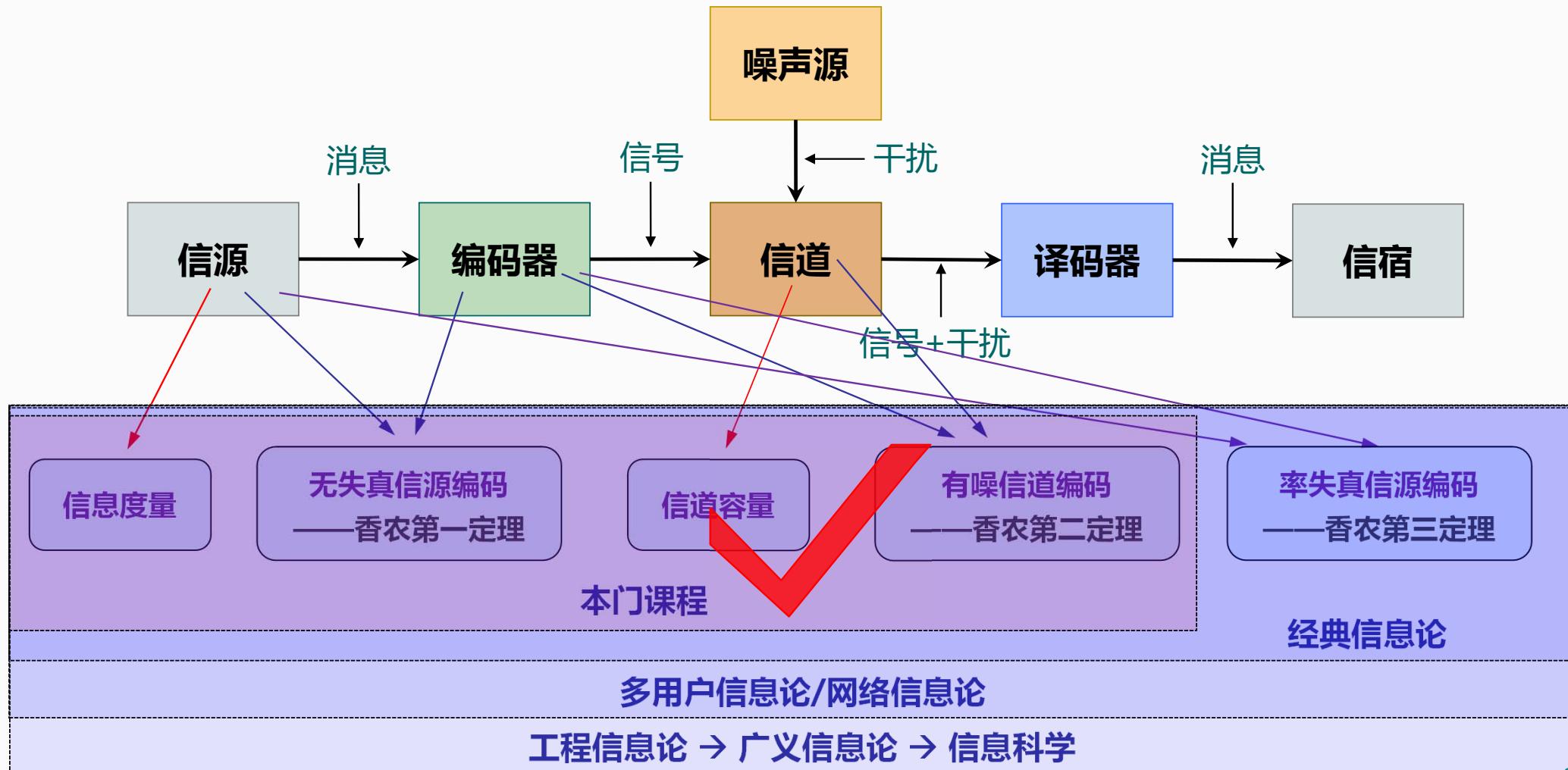
课程内容学习顺序



作为导论课，本课程只讨论“离散”信源、信道

1. “离散” → “连续”
  - (时间、取值)连续信源
  - (时间、取值)连续信道
2. 高斯信道、并行高斯信道等，容量与编码定理
3. 有损压缩，率失真定理（香农第三定理）
4. 数据（文本、图像、视频.....）压缩算法
5. 差错控制编码算法
6. .....

# 回顾：我们可以从这门课中学到什么？



# 回顾：为什么叫“导论”课？



- **大二下学期（承上启下时间点），需要有这么一门课**
- **通过此门课：**为后续信息类学科（通信、计算机……）知识的学习奠定基础。
- **16学时课程、不以“全面、系统、深入”为目标，期望结课时：**
  - 能有如下认识：1) 信息论这门课就是概率（随机变量/过程、期望……）、统计（弱大数定律……）的延伸；2) 信息论的典型研究范式是“极限 + 逼近极限的方法”，如“ $H(X)$ ,  $L \rightarrow H(X)$ ”、“ $C$ ,  $R \rightarrow C$ ”……
  - 能产生对信息基础理论的兴趣，理解一些基本概念、基本结论，为今后科研中可能碰到的一些基础问题留下线索……
  - 能感受到经典信息理论的简洁、优美，并以此为目标设立自己今后的学术、科研“质量标准”

# 关于期末课程报告



可在如下类型课程报告中任选一类：

- 关于经典信息论的局限性的思考 —— 副标题
- 信息论在信息学科前沿领域的研究与应用 —— 副标题
- 信息论在我#####竞赛/课题中的角色 —— 副标题
- 关于信息论课程中 “#####” 问题的思考 (注：可能会在课堂上留下一些开放性思考题)

报告形式与要求：

- ① 模板统一（由助教提供）
- ② 篇幅不限(建议主体内容1500字~3000字)，要有摘要、结论、参考文献

报告提交时间、提交方式：

- 提交时间：2025.6.19, 24:00之前，逾期不收（以助教邮箱接收时间戳为准）
- 提交方式：电子版报告附上个人电子签名，邮箱发送至助教邮箱，以收到助教邮件回复为“报告成功提交”确认依据



哈爾濱工業大學(深圳)

HARBIN INSTITUTE OF TECHNOLOGY, SHENZHEN

# 结束