论失配无认知中继

Michael Dikshtein、Nir Weinberger 和 Shlomo Shamai (Shitz)，Technion 电气与计算机工程系，海法 3200003，以色列 电子邮箱： {michaeldic@、nirwein@、sshlomo@ee.}technion.ac.il

arXiv:2305.00930v1 [cs.IT] 1 May 2023

摘要——我们研究借助中继通过离散无记忆信道（DMC）实现可靠通信的问题，并将其称为信息瓶颈（IB）信道。信源与目的端之间没有直接链路，信息通过两跳传输。第一跳是从信源到中继的噪声信道。第二跳是从中继到解码器的一条无噪声但容量受限的回传链路。我们进一步假设中继对传输码本无认知。我们考察两种失配场景。在第一种设置中，假设解码器受限于采用某一固定译码规则，而该规则与实际信道失配。在第二种设置中，假设中继受限于采用某一固定压缩度量，该度量同样与中继输入的统计特性失配。我们为这两种设置的随机码容量建立了界，其中一些界被证明是系综紧致的。

I. 引言

本文考虑点对点无感知中继信道 [1], ，其中中继通过噪声信道观测传输信号，并通过容量受限的链路向解码器传输数字信息。本文主要关注各种失配条件下，特别是采用随机码时的可达率结果。我们既考虑接收方的失配译码，也考虑中继的失配压缩。开展这项分析的动机在于，采用中继无感知处理的网络架构是现代通信系统的基本构建模块。关于无感知通信网络的最新全面综述可参见[2]。分析失配情形的动机在于，许多实际场景中（例如采用无感知中继的蜂窝通信上行链路），中继或解码器掌握的信道统计模型信息有限，或者只能使用某种特定的解码度量（通常会在更复杂的信道译码算法中使用该度量，例如置信传播 [3]）。

无认知中继信道的容量与信息瓶颈（IB）问题密切相关 [4], 。近来，该问题受到了广泛研究，主要是因为它与机器学习的最新进展有关，参见例如 [5], [6], 以及统计学习 [7]。在有损信源编码的经典率失真理论中，必须选择一种保真度量来量化压缩的质量[8]。IB 方法通过引入另一个相关随机变量来确定这种失真度量，该变量捕捉待压缩数据中的有意义信息，可以将其视为数据的上下文标记。具体而言，IB 方法通过压缩表示与该附加变量之间的互信息来评估压缩质量。结果表明，所得的 IB 恰好与所研究的无认知中继信道的容量相匹配。实际上，它也是远程信源编码情形下的单字母率失真公式 [9], [10], 其中失真度量为对数损失 [11]。文献 [12] 表明，对于二元分类问题，最小化对数损失可以最小化任意损失函数选择所对应的一个上界。还可以证明，对数损失实际上界定了一般失真度量（正如 Linder 所述，但据我们所知，该结果尚未发表）。

众所周知，失配解码问题极具挑战性，即使对于标准的点对点信道 [13], 也尚未得到彻底解决。尽管如此，对失配编码或失配解码下的随机码进行分析，可以得到易于处理的可达界，因此本文对所研究的无知中继信道（或信息瓶颈问题）采用这种分析方法。在随机编码方案下，采用失配解码器对离散无记忆信道进行编码，是在 [14] 和 [15], 中独立提出的；这两篇文献推导出了容量的一个下界（称为 LM 速率）。文献 [16] 表明，LM 界并不紧。文献 [17], 提出了一个更易于分析的下界，即广义互信息（GMI）；文献 [18]则证明了 LM 方案的随机编码系综紧致性。在我们的场景中，分析随机码还有进一步的动机：我们希望对不针对特定通信场景进行适配的码进行建模；此外，中继具有无认知特性，而中继通信系统通常还涉及安全因素。

本文其余部分的结构和我们的贡献如下。在第 II 节中，我们考虑接收方使用失配译码器时的信息瓶颈问题，并推导随机编码可达率，既包括 LM 界形式的可达率，也包括广义互信息界。随后，我们提出一种计算可达率的算法，并以四元信道为例展示其运行过程。接着，我们将广义互信息速率推广到连续字母表信道，并通过高斯衰落信道展示这一结果。之后，在第 III 节中，我们考虑采用失配压缩规则的中继场景。第 IV 节总结全文。

相关工作： [13] 对失配译码和编码的信息论基础进行了全面总结。除信道编码外，信源编码领域也研究了失配问题 [19], [20]。 [21] 提出了一种受限于高斯码本并采用最小欧几里得距离编码的渐进细化设置。近期， [22] 将其扩展到一般字母表。 [23], 研究了目的端具有全局信道知识的量化分布式接收设置，其中假设采用简单的复数二元符号量化。一个分布式解码通信网络

考虑了在加性高斯白噪声信道上采用 BPSK 传输的情形 [24]。 [25] 研究了基于正交频分多址的云无线接入网中的均匀量化。 [26] 研究了硬判决交换的分布式接收问题中的中断概率。

II. 具有失配解码器的信息瓶颈信道

A. 离散无记忆信道

本节考虑图 1 所示的带中继三节点点对点通信系统，其中发送方希望借助中继向接收方传递消息 M。我们将这种设置称为具有失配解码器 $\langle \mathcal{X}、\mathsf{P}^{\mathsf{Y}|\mathsf{X}}, \mathcal{Y}, \mathcal{Z}, V(z|x)\rangle$ 的离散无记忆信息瓶颈信道（DM-IBC）。该信道由三个有限集合 $\mathcal{X}$, $\mathcal{Y}$, $\mathcal{Z}$、定义在 $\mathcal{Y}$ 上的一组条件概率质量函数 $\mathsf{P}^{\mathsf{Y}|\mathsf{X}}$ （每个输入符号 x 对应一个），以及定义在 $\mathcal{Z}$ 上的解码度量 $V(z|x)$ 组成。

具有失配解码器的 DM-IBC 的 $(2^{nR}, 2^{nB}, n)$ 码由以下部分组成:

- 一个消息集 $\mathcal{M} = [1:\ 2^{nR}]$,
- 一个表示集 $\mathcal{W} = [1:\ 2^{nB}]$,
- 一个编码器，为每条消息 $m \in \mathcal{M}$指定一个码字 $x^n(m)$,
- 一个中继源编码器，为每个接收序列 $y^n \in \mathcal{Y}^n$指定一个索引 $w \in \mathcal{W}$，并对应重构序列 $z^n(w)$，以及
- 一个解码器，根据某个固定的失配度量，为每个接收表示索引 $w \in \mathcal{W}$ 指定一个估计值 $\hat{m}$ 或一条错误消息 e

$$\begin{aligned}\hat{m}(z^n) &= \arg\max_{m\in\mathcal{M}} V(z^n|x^n(m)) \\ &= \arg\max_{m\in\mathcal{M}} \prod_{i=1}^{n} V(z_i|x_i(m)). \quad (1)\end{aligned}$$

假设消息 M 在消息集 $\mathcal{M}$ 上均匀分布。

定义 2.1：考虑一个由 2^{nR} n 维序列构成的码本，$\{x^n(m)\}_{m=1}^{2^{nR}}$，其中每个序列均按照无记忆概率质量函数 P_{X} 随机生成，且与所有其他向量相互独立。再考虑一个由 2^{nB} n 维序列构成的压缩码本，$\{z^n(w)\}_{w=1}^{2^{nB}}$，其中每个序列均按照无记忆概率质量函数 P_{Z}随机生成，且与所有其他向量相互独立。这样一对信道-压缩码本称为随机码本。令 $P_e^{(n)} = \mathrm{P}\{\hat{\mathsf{M}} \neq \mathsf{M}\}$ 表示对随机码本取平均后的错误概率。若在压缩速率为 B、失配译码度量为 V的情况下，满足失配译码规则下的 $\lim_{n\to\infty} P_e^{(n)} = 0$，则称码率 R 对 DM-IBC 可达。DM-IBC 的随机编码容量 $C_V(B)$

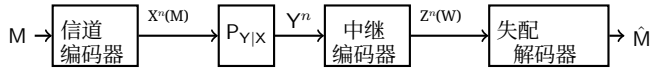

图 1：采用失配解码器的盲通信系统

采用失配译码器的 DM-IBC 的随机编码容量，是所有随机码本可达码率的上确界。

本节的主要结果由下述定理给出，它描述了无知中继设置下相应的 LM 速率 [14], [15] 。

定理 1：采用失配解码器的 DM-IBC 的随机编码容量为

$$\begin{aligned}C_V(B) = \quad & \max_{\mathsf{P}_{\mathsf{X}}} \max_{\mathsf{P}_{\mathsf{Z}|\mathsf{Y}}} \quad J_V(\mathsf{X};\mathsf{Z}) \\ & \text{subject to} \quad I(\mathsf{Y};\mathsf{Z}) \le B,\end{aligned} \tag{2}$$

其中

$$\begin{aligned}J_V(\mathsf{X};\mathsf{Z}) = \quad & \min_{\mathsf{Q}_{\mathsf{Z}|\mathsf{X}}\in\mathcal{P}_{\mathcal{X}\times\mathcal{Z}}} \quad I(\mathsf{P}_{\mathsf{X}}, \mathsf{Q}_{\mathsf{Z}|\mathsf{X}}) \\ & \text{subject to} \quad \sum_{x\in\mathcal{X}} \mathsf{Q}(x,z) = \mathsf{P}_{\mathsf{Z}}(z), \\ & \qquad\qquad \sum_{(x,z)\in\mathcal{X}\times\mathcal{Z}} \mathsf{Q}(x,z)\log V(z|x) \ge -D,\end{aligned} \tag{3}$$

with $\mathsf{Q}(x,z) = \mathsf{Q}_{\mathsf{Z}|\mathsf{X}}(z|x)\cdot\mathsf{P}_{\mathsf{X}}(x)$ and

$$D \triangleq -\sum_{x\in\mathcal{X}}\sum_{z\in\mathcal{Z}} \mathsf{P}_{\mathsf{XZ}}(x,z)\log V(z|x). \tag{4}$$

定理 1 给出了失配译码下的精确随机编码容量。具体而言，我们证明了一个上界和系综紧致性。

证明：直接部分的证明见附录 B，系综紧致性的证明见附录 C。 ■

定理 1 所得的容量表达式是在离散情形下证明的，不易扩展到连续信道。因此，下面我们基于广义互信息（GMI）编码方案 [17]给出一个可达随机编码速率。正如我们将看到的，该码率可以扩展到连续信道。请注意，定理 1 中的 LM 速率与下述定理给出的 GMI 速率之间的主要区别，在于放宽了边缘分布相等约束，即 Q_{X} =P_{X}。

定理 2：带有失配解码器的 DM-IBC 的随机编码容量的下界为

$$\begin{aligned}C_V(B) \ge \quad & \max_{\mathsf{P}_{\mathsf{X}},\mathsf{P}_{\mathsf{Z}|\mathsf{Y}}} \quad I_{\mathsf{GMI}}(\mathsf{P}_{\mathsf{XZ}}) \\ & \text{subject to} \quad I(\mathsf{Y};\mathsf{Z}) \le B,\end{aligned} \tag{5}$$

其中

$$\begin{aligned}I_{\mathsf{GMI}}(\mathsf{P}_{\mathsf{XZ}}) = \quad & \min_{\tilde{\mathsf{P}}_{\mathsf{XZ}}} \quad \mathsf{D}\left(\tilde{\mathsf{P}}_{\mathsf{XZ}} \| \mathsf{P}_{\mathsf{X}}\times\mathsf{P}_{\mathsf{Z}}\right) \\ & \text{s.t.} \quad \sum_{x\in\mathcal{X}} \tilde{\mathsf{P}}_{\mathsf{XZ}}(x,z) = \mathsf{P}_{\mathsf{Z}}(z) \\ & \qquad \mathbb{E}_{\tilde{\mathsf{P}}_{\mathsf{XZ}}}\left[V(\mathsf{Z}|\mathsf{X})\right] \ge \mathbb{E}_{\mathsf{P}_{\mathsf{XZ}}}\left[V(\mathsf{Z}|\mathsf{X})\right].\end{aligned} \tag{6}$$

此外，内层最小化问题具有如下对偶形式,

$$I_{\mathsf{GMI}}(\mathsf{P}_{\mathsf{XZ}}) = \max_{\lambda\ge 0} \sum_{(x,z)\in\mathcal{X}\times\mathcal{Z}} \mathsf{P}_{\mathsf{XZ}}(x,z)\log\frac{V(z|x)^\lambda}{\sum_{x'}\mathsf{P}_{\mathsf{X}}(x')V(z|x')^\lambda}. \tag{7}$$

Proof: The proof appears in App. D. ■

B. 计算高效算法

本节中，我们提出一种高效算法，通过求解定理 1 中的优化问题来计算$C_V(B)$。为此，需要注意的是，在 (3) 中计算 $J_V(\mathrm{X};\mathrm{Z})$ 的内层最小化问题是关于 $\mathrm{Q}_{\mathrm{Z}|\mathrm{X}}$ 的凸优化问题，因为互信息是信道的凸函数，而对 $\mathrm{Q}_{\mathrm{Z}|\mathrm{X}}$ 的约束条件是线性的。因此，可以使用标准凸优化求解器高效地计算 $J_V(\mathrm{X};\mathrm{Z})$。

相比之下，式 (2) 中关于 $\mathrm{P}_{\mathrm{Z}|\mathrm{Y}}$ 的外部最大化问题并非凹问题，求解它需要全局优化方法，例如网格搜索。具体而言，我们建议先在概率单纯形的粗网格上计算该最大值。然后，仅在粗略最大化所得解的局部邻域内，通过更细的网格对最大值进行精细搜索。以迭代方式重复这一细化过程，可使每一步都得到更优的解；最后一次迭代中可达率的增幅若可忽略不计，则可作为停止准则。尽管如此，我们强调，任意选择 $\mathrm{P}_{\mathrm{Z}|\mathrm{Y}}$ 都能得到码率的一个可达下界。因此，确保解有效性的关键优化步骤是式 (3) 中的凸最小化步骤。此外，还可以使用搜索方法优化输入分布 P_{X}。然而，在许多实际应用中，输入分布是任意选取的，例如采用均匀分布，这通常是基于问题的对称性。

所提出的算法包含一个交替最大化步骤，该步骤最初是为原始信息瓶颈问题提出的 [4] ，用于在解码器匹配的情况下寻找最优测试信道。众所周知，该算法基于 Blahut-Arimoto [27], [28] 算法，在此称为信息瓶颈交替最小化（IBAM）。其正式描述可见例如 [4, 定理 5]。我们的主要算法详见算法1。

C. 示例：四元信道

本节通过一个简单的场景展示定理 1 的结果以及算法 1的运行过程。具体而言，假设从 X 到 Y 的信道由以下条件概率质量函数定义：

$$\mathrm{P}_{\mathrm{Y}|\mathrm{X}}(y|x)=\mathrm{P}\{\mathrm{Y}=y|\mathrm{X}=x\}=\begin{cases}1-\epsilon, & y=x\\ \frac{\epsilon}{2}, & y=x\pm 1\\ 0, & \text{otherwise},\end{cases}\quad(8)$$

其中，所有加减运算均按模4计算。度量 V 是失配的，具体而言，它与另一种信道相匹配。该信道具有相同的正确符号转移概率 $1-\epsilon$，但错误可能以概率$\frac{\epsilon}{3}$ 转移到其他三个备选符号中的任意一个，即：

$$\mathrm{Q}_{\mathrm{Y}|\mathrm{X}}(y|x)=\mathrm{P}\{\mathrm{Y}=y|\mathrm{X}=x\}=\begin{cases}1-\epsilon, & y=x\\ \frac{\epsilon}{3}, & y\neq x.\end{cases}\quad(9)$$

由于对称性，我们假设最优的 P_{X} 是均匀分布，且 $\mathrm{P}_{\mathrm{Z}|\mathrm{Y}}$ 是对称的。在这种情况下，$\mathrm{P}_{\mathrm{Z}|\mathrm{Y}}(z|y)$ 是一个模加信道。我们比较采用匹配

算法1：MMIB 算法

输入：$\mathrm{P}_{\mathrm{XY}}, V(z|x)$ B, RES

$\mathrm{P}^{opt}_{\mathrm{Z}|\mathrm{Y}}=\mathrm{IBAM}(\mathrm{P}_{\mathrm{XY}}, C, args)$

$\mathcal{P}_{\mathcal{Z}|\mathcal{Y}}=\{\mathrm{P}_{\mathrm{Z}|\mathrm{Y}}\colon\ I(\mathrm{Y};\mathrm{Z})=B\}$

$\mathcal{P}^{\mathrm{coarse}}_{\mathcal{Z}|\mathcal{Y}}=\mathcal{P}_{\mathcal{Z}|\mathcal{Y}}\cap \mathrm{RES}\cdot\mathbb{Z}^{|\mathcal{Z}|}$

对于 $\mathrm{P}_{\mathrm{Z}|\mathrm{Y}}\in\mathcal{P}^{coarse}_{\mathcal{Z}|\mathcal{Y}}$，执行 do

- 根据 (4) 计算 D，其中 $\mathrm{P}=\mathrm{P}_{\mathrm{Z}|\mathrm{Y}}\cdot\mathrm{P}_{\mathrm{XY}}$
- $J_V(\mathrm{X};\mathrm{Z})=\min_Q I(\mathrm{P}_{\mathrm{X}},\mathrm{Q}_{\mathrm{Z}|\mathrm{X}})$
- 满足：
- $\mathbb{E}_{\mathrm{P}_Q}[\log V(\mathrm{Z}|\mathrm{X})]\geq -D$

求使 $J_V(\mathrm{X};\mathrm{Z})$ 最大化的 $\mathrm{P}^*_{\mathrm{Z}|\mathrm{Y}}\in\mathcal{P}^{\mathrm{coarse}}_{\mathcal{Z}|\mathcal{Y}}$。

当 $\Delta R\geq\epsilon$ 时执行

- $\mathrm{RES}=\mathrm{RES}*\mathrm{FINE}$
- $\mathcal{P}^{\mathrm{fine}}_{\mathcal{Z}|\mathcal{Y}}=\mathcal{P}_{\mathcal{Z}|\mathcal{Y}}\cap\mathrm{FINE}\cdot\mathbb{Z}^{|\mathcal{Z}|}\cap\mathcal{B}(\mathrm{P}^*_{\mathrm{Z}|\mathrm{Y}},1/\mathrm{RES})$
- 对于 $\mathrm{P}_{\mathrm{Z}|\mathrm{Y}}\in\mathcal{P}^{fine}_{\mathcal{Z}|\mathcal{Y}}$ **do**
  - 根据 (4) 计算 D，其中 $\mathrm{P}=\mathrm{P}_{\mathrm{Z}|\mathrm{Y}}\cdot\mathrm{P}_{\mathrm{XY}}$
  - $J_V(\text{互信息};\mathrm{Z})=\min_Q I(\mathrm{P}_{\mathrm{X}},\mathrm{Q}_{\mathrm{Z}|\mathrm{X}})$
  - 约束条件：
  - $\mathbb{E}_{\mathrm{P}_Q}[\log V(\mathrm{Z}|\mathrm{X})]\geq -D$
- 求使互信息$(\mathrm{X};\mathrm{Z})$ 最大的 $\mathrm{P}^*_{\mathrm{Z}|\mathrm{Y}}\in\mathcal{P}^{\mathrm{fine}}_{\mathcal{Z}|\mathcal{Y}}$。

输出：$\mathrm{P}^*_{\mathrm{Z}|\mathrm{Y}}$, R

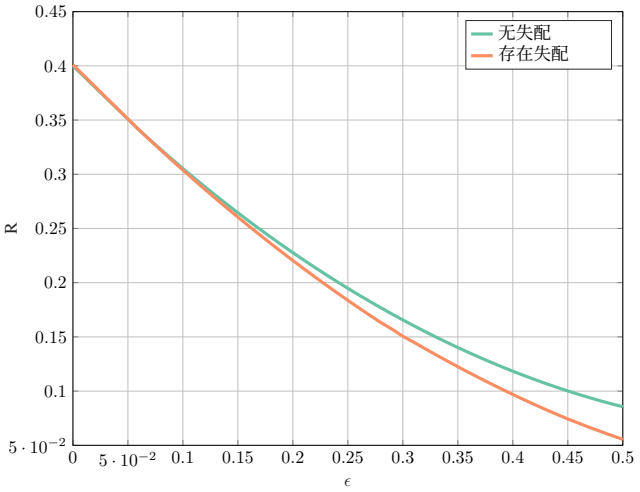

图 2：四元示例的失配性能。

与失配解码器所获得的码率，如图2所示，并将其表示为 ϵ 的函数。正如预期的那样，随着 $\epsilon\to 0$，码率之间的差异逐渐消失，因为随着 $\epsilon\to 0$，信道之间的失配会减轻。

D. 连续字母表无记忆信道

本节将上一节的下界修改为适用于连续字母表高斯信道的形式。如前所述，离散字母表这一假设对于定理 1 中 LM 速率的推导至关重要。事实上，在标准的点对点通信场景中，若不存在失配，则连续字母表的编码定理可通过对

信道的连续输入和输出进行精细量化并取极限得到。遗憾的是，这种方法不适用于失配情形，因为对输出进行量化会改变解码器，而根据假设，解码器应保持固定；此外，要追踪并评估这一变化对最终结果的影响也极具挑战性。为此，我们还在定理 2 中推导了相应的广义互信息速率，作为一个下界。正如我们将看到的，这一下界便于从离散字母表信道推广到连续字母表信道。

从技术记号的角度来看，我们首先将所有概率质量函数改写为概率密度的形式。此外，解码度量 $V(z|x)$ 也定义在 $\mathbb{R}\times\mathbb{R}$ 上。另外，我们对信道输入添加约束；否则，正如大多数连续信道中的情况一样，容量可能无界。因此，每个传输序列x^n 都必须满足$\frac{1}{n}\sum_{i=1}^{n} c(x_i) \le \Gamma$，其中代价函数为 $c(x)$，阈值为 Γ。通常，我们取 $c(x)=x^2$ 和 Γ，二者均表示功率约束以及每符号所允许的最大功率。

定理 3：对于任意采用失配解码度量 V、输入代价函数 $c(\cdot)$ 和输入代价阈值 Γ 的连续无知中继信道，若码率 R 满足：则随机编码错误概率趋于零。

$$\begin{aligned} R \le \quad & \max_{f_{\mathsf{XZ}},\, f_{\mathsf{Z}|\mathsf{Y}}} \quad I_{\mathsf{GMI}}(f_{\mathsf{XZ}}) \\ & \text{s.t.} \qquad\quad I(\mathsf{Y};\mathsf{Z}) \le B, \quad \mathbb{E}\left[c(\mathsf{X})\right] \le \Gamma. \end{aligned} \tag{10}$$

其中

$$I_{\mathsf{GMI}}(\mathsf{P}_{\mathsf{XZ}}) = \max_{\lambda\ge 0} \int \mathrm{d}x\mathrm{d}z\, f_{\mathsf{XZ}}(x,z) \log \frac{V(z|x)^{\lambda}}{\int \mathsf{P}_{\mathsf{X}}(x')V(z|x')^{\lambda}\mathrm{d}x'}. \tag{11}$$

证明：式 (7) 中的对偶表达式也可以直接推导出来（而非如 D 节所示先推导对偶优化问题），方法是采用类似于 Gallager 针对最大似然解码所用的分析。前一种方法需要将求和替换为积分，并使用标准的清除论证，从随机生成的码本中构造一个由符合约束的码字组成的子码本。 ■

E. 示例：衰落信道

接下来，我们以衰落信道为例说明我们的结果。衰落信道是一种基础的无线通信信道模型。通常，我们假设信道为复值信道，且加性噪声为圆对称复高斯（CSCG）噪声。具体而言，我们考虑如下形式的无记忆时变快衰落模型：$\mathsf{Y}_i = \mathsf{H}_i\mathsf{X}_i + \mathsf{N}_i$，其中 $\mathsf{X}_i \in \mathbb{C}$为输入，$\mathsf{N}_i \in \mathbb{C}$ 为加性噪声，$\mathsf{H}_i \in \mathbb{C}$ 为衰落系数。我们假设 $\{\mathsf{N}_i\}_{i=1}^{n}$ 独立同分布，服从 $\mathcal{CN}(0,\sigma^2)$，且 H_i 是一个独立同分布序列，其密度函数为 S_{H}。

1) 完美信道知识：如果解码器完美知道每个随机实现值 $\mathsf{H}_i = h_i$，那么由于噪声的高斯性，以及在假设输入分布为高斯分布时高斯压缩的最优性，最优译码规则将是以下最近邻规则的加权版本：

$$\hat{m} = \underset{j=1,\ldots,M}{\arg\min} \sum_{i=1}^{n} |z_i - h_i x_i^{(j)}|^2. \tag{12}$$

类似地，在功率约束 $\mathbb{E}\left[|\mathsf{X}|^2\right] \le \Gamma$（即 $|x|^2 \quad c\ x\ =$ ( ) ）和输入分布为高斯分布的假设下，使用从 Y 到 Z 的高斯测试信道即可达到最优码率，其表达式为

$$R_{\mathrm{IB}}^{CG}(\Gamma,\sigma^2,S_{\mathsf{H}}) = \mathbb{E}\left[\log\frac{|\mathsf{H}|^2\Gamma+\sigma^2+q}{\sigma^2+q}\right]. \tag{13}$$

显然，在快衰落信道中，假设解码是匹配的并不现实，因为这要求在任意时刻都完美地知道 H_i。

2) 不完美信道知识：如前所述，在假设采用高斯信号的情况下，(X,Y) 是联合高斯的。在不失配的标准信息瓶颈设置中，此时从 Y 到 Z 的最优测试信道也是高斯的。因此，我们也将此测试信道用于失配设置，并假设从 Y 到 Z 的信道为高斯信道，即存在一个 $\mathsf{W}\sim\mathcal{CN}(0,q)$，使得 $\mathsf{Z}=\mathsf{Y}+\mathsf{W}=\mathsf{HX}+\mathsf{N}+\mathsf{W}$。其中，$q$ 的取值由互信息约束方程的解确定，即：

$$B \ge I(\mathsf{Y};\mathsf{Z}) = \mathbb{E}_{\mathsf{H}}\left[\log\frac{|\mathsf{H}|^2\Gamma+\sigma^2+q}{q}\right]. \tag{14}$$

q^* 为上述方程的解。我们利用它通过对偶形式求出定理 2 的广义互信息速率。

我们采用一个简单的不确定性模型，其中

$$\mathsf{H}_i = \hat{\mathsf{H}}_i + \Delta_i, \quad \mathbb{E}\left[\Delta|\hat{\mathsf{H}}_i\right] = 0, \tag{15}$$

其中，$\hat{\mathsf{H}}_i$ 是解码器已知的、可能为随机的 H 估计值，而 Δ 表示一个条件均值为零的未知误差项。为简化起见，我们假设$\hat{}$ $\{(\mathsf{H}_i,\Delta_i)\}_{i=1}^{n}$ 这些样本对相对于 $i=1,\ldots,n$ 独立同分布，且与信道输入和噪声相互独立。

在 $(\mathsf{H},\hat{\mathsf{H}})$ 的联合概率密度函数未知时（即使该函数已知，但难以设计相应的最优编码方案时亦然），采用加权最近邻编码是很自然的选择

$$\hat{m} = \underset{j=1,\ldots,M}{\arg\min} \sum_{i=1}^{n} |z_i - \hat{h}_i x_i^{(j)}|^2. \tag{16}$$

这是一种失配解码规则，因为在形式为 $\mathsf{Y}=\mathsf{HX}+\mathsf{N}$ 的模型下，它会是最优的。相应的解码度量为 $V(x,z)=e^{-|z-\hat{h}x|^2}$。

定理 4：考虑复值信道衰落估计设置，其中输出端已知估计值 $|\mathsf{H}|$，且误差项 Δ 的条件均值为零。在采用 X $\sim\ \mathcal{CN}(0,\Gamma)$ 的独立同分布随机编码以及加权最近邻解码时，广义互信息速率为

$$\begin{aligned} C(B) = \quad & \max_{q} \quad \mathbb{E}\left[\log\left(1+\frac{|\hat{\mathsf{H}}|^2\Gamma}{\mathbb{E}\left[|\Delta|^2|\hat{\mathsf{H}}\right]\Gamma+\sigma^2+q}\right)\right] \\ & \text{s.t.} \qquad \mathbb{E}_{\mathsf{H}}\left[\log\frac{|\mathsf{H}|^2\Gamma+\sigma^2+q}{q}\right] \le B. \end{aligned} \tag{17}$$

证明：由于篇幅有限，此处省略证明。我们通过一个数值示例来说明定理 4。我们选取 $\Gamma=\sigma=1$ 和 H 以及 Δ，使其为相互独立的瑞利分布 ■

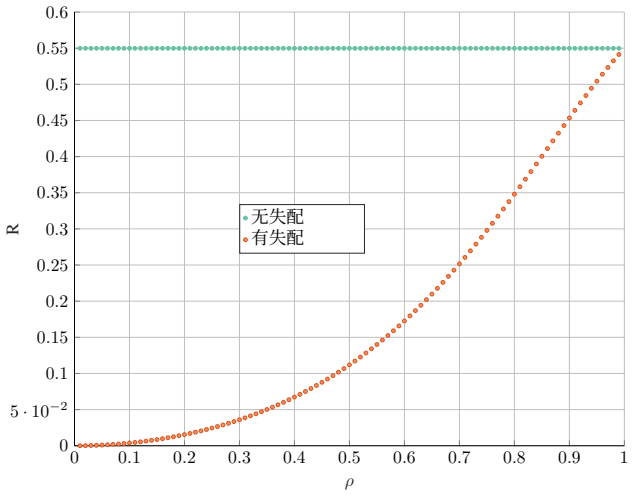

图3：衰落示例的失配性能。

随机变量，其中 $\sigma_{\mathsf{H}}^2 = \rho^2$ 和 $\sigma_{\Delta}^2 = 1 - \rho^2$。结果如图3所示。

III. 带失配中继的信息瓶颈信道

本节中，我们考虑一种具有失配的不同中继模型。具体而言，考虑图4所示的带失配中继的三节点点对点通信系统。在该模型中，发送方希望在中继的帮助下将消息 M 传送给接收方。我们考虑带失配中继的离散无记忆信息瓶颈信道（DM-IBC）$\langle \mathsf{X}, \mathsf{P}_{\mathsf{Y}|\mathsf{X}}, \mathcal{Y}, \mathcal{Z}, d_0(y,z)\rangle$，它由三个有限集合 $\mathcal{X}$，$\mathcal{Y}$，$\mathcal{Z}$、定义在 $\mathcal{Y}$ 上的一组条件概率质量函数 $\mathsf{P}_{\mathsf{Y}|\mathsf{X}}$（每个 x对应一个），以及定义在 $\mathcal{Y} \times \mathcal{Z}$上的编码度量 $d_0(y,z)$组成。

带失配中继的 DM-IBC 的 $(2^{nR}, 2^{nB}, n)$ 码定义方式与第二节类似，但有以下例外：

- 一个失配中继编码器，根据$w = \arg\min_{w \in \mathcal{W}} d_0^n(y^n, z^n(w))$ 将索引$\hat{w} \in \mathcal{W}$分配给每个接收序列 $y^n \in \mathcal{Y}^n$，其中

$$d_0^n(y^n, z^n) = \frac{1}{n}\sum_{i=1}^{n} d_0(y_i, z_i), \tag{18}$$

- 一个解码器，为每个接收表示索引 $w \in \mathcal{W}$ 分配一个估计值 $\hat{m}$ 或一条错误消息 e。

我们假设解码器知道信道以及失配中继的码本。此外，假设消息 M 在消息集 $\mathcal{M}$ 上均匀分布。

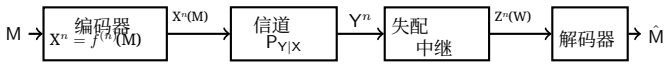

图 4：带失配中继的盲通信系统

注 1：如果中继处的失配是由压缩所用的测试信道错误造成的，那么失配压缩问题会相对简单一些；也就是说，中继受到限制，只能使用 $\mathsf{Q}_{\mathsf{Z}|\mathsf{Y}}$ 进行联合典型性编码。这等价于在标准信息瓶颈问题中选择了非最优测试信道。在这种情况下，所得容量为

$$C(B, \mathsf{Q}_{\mathsf{Z}|\mathsf{Y}}) = \begin{cases} \max_{\mathsf{P}_{\mathsf{X}}} I(\mathsf{X};\mathsf{Z}), & I(\mathsf{P}_{\mathsf{Y}}, \mathsf{Q}_{\mathsf{Z}|\mathsf{Y}}) \le B \\ 0, & \text{otherwise} \end{cases} \tag{19}$$

其中 $\mathsf{Q}_{\mathsf{Z}|\mathsf{X}}(z|x) = \sum_{y \in \mathcal{Y}} \mathsf{Q}_{\mathsf{Z}|\mathsf{Y}}(z|y)\mathsf{P}_{\mathsf{Y}|\mathsf{X}}(y|x)$。请注意，此设置与失配中继压缩设置之间的区别。在标准信息瓶颈问题中，优化测试信道 $\mathsf{P}_{\mathsf{Z}|\mathsf{Y}}$，并根据 P_{Z}生成随机压缩码本。给定该码本后，根据给定的 P_{YZ}，通过联合典型性编码选择压缩索引。相比之下，在失配中继场景中，压缩码字根据给定的固定（失配）度量进行选择，该度量不一定与 P_{YZ}匹配。

该设置下的主要结果由以下定理给出。

定理 5：失配中继信道的容量下界如下：

$$C(B) \ge \max_{\mathsf{P}_{\mathsf{X}}, \mathsf{P}_{\mathsf{Z}}} \min_{\mathsf{Q}_{\mathsf{YZ}} \in \mathcal{Q}} I(\mathsf{X};\mathsf{Z}) \tag{20}$$

其中 $\mathcal{Q} = \arg\min_{\mathsf{Q}_{\mathsf{YZ}} \in \mathcal{D}} \sum_{y \in \mathcal{Y}} \sum_{z \in \mathcal{Z}} \mathsf{Q}_{\mathsf{YZ}}(y,z) d_0(y,z)$，且

$$\mathcal{D} = \left\{ \mathsf{Q}_{\mathsf{YZ}} : \mathsf{Q}_{\mathsf{Y}} = \mathsf{P}_{\mathsf{Y}}, \mathsf{Q}_{\mathsf{Z}} = \mathsf{P}_{\mathsf{Z}}, \mathsf{Q}_{\mathsf{XZ}}(x,z) = \sum_{y \in \mathcal{Y}} \mathsf{P}_{\mathsf{X}|\mathsf{Y}}(x|y)\mathsf{Q}_{\mathsf{YZ}}(y,z), I(\mathsf{Y};\mathsf{Z}) \le B \right\}. \tag{21}$$

证明：证明见附录 E。 ■

IV. 总结与展望

我们研究了存在失配的点对点无知中继通信系统中的可靠通信问题。具体而言，我们研究了中继端或解码器端的失配问题。我们建立了系综紧致的可达率及其对偶表示。此外，我们还将这些结果应用于若干特定实例：四元信道和衰落信道。我们提出了一种交替算法来求得这些速率。

未来的工作可以考虑这一问题的逆界，例如采用[13]和 [30]中描述的方法。或者，也可以探究信道 $\mathsf{P}_{\mathsf{Y}|\mathsf{X}}$的失配容量与整个信道 $\mathsf{P}_{\mathsf{Z}|\mathsf{X}}$的失配容量之间的关系。另一种可能是将结果推广到状态依赖信道 [31], [32],，其中中继知道状态序列（也可以假设该序列独立同分布），并将状态描述作为其消息的一部分，以便向接收方提供信道状态信息，从而提升其解码性能。

致谢

本工作得到欧盟“地平线 2020”研究与创新计划（资助协议编号：694630）、ISF 资助项目 1791/17 以及通过以色列经济与科学部长开展的 WIN 联盟的支持。

参考文献

[1] A. Sanderovich、S. Shamai、Y. Steinberg 和 G. Kramer，“通过去中心化处理进行通信”，IEEE 信息论汇刊，第 54 卷，第 3008–3023 页，2008 年 7 月。[2] I. Estella Aguerri、A. Zaidi、G. Caire 和 S. Shamai Shitz，“关于具有无认知中继的云无线接入网络的容量”，IEEE 信息论汇刊，第 65 卷，第 4575–4596 页，2019 年 7 月。[3] T. Richardson 和 R. Urbanke，《现代编码理论》。剑桥，英国：剑桥大学出版社，2008 年。[4] N. Tishby、F. C. N. Pereira 和 W. Bialek，“离散信源在保真度准则下的编码定理”，载于第 37 届年度 Allerton 会议：通信、控制与计算，第 368—-377 页，1999 年 9 月。[5] A. Zaidi、I. Estella-Aguerri 和 S. Shamai (Shitz)，“关于信息瓶颈问题：模型、联系、应用与信息论视角”，《熵》，第 22 卷，第 2 期，2020 年。[6] Z. Goldfeld 和 Y. Polyanskiy，“信息瓶颈问题及其在机器学习中的应用”，IEEE 信息论精选领域期刊，第 1 卷，第 19–38 页，2020 年 5 月。[7] I. E. Aguerri 和 A. Zaidi，“分布式变分表示学习”，IEEE 模式分析与机器智能汇刊，第 43 卷，第 120–138 页，2021 年 1 月。[8] C. E. Shannon 等，“具有保真度准则的离散信源编码定理”，IRE 全国会议记录，第 4 卷，第 142-163 号，第 1 页，1959 年。[9] R. Dobrushin 和 B. Tsybakov，“带有附加噪声的信息传输”，IRE 信息论汇刊，第 8 卷，第 293–304 页，1962 年 9 月。[10] J. Wolf 和 J. Ziv，“将噪声信息传输给具有最小失真的噪声接收机”，IEEE 信息论汇刊，第 16 卷，第 406–411 页，1970 年 7 月。[11] T. A. Courtade 和 T. Weissman，“对数损失下的多终端信源编码”，IEEE 信息论汇刊，第 60 卷，第 740–761 页，2014 年 1 月。[12] A. Painsky 和 G. Wornell，“逻辑损失函数的普适性”，载于 2018 IEEE 国际信息论研讨会（ISIT），第 936–940 页，2018 年。[13] J. Scarlett、A. G. i F`abregas、A. Somekh-Baruch 和 A. Martinez，“失配译码的信息论基础”，《通信与信息论基础与趋势》，第 17 卷，第 2–3 期，第 149–401 页，2020 年。[14] J. Y. N. Hui，《多址接入的基本问题》。博士论文，麻省理工学院，1983 年。[15] I. Csiszar 和 J. Korner，“图分解：编码定理的新关键”，IEEE 信息论汇刊，第 27 卷，第 5–12 页，1981 年 1 月。[16] I. Csiszar 和 P. Narayan，“给定解码度量下的信道容量”，IEEE 信息论汇刊，第 41 卷，第 35–43 页，1995 年 1 月。[17] G. Kaplan 和 S. Shamai，“复合信道的信息率与错误指数及其在衰落环境中反极性信号传输中的应用”，AEU。《电子与传输技术档案》，第 47 卷，第 4 期，第 228–239 页，1993 年。¨N. Merhav、G. Kaplan、A. Lapidoth 和 S. Shamai Shitz，“关于失配译码器的信息率”，IEEE 信息论汇刊，第 40 卷，第 1953–1967 页，1994 年 11 月。[18] A. Lapidoth，“失配在率失真理论中的作用”，IEEE 信息论汇刊，第 43 卷，第 38–47 页，1997 年 1 月。[19] L. Zhou、V. Y. F. Tan 和 M. Motani，“使用高斯码本对任意信源进行率失真的精细渐近分析”，IEEE 信息论汇刊，第 65 卷，第 3145–3159 页，2019 年 5 月。[20] L. Bai、Z. Wu 和 L. Zhou，“使用高斯码本进行渐进细化的可达精细渐近结果”，arXiv 预印本 arXiv:2208.03926，2022 年。[21] M. Kanabar 和 J. Scarlett，“失配率失真理论：系综、界与一般字母表”，arXiv 预印本 arXiv:2203 15193，2022 年。[22] J. Choi、D. J. Love、D. R. Brown 和 M. Boutin，“使用空间复用的 MIMO 无线系统中的量化分布式接收”，IEEE 信号处理汇刊，第 63 卷，第 3537–3548 页，2015 年 7 月。[23] D. R. Brown、M. Ni、U. Madhow 和 P. Bidigare，“粗量化观测交换下的分布式接收”，载于 2013 年第 47 届信息科学与系统年度会议（CISS），第 1–6 页，2013 年。[24] L. Liang、S. Bi 和 R. Zhang，“基于 OFDMA 的云无线接入网络中用于吞吐量最大化的联合功率控制与前传速率分配”，arXiv 预印本 arXiv:1407.3855，2014 年。[25] R. Wang、D. Richard Brown、M. Ni、U. Madhow 和 P. Bidigare，“硬判决交换下分布式接收的中断概率分析”，载于 2013 年 Asilomar 信号、系统与计算会议，第 597–601 页，2013 年。[27] R. Blahut，“信道容量与率失真函数的计算”，IEEE 信息论汇刊，第 18 卷，第 460–473 页，1972 年 7 月。[28] S. Arimoto，“一种计算任意离散无记忆信道容量的算法”，IEEE 信息论汇刊，第 18 卷，第 14–20 页，1972 年 1 月。[29] R. G. Gallager，《信息论与可靠通信》，Robert G. Gallager 著。纽约：Wiley，1968 年。[30] A. Somekh-Baruch，“利用神谕接收机推导失配可靠性函数与容量的上界”，arXiv 预印本 arXiv:2203.08524，2022 年。[31] H. Xu、T. Yang、G. Caire 和 S. Shamai (Shitz)，“具有无认知中继的瑞利衰落 MIMO 信道的信息瓶颈”，Information，第 12 卷，第 4 期，2021 年。[32] H. Xu、T. Yang、G. Caire 和 S. Shamai Shitz，“具有信道状态信息的无认知中继的信息瓶颈：向量情形”，载于 2021 年 IEEE 国际信息论研讨会（ISIT），第 2483–2488 页，2021 年。[33] A. El Gamal 和 Y.-H. Kim，《网络信息论》。剑桥大学出版社，2011 年。[34] S. Boyd 和 L. Vandenberghe，《凸优化》。剑桥大学出版社，2004 年。[35] T. M. Cover 和 J. A. Thomas，《信息论基础》。美国新泽西州霍博肯：Wiley，2006 年。

附录 A 符号约定

全文中，随机变量使用无衬线字体表示，例如 X；其取值用相应的小写字母表示，例如 x；其字母表用相应的花体字母表示，例如 $\mathcal{X}$。令$\mathcal{X}^n$ 表示由 $\mathcal{X}$中的元素组成的所有 n元组的集合。 $\mathcal{X}^n$ 中的元素表示为 $x^n = (x_1, x_2, \ldots, x_n)$。为简化表述，本文将频繁使用向量记法 $\mathbf{x} = x^n$ 。有限集（例如$\mathcal{X}$）的基数表示为 $|\mathcal{X}|$。X 的概率质量函数（概率质量函数）、X 与 Y 的联合概率质量函数，以及给定 Y 时 X 的条件概率质量函数，分别表示为 P_{X}、P_{XY} 和 $\mathsf{P}^{\mathsf{X|Y}}$。$\mathsf{X}$ 的期望表示为 $\mathbb{E}[\mathsf{X}]$。事件 $\mathcal{E}$ 的概率表示为 $\mathsf{P}\{\mathcal{E}\}$。除非另有说明，本文中的所有对数均以 2 为底。

X 与 Y 之间的互信息定义为

$$I(\mathsf{X};\mathsf{Y}) = \sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} \mathsf{P}_{\mathsf{XY}}(x,y)\log\frac{\mathsf{P}_{\mathsf{XY}}(x,y)}{\mathsf{P}_{\mathsf{X}}(x)\mathsf{P}_{\mathsf{Y}}(y)}. \tag{22}$$

概率测度 P 和 Q 之间的 Kullback-Leibler（KL）散度定义为

$$\mathsf{D}(\mathsf{P}\|\mathsf{Q}) \triangleq \mathbb{E}_{\mathsf{P}}\left[\log\frac{\mathsf{P}}{\mathsf{Q}}\right]. \tag{23}$$

$\hat{}$ x^n 的经验分布记为 $\mathsf{P}_{\mathbf{x}}(a)$，对于每个 $a\in\mathcal{X}$，即

$$\hat{\mathsf{P}}_{\mathbf{x}}(a) \triangleq \frac{1}{n}\sum_{i=1}^{n}\mathbf{1}\{x_i = a\}, \tag{24}$$

其中，$\mathbf{1}\{\cdot\}$ 是事件指示函数。

对于由边缘概率分布 P_{X} 和条件概率分布 $\mathsf{P}_{\mathsf{Y|X}}$ 定义的一对随机变量 (X,Y)，我们将边缘化运算 $\circ$ 定义如下

$$\mathsf{P}_{\mathsf{Y}} = \mathsf{P}_{\mathsf{Y|X}}\circ\mathsf{P}_{\mathsf{X}} \rightleftharpoons \mathsf{P}_{\mathsf{Y}}(y) = \sum_{x\in\mathcal{X}}\mathsf{P}_{\mathsf{Y|X}}(y|x)\mathsf{P}_{\mathsf{X}}(x). \tag{25}$$

附录 B 定理 1 的可达性证明

在继续描述编码方案之前，先考虑失配解码规则 (1)。注意，

$$V^n(\mathbf{z}|\mathbf{x}) = \prod_{i=1}^{n} V(z_i|x_i) \tag{26}$$

$$= 2^{n\sum_{(a,c)\in\mathcal{X}\times\mathcal{Z}} \hat{\mathsf{P}}_{\mathbf{xz}}(a,c)\log V(c|a)}. \tag{27}$$

因此，式 (1) 中的失配解码规则等价于

$$\hat{m}(\mathbf{z}) = \arg\max_{m\in\mathcal{M}} \sum_{(a,c)\in\mathcal{X}\times\mathcal{Z}} \hat{\mathsf{P}}_{\mathbf{x}(m)\mathbf{z}}(a,c)\log V(c|a). \tag{28}$$

现在考虑以下阈值译码规则。固定某个 $\epsilon > 0$。如果消息 $\tilde{m}$ 是满足以下条件的唯一消息，则解码器判定发送的是该消息

$$(\mathbf{x}(\tilde{m}), \mathbf{z}(w)) \in \mathcal{T}_{\epsilon,\theta}^{(n)}, \tag{29}$$

其中

$$\mathcal{T}_{\epsilon,\theta}^{(n)} \triangleq \left\{ (\mathbf{x},\mathbf{z}) \colon \sum_{(a,c)\in\mathcal{X}\times\mathcal{Z}} \hat{\mathsf{P}}_{\mathbf{xz}}(a,c)\log V(c|a) \geq \theta - \epsilon \right\}. \tag{30}$$

注意，如果存在这样的 $\tilde{m}$ ，那么它也是式 (28) 中真实译码规则的解，即 $\tilde{m} = \hat{m}$ 因此，式 (28) 中失配译码器的错误概率由阈值译码器的错误概率上界所限制。特别地，我们有

$$\mathsf{P}_c = \mathsf{P}\left\{\hat{\mathsf{M}} = \mathsf{M}\right\} \tag{31}$$

$$\geq \mathsf{P}\left\{\hat{\mathsf{M}} = \mathsf{M}, (\mathsf{X}^n(\mathsf{M}), \mathsf{Z}^n(W)) \in \mathcal{T}_{\epsilon,\theta}^{(n)}\right\} \tag{32}$$

$$= \mathsf{P}\left\{\text{threshold decoder is correct}\right\}. \tag{33}$$

在下文中，我们设 $\theta = \sum_{(a,\ c)\in\mathcal{X}\times\mathcal{Z}} \mathsf{P}_{\mathsf{XZ}}(a,c)\log\ V(c|a)$。

我们采用随机编码。令 $0 < \epsilon'' < \epsilon' < \epsilon$。固定一个概率质量函数 P_X 和一个条件概率质量函数 $\mathsf{P}_{\mathsf{Z}|\mathsf{Y}}$，它们诱导出边缘概率质量函数 $\mathsf{P}_\mathsf{Y}(y) = \sum_{x\in\mathcal{X}} \mathsf{P}_{\mathsf{Y}|\mathsf{X}}(y|x)\mathsf{P}_\mathsf{X}(x)$ 以及 $\mathsf{P}_\mathsf{Z}(z) = \sum_{y\in\mathcal{Y}} \mathsf{P}_{\mathsf{Z}|\mathsf{Y}}(z|y)\mathsf{P}_\mathsf{Y}(y)$。

a) 发射机的恒定组成码本生成：令 $\mathsf{Q}_{\mathsf{X}^n}$ 表示一个与 P_X具有相同支撑集、且满足 $\|\mathsf{P}_\mathsf{X} - \mathsf{Q}_{\mathsf{X}^n}\| \leq \frac{1}{n}$的任意类型。码字独立地从以下分布中抽取

$$\mathsf{Q}_{\mathsf{X}^n}(x^n) = \frac{1}{|\mathcal{T}^n(\mathsf{Q}_{\mathsf{X}^n})|}\mathbf{1}\left\{x^n \in \mathcal{T}^{(n)}(\mathsf{Q}_{\mathsf{X}^n})\right\}. \tag{34}$$

也就是说，在经验分布为 $\mathsf{Q}_{\mathsf{X}^n}$的序列集合上，每个码字出现的概率相等。码本会提供给编码器和解码器，但不会提供给中继。

b) 中继码本生成与压缩：根据 P_Z生成随机码本 $\mathcal{C}_R = \{z^n(w)\}$, $w \in [1,$ 码率为 B 和 n。该码本对中继和解码器公开，但不对发射机公开。接收y^n 后，中继寻找一个索引 w ，使得 $z^n(w)$与 y^n联合典型。当 $B > I(\mathsf{Y};\mathsf{Z}) + \delta(\epsilon')$时，这样的 w 存在，其中 $\delta(\epsilon')$随着 $\epsilon' \to 0$趋于零。

c) 失配译码：获得 w后，接收方寻找唯一的消息 $\hat{m}$ ，使得

$$\hat{m} = \arg\max_{m\in\mathcal{M}} V^n(z^n(w)|x^n(m)). \tag{35}$$

如果符合条件的消息不存在或多于一个，则声明发生错误事件 $\mathcal{E}$。

d) 失配译码器处错误概率的分析：不失一般性地，假设发送了 $\mathsf{M} = 1$ ，且中继选择了 $\mathsf{W} = g(\mathsf{Y}^n)$。如果对于某个 $m \neq 1$，有$V^n(\mathsf{Z}^n(\mathsf{W})|\mathsf{X}^n(m)) > V^n(\mathsf{Z}^n(\mathsf{W})|\mathsf{X}^n(1))$，则解码器会出错。因此，错误概率的上界如下

$$P_e = \mathsf{P}\left\{\bigcup_{m=2}^{2^{nR}} V^n(\mathsf{Z}^n(\mathsf{W})|\mathsf{X}^n(m)) \geq V^n(\mathsf{Z}^n(\mathsf{W})|\mathsf{X}^n(1))\right\} \tag{36}$$

$$\leq (2^{nR}-1)\mathsf{P}\left\{V^n(\mathsf{Z}^n(\mathsf{W})|\mathsf{X}^n(2)) \geq V^n(\mathsf{Z}^n(\mathsf{W})|\mathsf{X}^n(1))\right\}. \tag{37}$$

我们分别记作 X^n 、$\mathsf{X}^n(1)$ 和 $\bar{}\mathsf{X}^n$ 、$\mathsf{X}^n(2)$。因此，

$$P_e \leq 2^{nR}\cdot\mathsf{P}\left\{V^n(\mathsf{Z}^n(\mathsf{W})|\bar{\mathsf{X}}^n) \geq V^n(\mathsf{Z}^n(\mathsf{W})|\mathsf{X}^n)\right\}. \tag{38}$$

进一步注意，$\log\ V^n(z^n|x^n) = n\mathbb{E}_{\hat{\mathsf{P}}_{\mathbf{xz}}}[\log\ V(\mathsf{Z}|\mathsf{X})]$。记 $\mathsf{P} \triangleq \mathsf{P}_{\mathsf{X}^n,\mathsf{Z}^n}$和 $\tilde{}\mathsf{P} \triangleq \mathsf{P}_{\bar{\mathsf{X}}^n,\mathsf{Z}^n}$。注意，根据构造，$\tilde{}\mathsf{P}_\mathsf{X} = \mathsf{P}_\mathsf{X}$ 和 $\tilde{}\mathsf{P}_\mathsf{Z} = \mathsf{P}_\mathsf{Z}$。由此可得

$$\mathsf{P}\left\{V^n(z^n(W)|\bar{\mathsf{X}}^n) \geq V^n(z^n(W)|x^n)\right\} \tag{39}$$

$$= \mathsf{P}\left\{\mathbb{E}_{\tilde{\mathsf{P}}}[\log V(\mathsf{Z}|\mathsf{X})] > \mathbb{E}_{\mathsf{P}}[\log V(\mathsf{Z}|\mathsf{X})]\right\} \tag{40}$$

$$= \sum_{\substack{\tilde{\mathsf{P}}\in\mathcal{P}_{\mathcal{X}\times\mathcal{Z}}:\ \tilde{\mathsf{P}}_\mathsf{X}=\mathsf{P}_\mathsf{X},\tilde{\mathsf{P}}_\mathsf{Z}=\mathsf{P}_\mathsf{Z} \\ \mathbb{E}_{\tilde{\mathsf{P}}}[\log V(\mathsf{Z}|\mathsf{X})]>\mathbb{E}_{\mathsf{P}}[\log V(\mathsf{Z}|\mathsf{X})]}} \mathsf{P}\left\{(\bar{\mathsf{X}}^n, z^n) \in \mathcal{T}^{(n)}\left(\tilde{\mathsf{P}}\right)\right\} \tag{41}$$

$$\overset{(a)}{\leq} \sum_{\substack{\tilde{\mathsf{P}}\in\mathcal{P}_{\mathcal{X}\times\mathcal{Z}}:\ \tilde{\mathsf{P}}_\mathsf{X}=\mathsf{P}_\mathsf{X},\tilde{\mathsf{P}}_\mathsf{Z}=\mathsf{P}_\mathsf{Z} \\ \mathbb{E}_{\tilde{\mathsf{P}}}[\log V(\mathsf{Z}|\mathsf{X})]>\mathbb{E}_{\mathsf{P}}[\log V(\mathsf{Z}|\mathsf{X})]}} e^{-n\mathsf{D}\left(\tilde{\mathsf{P}}\|\mathsf{P}_\mathsf{X}\times\mathsf{P}_\mathsf{Z}\right)} \tag{42}$$

$$\leq (n+1)^{|\mathcal{X}|\cdot|\mathcal{Z}|} \max_{\substack{\tilde{\mathsf{P}}\in\mathcal{P}_{\mathcal{X}\times\mathcal{Z}}:\ \tilde{\mathsf{P}}_\mathsf{X}=\mathsf{P}_\mathsf{X},\tilde{\mathsf{P}}_\mathsf{Z}=\mathsf{P}_\mathsf{Z} \\ \mathbb{E}_{\tilde{\mathsf{P}}}[\log V(\mathsf{Z}|\mathsf{X})]>\mathbb{E}_{\mathsf{P}}[\log V(\mathsf{Z}|\mathsf{X})]}} e^{-n\mathsf{D}\left(\tilde{\mathsf{P}}\|\mathsf{P}_\mathsf{X}\times\mathsf{P}_\mathsf{Z}\right)} \tag{43}$$

$$= (n+1)^{|\mathcal{X}|\cdot|\mathcal{Z}|} e^{-nJ_V(\mathsf{X};\mathsf{Z})}. \tag{44}$$

其中，由于 X 与 $\tilde{}z^n$ 相互独立，因此有 (a)。

问题在于 P 是什么。下一节将说明，我们可以以高概率将 P 取为 P_{XZ},

A. 马尔可夫三元组的联合典型性

由于添加约束条件会降低概率，因此

$$\mathsf{P}\left\{(\mathsf{X}^n(1), \mathsf{Z}^n(\mathsf{W}(\mathsf{Y}^n))) \in \mathcal{T}_\epsilon^{(n)}(\mathsf{P}_{\mathsf{XZ}})\right\}$$

$$\geq \mathsf{P}\left\{(\mathsf{X}^n(1), \mathsf{Y}^n, \mathsf{Z}^n(\mathsf{W}(\mathsf{Y}^n))) \in \mathcal{T}_\epsilon^{(n)}(\mathsf{P}_{\mathsf{XYZ}})\right\}. \tag{45}$$

中继处的编码器采用联合典型性编码，因此

$$\lim_{n\to\infty} \mathsf{P}\left\{(y^n, \mathsf{Z}^n) \in \mathcal{T}_{\epsilon'}^{(n)}(\mathsf{P}_{\mathsf{YZ}})\right\} = 1. \tag{46}$$

此外，对于每个 $z^n \in \mathcal{T}_{\epsilon'}^{(n)}(\mathsf{Z}|y^n)$

$$\mathsf{P}\{\mathsf{Z}^n(\mathsf{W}) = z^n|\mathsf{Y}^n = y^n\} \tag{47}$$
$$= \mathsf{P}\left\{\mathsf{Z}^n(\mathsf{W}) = z^n, \mathsf{Z}^n(\mathsf{W}) \in \mathcal{T}_{\epsilon'}^{(n)}(\mathsf{Z}|y^n)|\mathsf{Y}^n = y^n\right\} \tag{48}$$
$$\leq \mathsf{P}\left\{\mathsf{Z}^n(\mathsf{W}) = z^n|\mathsf{Y}^n = y^n, \mathsf{Z}^n(\mathsf{W}) \in \mathcal{T}_{\epsilon'}^{(n)}(\mathsf{Z}|y^n)\right\} \tag{49}$$
$$= 2^{-n(H(\mathsf{Z}|\mathsf{Y})-\delta(\epsilon'))}. \tag{50}$$

同理

$$\mathsf{P}\{\mathsf{Z}^n(\mathsf{W}) = z^n|\mathsf{Y}^n = y^n\} \geq (1-\epsilon')2^{-n(H(\mathsf{Z}|\mathsf{Y})-\delta(\epsilon'))}. \tag{51}$$

因此，Markov 引理 [33, 引理 12.1] 的条件均已满足，因此

$$\lim_{n\to\infty} \mathsf{P}\left\{(\mathsf{X}^n(1), \mathsf{Y}^n, \mathsf{Z}^n(\mathsf{W}(\mathsf{Y}^n))) \in \mathcal{T}_\epsilon^{(n)}\right\} = 1. \tag{52}$$

附录 C
定理 1 的系综紧致性证明

假设码率 R 大于随机编码容量。下面证明随机编码错误概率趋于 1。也就是说，假定可达率 R 大于所有满足 $I(\mathsf{Y};\mathsf{Z}) \leq B$ 的 $J_V(\mathsf{X};\mathsf{Z})$。与常规情形相比，额外的约束 $I(\mathsf{Y};\mathsf{Z}) \leq B$ 带来了困难。

我们在发射机和中继处均采用随机编码。固定 $\epsilon > 0$，令 $\mathcal{T}_\epsilon^{(n)}(\mathsf{P}_{\mathsf{XY}})$ 和 $\mathcal{T}_\epsilon^{(n)}(\mathsf{P}_{\mathsf{Z}})$ 分别表示与输入和输出边缘分布相关的 ϵ–典型集。同理，固定 $\delta > 0$，并令

$$\mathcal{S}_{\delta^-}^{(n)} \triangleq \{(x^n, z^n)\colon \log V(z^n|x^n) \leq n(-D+\delta)\}. \tag{53}$$

不失一般性，假设消息 $\mathsf{M} = 1$ 已发送，且中继选择了索引 W。平均正确消息译码概率可按如下方式给出上界：

$$1 - \bar{P}_e \tag{54}$$
$$= \mathsf{P}\{V(\mathsf{Z}^n|\mathsf{X}^n(m)) < V(\mathsf{Z}^n|\mathsf{X}^n(1)) \forall m \neq 1\} \tag{55}$$
$$\leq \mathsf{P}\left\{\begin{smallmatrix} V(\mathsf{Z}^n|\mathsf{X}^n(m)) < V(\mathsf{Z}^n|\mathsf{X}^n(1)) \forall m \in [2:2^{nR}], \\ (\mathsf{X}^n(1), \mathsf{Z}^n) \in \mathcal{S}_{\delta^-}^{(n)}, \mathsf{Y}^n \in \mathcal{T}_\epsilon^{(n)}(\mathsf{P}_{\mathsf{Y}}), \mathsf{Z}^n \in \mathcal{T}_\epsilon^{(n)}(\mathsf{P}_{\mathsf{Z}}) \end{smallmatrix}\right\}$$
$$+ \mathsf{P}\left\{\mathsf{Y}^n \notin \mathcal{T}_\epsilon^{(n)}(\mathsf{P}_{\mathsf{Y}})\right\} + \mathsf{P}\left\{\mathsf{Z}^n \notin \mathcal{T}_\epsilon^{(n)}(\mathsf{P}_{\mathsf{Z}})\right\}$$
$$+ \mathsf{P}\left\{(\mathsf{X}^n(1), \mathsf{Z}^n) \notin \mathcal{S}_{\delta^-}^{(n)}\right\}. \tag{56}$$

由于 $\mathsf{P}_{\mathsf{Y}^n}(y^n) = \prod_{i=1}^n \mathsf{P}_{\mathsf{Y}}(y_i)$ 且 $\mathsf{P}_{\mathsf{Z}^n}(z^n) = \prod_{i=1}^n \mathsf{P}_{\mathsf{Z}}(z_i)$，根据大数定律，$\mathsf{P}\left\{\mathsf{Y}^n \notin \mathcal{T}_{(n)\epsilon}(\mathsf{P}_{\mathsf{Y}})\right\}$ 和 $\mathsf{P}\left\{\mathsf{Z}^n \notin \mathcal{T}_{(n)\epsilon}(\mathsf{P}_{\mathsf{Z}})\right\}$ 在 $n \to \infty$ 时趋于零。

考虑最后一项的补集。

$$\mathsf{P}\left\{(\mathsf{X}^n(1), \mathsf{Z}^n) \in \mathcal{S}_{\delta^-}^{(n)}\right\} \tag{57}$$
$$\geq \mathsf{P}\left\{(\mathsf{X}^n(1), \mathsf{Z}^n) \in \mathcal{S}_{\delta^-}^{(n)}, (\mathsf{X}^n(1), \mathsf{Y}^n, \mathsf{Z}^n(\mathsf{W})) \in \mathcal{T}_\epsilon^{(n)}\right\} \tag{58}$$
$$= \mathsf{P}\left\{(\mathsf{X}^n(1), \mathsf{Y}^n, \mathsf{Z}^n(\mathsf{W})) \in \mathcal{T}_\epsilon^{(n)}(\mathsf{P}_{\mathsf{XYZ}})\right\} \tag{59}$$
$$\times \mathsf{P}\left\{(\mathsf{X}^n(1), \mathsf{Z}^n) \in \mathcal{S}_{\delta^-}^{(n)} \middle| (\mathsf{X}^n(1), \mathsf{Y}^n, \mathsf{Z}^n(\mathsf{W})) \in \mathcal{T}_\epsilon^{(n)}\right\}.$$

考虑第一个乘数的补事件，

$$\mathsf{P}\left\{(\mathsf{X}^n(1), \mathsf{Y}^n, \mathsf{Z}^n(\mathsf{W})) \notin \mathcal{T}_\epsilon^{(n)}(\mathsf{P}_{\mathsf{XYZ}})\right\} \tag{60}$$
$$= \mathsf{P}\left\{\begin{smallmatrix} \{(\mathsf{X}^n(1), \mathsf{Y}^n) \notin \mathcal{T}_\epsilon^{(n)}\} \\ \cup\{\{(\mathsf{X}^n(1), \mathsf{Y}^n, \mathsf{Z}^n(\mathsf{W})) \notin \mathcal{T}_\epsilon^{(n)}\} \cap \{(\mathsf{X}^n(1), \mathsf{Y}^n) \in \mathcal{T}_\epsilon^{(n)}\}\} \end{smallmatrix}\right\} \tag{61}$$
$$\leq \mathsf{P}\left\{\left\{(\mathsf{X}^n(1), \mathsf{Y}^n) \notin \mathcal{T}_\epsilon^{(n)}\right\}\right\}$$
$$+ \mathsf{P}\left\{\begin{smallmatrix} \{(\mathsf{X}^n(1), \mathsf{Y}^n, \mathsf{Z}^n(\mathsf{W})) \notin \mathcal{T}_\epsilon^{(n)}\} \\ \cap\{(\mathsf{X}^n(1), \mathsf{Y}^n) \in \mathcal{T}_\epsilon^{(n)}\} \end{smallmatrix}\right\}. \tag{62}$$

由于 $\mathsf{P}_{\mathsf{X}^n(1)\mathsf{Y}^n}(x^n, y^n) = \prod_{i=1}^n \mathsf{P}_{\mathsf{XY}}(x_i, y_i)$，根据大数定律，第一项趋于零（当 $n \to \infty$ 时）。此外，由于 $(\mathsf{X}^n(1), \mathsf{Y}^n) \in \mathcal{T}_{(n)\epsilon}(\mathsf{P}_{\mathsf{XY}})$ 且 $\mathsf{Z}^n(\mathsf{W}) \sim \mathsf{P}_{\mathsf{Z}^n|\mathsf{Y}^n}$，根据马尔可夫引理 [33, 引理 12.1], ，第二项也趋于零（当 $n \to \infty$ 时）。接下来考虑第二个乘数，并取ϵ 足够小，

$$\mathsf{P}\left\{(\mathsf{X}^n(1), \mathsf{Z}^n) \in \mathcal{S}_{\delta^-}^{(n)} \middle| (\mathsf{X}^n(1), \mathsf{Y}^n, \mathsf{Z}^n(\mathsf{W})) \in \mathcal{T}_\epsilon^{(n)}\right\} = 1, \tag{63}$$

最后一个等号成立是因为 (x^n, z^n) 联合典型。因此，剩下的任务是为第一项（下文记为 $\mathsf{P}\{\mathcal{A}\}$）也给出一个趋于零的上界。注意，

$$\mathsf{P}\{\mathcal{A}\} \tag{64}$$
$$\leq \mathsf{P}\left\{\log V(\mathsf{Z}^n|\mathsf{X}^n(m)) \leq n(-D+\delta) \forall m \neq 1, \mathsf{Z}^n \in \mathcal{T}_\epsilon^{(n)}\right\} \tag{65}$$
$$= \sum_{z^n \in \mathcal{T}_\epsilon^{(n)}} \mathsf{P}_{\mathsf{Z}^n}(z^n) \mathsf{P}\{\log V(z^n|\mathsf{X}^n(m)) \leq n(-D+\delta) \forall m \neq 1\} \tag{66}$$
$$= \sum_{z^n \in \mathcal{T}_\epsilon^{(n)}} \mathsf{P}_{\mathsf{Z}^n}(z^n) \left[1 - \sum_{x^n\colon \log V(z^n|x^n) > n(-D+\delta)} \mathsf{P}_{\mathsf{X}^n}(x^n)\right]^{2^{nR}-1}. \tag{67}$$

记 $h(z^n) \sum_{x^n\colon \log V(z^n|x^n) > n(-D+\delta)} \mathsf{P}_{\mathsf{X}^n}(x^n)$，并利用 $1 - \alpha \leq e^{-\alpha}$ 这一事实，可得

$$\mathsf{P}\{\mathcal{A}\} \leq \sum_{z^n \in \mathcal{T}_\epsilon^{(n)}(\mathsf{P}_{\mathsf{Z}})} \mathsf{P}_{\mathsf{Z}^n}(z^n) \exp\left(-h(z^n) \cdot 2^{nR}\right) \tag{68}$$
$$\leq \exp\left(-\min_{z^n \in \mathcal{T}_\epsilon^{(n)}(\mathsf{P}_{\mathsf{Z}})} h(z^n) \cdot 2^{nR}\right). \tag{69}$$

为完成证明，我们需要证明，对于每个 $z^n \in \mathcal{T}_\epsilon^n$ Z h z^n 2^{-nI_V} $()(\mathsf{P})$，当 ϵ趋于零时，$()$ 指数级地不小于 $(\mathsf{X};\mathsf{Z})$，因此对于每个 $R > I_V(\mathsf{X};\mathsf{Z})$，$\mathsf{P}\{\mathcal{A}\}$ 本质上小于 $\exp\left(-e^{n(R-I_V(\mathsf{X};\mathsf{Z}))}\right) \to 0$。为此，

我们进一步给 $h(z^n)$ 找一个下界。首先注意，对于每个 $z^n \in \mathcal{T}_\epsilon^{(n)}(\mathsf{P}_\mathsf{Z})$,

$$h(z^n) = \sum_{x^n:\ \log V(z^n|x^n)>n(-D+\delta)} \mathsf{P}_{\mathsf{X}^n}(x^n) \tag{70}$$

$$= \sum_{\mathcal{T}(\hat{\mathsf{P}}_{\mathbf{x}|\mathbf{z}}):\ \log V(\mathbf{z}|\mathbf{x})\geq n(-D+\delta)} |\mathcal{T}(\hat{\mathsf{P}}_{\mathbf{x}|\mathbf{z}})| \cdot 2^{-n[\mathsf{D}(\hat{\mathsf{P}}_\mathbf{x}\|\mathsf{P}_\mathsf{X})+H(e\mathsf{P}_\mathbf{x})]} \tag{71}$$

$$\geq \max_{\mathcal{T}(\hat{\mathsf{P}}_{\mathbf{x}|\mathbf{z}}):\ \log V(\mathbf{z}|\mathbf{x})\geq n(-D+\delta)} |\mathcal{T}(\hat{\mathsf{P}}_{\mathbf{x}|\mathbf{z}})| \cdot 2^{-n[\mathsf{D}(\mathsf{P}_\mathbf{x}\|\mathsf{P}_\mathsf{X})+H(\hat{\mathsf{P}}_\mathbf{x})]}. \tag{72}$$

^ ^ ^ 注意，由于给定了 z^n， $\mathcal{T}(\mathsf{P}_{\mathbf{x}|\mathbf{z}})$ 同时定义了 $\mathsf{P}_\mathbf{x}$ 和 $\mathsf{P}_{\mathbf{x}|\mathbf{z}}$。现在，令

$$\mathcal{S} \triangleq \{x^n:\ \log V(z^n|x^n) > n(-D+\delta)\} \cap \mathcal{T}_\epsilon^{(n)}(\mathsf{P}_\mathsf{X}). \tag{73}$$

因此

$$h(z^n) \geq \max_{\mathcal{T}(\hat{\mathsf{P}}_{\mathbf{x}|\mathbf{z}})\in\mathcal{S}} |\mathcal{T}(\hat{\mathsf{P}}_{\mathbf{x}|\mathbf{z}})| \cdot 2^{-n[\mathsf{D}(\mathsf{P}_\mathbf{x}\|\mathsf{P}_\mathsf{X})+H(\hat{\mathsf{P}}_\mathbf{x})]} \tag{74}$$

$$\geq \max_{\mathcal{T}(\hat{\mathsf{P}}_{\mathbf{x}|\mathbf{z}})\in\mathcal{S}} 2^{n[H(\hat{\mathsf{P}}_{\mathbf{x}|\mathbf{z}})-\delta(\epsilon)]} \cdot 2^{-n[H(\mathsf{X})+\epsilon']} \tag{75}$$

$$= 2^{-n[\min_{\mathsf{Q}_{\mathsf{X}|\mathsf{Z}}\in\mathcal{S}} I(\mathsf{Q}_\mathsf{X},\mathsf{P}_{\mathsf{X}|\mathsf{Z}})+\epsilon'+\delta(\epsilon)]}. \tag{76}$$

注意，按照构造，最后一个界对每个 $z^n \in$ $()$(P) 都成立，因此，Z 的边缘分布必定在 P$\sum_{x\ y\ \in\mathcal{X}\times\mathcal{Y}}$ Z z $z|y$ $\mathsf{Y}|\mathsf{X}$ $y|x$ ( ) $\mathcal{T}_\epsilon^{\ n}$ Z = $\sum_{(,\ )}$P (z|y)Q (y|x) $_{(,\ )}$P ( )P ( ) 的附近。由此得到以下等式:

$$\min_{\mathsf{Q}_{\mathsf{X}|\mathsf{Z}}\in\mathcal{S}\cap\mathcal{T}_\epsilon^{(n)}(\mathsf{P}_\mathsf{Z})} I(\mathsf{P}_\mathsf{X},\mathsf{Q}_{\mathsf{X}|\mathsf{Z}}) = J_V(\mathsf{X};\mathsf{Z}). \tag{77}$$

这就完成了逆向证明。

Appendix D
Proof of achievability for Theorem 2

A. 原始形式

事件 $\{V^n(x^n(m)\ z^n) > V^n(x^n(1)\ z^n)\}$ 等价于事件

$$\left\{\mathbb{E}_{\hat{\mathsf{P}}_{\mathbf{x}(m)\mathbf{z}}}[\log V(\mathsf{X},\mathsf{Z})] > \mathbb{E}_{\hat{\mathsf{P}}_{\mathbf{x}(1)\mathbf{z}}}[\log V(\mathsf{X},\mathsf{Z})]\right\}. \tag{78}$$

我们接下来界定解码器处的错误概率。将某些序列对的联合经验分布记为

˜$(x^n(1)\ z^n)$ 记为 P，并将 $(x^n(2)\ z^n)$ 记为 P。

$$\mathsf{P}\{\mathcal{E}_4 \cap \mathcal{E}_2^c\} \tag{79}$$

$$= \mathsf{P}\left\{\bigcup_{m=2}^{2^{nR}} V^n(\mathsf{X}^n(m),\mathsf{Z}^n) > V^n(\mathsf{X}^n(1),\mathsf{Z}^n)\right\} \tag{80}$$

$$\leq 2^{nR}\cdot \mathsf{P}\{V^n(\mathsf{X}^n(2),\mathsf{Z}^n) > V^n(\mathsf{X}^n(1),\mathsf{Z}^n)\} \tag{81}$$

$$= \mathsf{P}\left\{\mathbb{E}_{\tilde{\mathsf{P}}}[\log V(\mathsf{X},\mathsf{Z})] > \mathbb{E}_{\mathsf{P}}[\log V(\mathsf{X},\mathsf{Z})]\right\} \tag{82}$$

$$= \sum_{\substack{\tilde{\mathsf{P}}\in\mathcal{P}_n(\mathcal{X}\times\mathcal{Z})\\ \mathbb{E}_{\tilde{\mathsf{P}}}[\log V(\mathsf{X},\mathsf{Z})]>\mathbb{E}_{\mathsf{P}}[\log V(\mathsf{X},\mathsf{Z})]}} \mathsf{P}\left\{(\mathsf{X}^n(2),\mathsf{Z}^n)\in\mathcal{T}^{(n)}\left(\tilde{\mathsf{P}}\right)\right\} \tag{83}$$

$$\leq \sum_{\substack{\tilde{\mathsf{P}}\in\mathcal{P}_n(\mathcal{X}\times\mathcal{Z})\\ \mathbb{E}_{\tilde{\mathsf{P}}}[\log V(\mathsf{X},\mathsf{Z})]>\mathbb{E}_{\mathsf{P}}[\log V(\mathsf{X},\mathsf{Z})]}} e^{-n\mathsf{D}\left(\tilde{\mathsf{P}}\|\mathsf{Q}_\mathsf{X}\times\mathsf{P}_\mathsf{Z}\right)} \tag{84}$$

$$\leq (n+1)^{|\mathcal{X}|\cdot|\mathcal{Z}|} \max_{\substack{\tilde{\mathsf{P}}\in\mathcal{P}_n(\mathcal{X}\times\mathcal{Z})\\ \mathbb{E}_{\tilde{\mathsf{P}}}[\log V(\mathsf{X},\mathsf{Z})]>\mathbb{E}_{\mathsf{P}}[\log V(\mathsf{X},\mathsf{Z})]}} e^{-n\mathsf{D}\left(\tilde{\mathsf{P}}\|\mathsf{Q}_\mathsf{X}\times\mathsf{P}_\mathsf{Z}\right)}. \tag{85}$$

因此，当 n 趋于无穷大时，$\mathsf{P}\{\mathcal{E}_4\cap\mathcal{E}_2^c\}$ 趋于零，如果

$$R < \min_{\substack{\tilde{\mathsf{P}}\in\mathcal{P}_n(\mathcal{X}\times\mathcal{Z})\\ \mathbb{E}_{\tilde{\mathsf{P}}}[\log V(\mathsf{X},\mathsf{Z})]>\mathbb{E}_{\mathsf{P}}[\log V(\mathsf{X},\mathsf{Z})]}} \mathsf{D}\left(\tilde{\mathsf{P}}\|\mathsf{Q}_\mathsf{X}\times\mathsf{P}_\mathsf{Z}\right). \tag{86}$$

至此，定理 2 的原始部分证明完毕。

B. 对偶形式

考虑式 (86) 中的最小化问题。由于目标函数是凸的（它是散度），且约束条件是线性的，因此这是一个凸优化问题。后者进一步意味着强对偶性成立[34]。相应的拉格朗日函数为

$$L(\tilde{\mathsf{P}},\lambda) = \mathsf{D}\left(\tilde{\mathsf{P}}\|\mathsf{P}_\mathsf{X}\cdot\mathsf{P}_\mathsf{Z}\right) + \lambda\mathbb{E}_{\mathsf{P}}[\log V(\mathsf{X},\mathsf{Z})] \tag{87}$$
$$- \lambda\mathbb{E}_{\tilde{\mathsf{P}}}[\log V(\mathsf{X},\mathsf{Z})]$$
$$+ \sum_{z\in\mathcal{Z}}\mu(z)\sum_{x\in\mathcal{X}}\left[\tilde{\mathsf{P}}(x,z)-\mathsf{P}_\mathsf{Z}(z)\right]$$
$$= \mathbb{E}_{\tilde{\mathsf{P}}}\left[\log\frac{\tilde{\mathsf{P}}}{\mathsf{P}_\mathsf{X}\cdot\mathsf{P}_\mathsf{Z}}\right] + \lambda\mathbb{E}_{\mathsf{P}}[\log V(\mathsf{X},\mathsf{Z})] \tag{88}$$
$$- \lambda\mathbb{E}_{\tilde{\mathsf{P}}}[\log V(\mathsf{X},\mathsf{Z})]$$
$$+ \sum_{z\in\mathcal{Z}}\mu(z)\left[\sum_{x\in\mathcal{X}}\tilde{\mathsf{P}}(x,z)-\mathsf{P}_\mathsf{Z}(z)\right].$$

对偶目标函数为

$$q(\lambda) = \min_{\tilde{\mathsf{P}}\in\mathcal{P}_n(\mathcal{X}\times\mathcal{Z})} L(\tilde{\mathsf{P}},\lambda). \tag{89}$$

我们通过确定驻点来求得最小值。

$$\frac{\partial L}{\partial\tilde{\mathsf{P}}(x,z)} = \log\frac{\tilde{\mathsf{P}}(x,z)}{\mathsf{P}_\mathsf{X}(x)\cdot\mathsf{P}_\mathsf{Z}(z)} + 1 - \lambda\log V(x,z) + \mu(z) = 0. \tag{90}$$

因此，

$$\tilde{\mathsf{P}}^*(x,z) = \frac{\mathsf{P}_\mathsf{X}(x)\mathsf{P}_\mathsf{Z}(z)V(x,z)^\lambda}{\sum_{x'\in\mathcal{X}}\mathsf{P}_\mathsf{X}(x')V(x',z)^\lambda}. \tag{91}$$

˜将 P^* 代入拉格朗日函数，我们得到以下对偶目标函数

$$q(\lambda)=\sum_{(x,z)\in\mathcal{X}\times\mathcal{Z}}\mathsf{P}(x,z)\log\frac{V(x,z)^{\lambda}}{\sum_{x'\in\mathcal{X}}\mathsf{P}_{\mathsf{X}}(x')V(x',z)^{\lambda}}. \tag{92}$$

最后，由于强对偶性，我们得到以下等式

$$I_{\mathsf{GMI}}(\mathsf{P}_{\mathsf{XZ}})=\max_{\lambda\geq 0}\sum_{(x,z)}\mathsf{P}(x,z)\log\frac{V(x,z)^{\lambda}}{\sum_{x'\in\mathcal{X}}\mathsf{P}_{\mathsf{X}}(x')V(x',z)^{\lambda}}, \tag{93}$$

其中 $I_{\mathsf{GMI}}(\mathsf{P}_{\mathsf{XZ}})$定义见式 (11)。

附录 E 定理 5 可达性证明

我们采用随机编码。令 $\epsilon'<\epsilon$。固定边缘概率质量函数 P_{X} 和 P_{Z}。由此得到 Y 的边缘概率质量函数，即 $\mathsf{P}_{\mathsf{Y}}(y)=\sum_{x\in\mathcal{X}}\mathsf{P}_{\mathsf{Y}|\mathsf{X}}(y|x)\mathsf{P}_{\mathsf{X}}(x)$。

a) 发射机的码本生成与编码：根据 P_{X}生成随机码本 $\mathcal{C}=\{x^n(m)\}$, $m\in[1:2^{nR}]$, 码率为 R 和 n。该码本会提供给编码器和解码器，但不会提供给中继。

b) 失配中继的码本生成：根据 P_{Z}生成随机码本 $\mathcal{C}_R=\{z^n(w)\}$, $w\in[1:2^{nB}]$，码率为 B 和 n。该码本会提供给中继和解码器，但不会提供给发射机。

c) 失配中继编码：观察到 y^n后，中继找到一个索引 $\hat{}w\in[1:\ 2^{nB}]$，使得

$$\hat{w}=\arg\min_{w\in\mathcal{W}}d_0^n(y^n,z^n(w)). \tag{94}$$

中继将选定的索引 $\hat{}w$ 传递给解码器。

d) 解码：请注意，解码器不知道从 X 到 Z 的真实信道实现。由于中继处的编码器采用固定的解码度量，而非联合典型性编码，因此不会像基于典型性的方案通常那样产生固有的“测试信道”。这种情况类似于通用解码器的设置，在这种设置下，解码器只知道码本。有鉴于此，我们提出以下解码器。从中继处获得 $\hat{}w$ 后，接收方找到一个唯一的消息 $\hat{}m$，使得

$$\hat{m}=\arg\max_{m\in\mathcal{M}}I(\hat{\mathsf{P}}_{\mathbf{x}(m)\mathbf{z}(\hat{w})}), \tag{95}$$

其中，$I(\hat{}\mathsf{P}_{\mathbf{x}(m)\mathbf{z}(\hat{w})})$表示针对 m号码字 $x^n(m)$与 $z^n(\hat{}w)$的联合经验分布计算出的互信息。如果 $\hat{}m\neq m$，则宣告发生错误事件 $\mathcal{E}$。

e) 失配中继性能分析：中继观测到 y^n。由于 $\mathsf{P}_{\mathsf{Y}^n}(y^n)=\prod_{i=1}^n\mathsf{P}_{\mathsf{Y}}(y_i)$，根据大数定律，以高概率有 $y^n\in\mathcal{T}_{\epsilon}^{(n)}(\mathsf{P}_{\mathsf{Y}})$。我们将 $\mathsf{P}_{\mathsf{Z},n}\in\mathcal{P}_n(\mathcal{Z})$ 定义为一个任意类型，它与 P_{Z} 具有相同的支撑集，并满足：$\|\mathsf{P}_{\mathsf{Z}}-\mathsf{P}_{\mathsf{Z},n}\|_{\infty}\leq\frac{1}{n}$。注意，对于每个 $w\in\mathcal{W}$, $z^n(w)\in\mathcal{T}_{\epsilon}^{(n)}(\mathsf{P}_{\mathsf{Z}})$，都有 $\hat{}\mathsf{P}_{z^n(w)}=\mathsf{P}_{\mathsf{Z},n}$。固定 y^n和 w，并考虑失真度量：

$$d_0(y^n,z^n(w))=\frac{1}{n}\sum_{i=1}^n d_0(y_i,z_i)=\mathbb{E}_{\hat{\mathsf{P}}_{\mathbf{yz}(w)}}\left[d_0(\mathsf{Y},\mathsf{Z})\right]. \tag{96}$$

$\hat{}$ 接下来我们将证明，$\mathsf{P}_{\mathbf{yz}}$ 必须满足某种特定结构。

分析的其余部分沿用 [13, 第 4 节5]中的分析。定义以下集合:

$$\mathcal{S}_n^{(-)}=\left\{\mathsf{Q}_{\mathsf{YZ}}\in\mathcal{P}^n_{\mathcal{Y}\times\mathcal{Z}}:\mathsf{Q}_{\mathsf{Y}}=\mathsf{P}_{\mathsf{Y}},\mathsf{Q}_{\mathsf{Z}}=\mathsf{P}_{\mathsf{Z},n},I(\mathsf{Q}_{\mathsf{YZ}})\geq B+\delta\right\},$$

and

$$\mathcal{S}_n^{(+)}=\left\{\mathsf{Q}_{\mathsf{YZ}}\in\mathcal{P}^n_{\mathcal{Y}\times\mathcal{Z}}:\mathsf{Q}_{\mathsf{Y}}=\mathsf{P}_{\mathsf{Y}},\mathsf{Q}_{\mathsf{Z}}=\mathsf{P}_{\mathsf{Z},n},I(\mathsf{Q}_{\mathsf{YZ}})\leq B-\delta\right\},$$

其中，$\mathcal{P}^n_{\mathcal{Y}\times\mathcal{Z}}$ 是 $\mathcal{Y}\times\mathcal{Z}$ 上与某个长度为 n 的序列 $(y^n,\ z^n)$ 的联合类型相对应的联合概率质量函数子集。

引理 6：对偶 $(y^n,z^n(w))$ 的联合类型的并集满足

1) $\lim_{n\to\infty}\mathsf{P}\left\{\bigcup_{w=1}^{2^{nB}}\hat{\mathsf{P}}_{\mathbf{yz}(w)}\in\mathcal{S}_n^{(-)}\right\}=0.$
2) $\lim_{n\to\infty}\mathsf{P}\left\{\bigcup_{w=1}^{2^{nB}}\hat{\mathsf{P}}_{\mathbf{yz}(w)}\in\mathcal{S}_n^{(+)}\right\}=1.$

Proof: Cary,

$$\mathsf{P}\left\{\bigcup_{w=1}^{2^{nB}}\left\{(y^n,\mathsf{Z}^n(w))\in\mathcal{T}^{(n)}(\mathsf{Q}_{\mathsf{YZ}})\right\}\right\}$$

$$=1-\prod_{w=1}^{2^{nB}}\mathsf{P}\left\{(y^n,\mathsf{Z}^n(w))\notin\mathcal{T}^{(n)}(\mathsf{Q}_{\mathsf{YZ}})\right\} \tag{97}$$

$$=1-\left[1-\mathsf{P}\left\{(y^n,\mathsf{Z}^n)\in\mathcal{T}^{(n)}(\mathsf{Q}_{\mathsf{YZ}})\right\}\right]^{2^{nB}}. \tag{98}$$

注意，$(1-x)^k\leq e^{-kx}$ 对于 $x\in[0,1]$ 和 $k\geq 0$。此外，

$$\lim_{n\to\infty}\left(1+\frac{z}{n}\right)^n=e^z. \tag{99}$$

考虑以下几种情况:

- 若 $\mathsf{P}\left\{(y^n,\mathsf{Z}^n)\in\mathcal{T}^{(n)}(\mathsf{Q}_{\mathsf{YZ}})\right\}\geq 2^{-n(B-\epsilon)}$，则 (98) 右侧的下界为

$$\mathsf{P}\left\{\bigcup_{w=1}^{2^{nB}}\left\{(y^n,\mathsf{Z}^n(w))\in\mathcal{T}^{(n)}(\mathsf{Q}_{\mathsf{YZ}})\right\}\right\}$$
$$\geq 1-[(1-2^{-n(B-\epsilon)})^{2^{n(B-\epsilon)}}]^{2^{n\epsilon}}. \tag{100}$$

令 n 趋于 ∞，可得

$$\lim_{n\to\infty}\mathsf{P}\left\{\bigcup_{w=1}^{2^{nB}}\left\{(y^n,\mathsf{Z}^n(w))\in\mathcal{T}^{(n)}(\mathsf{Q}_{\mathsf{YZ}})\right\}\right\}$$
$$\geq\lim_{n\to\infty}1-[(1-2^{-n(B-\epsilon)})^{2^{n(B-\epsilon)}}]^{2^{n\epsilon}} \tag{101}$$
$$=\lim_{n\to\infty}1-e^{-2^{n\epsilon}}=1. \tag{102}$$

- 若 $\mathsf{P}\left\{(y^n,\mathsf{Z}^n)\in\mathcal{T}^{(n)}(\mathsf{Q}_{\mathsf{YZ}})\right\}\leq 2^{-n(B+\epsilon)}$，则 (98) 的右端不超过

$$\mathsf{P}\left\{\bigcup_{w=1}^{2^{nB}}\left\{(y^n,\mathsf{Z}^n(w))\in\mathcal{T}^{(n)}(\mathsf{Q}_{\mathsf{YZ}})\right\}\right\}$$
$$\leq 1-[(1-2^{-n(B+\epsilon)})^{2^{n(B+\epsilon)}}]^{2^{-n\epsilon}}. \tag{103}$$

取 n 的极限趋于 ∞，我们有

$$\lim_{n\to\infty} \mathsf{P}\left\{\bigcup_{w=1}^{2^{nB}} \left\{(y^n, \mathsf{Z}^n(w)) \in \mathcal{T}^{(n)}(\mathsf{Q}_{\mathsf{YZ}})\right\}\right\}$$

$$\leq \lim_{n\to\infty} 1 - [(1-2^{-n(B+\epsilon)})^{2^{n(B+\epsilon)}}]^{2^{-n\epsilon}} \quad (104)$$

$$= \lim_{n\to\infty} 1 - e^{-2^{-n\epsilon}} = \lim_{n\to\infty} 1 - 1 + 2^{-n\epsilon} = 0. \quad (105)$$

ˆ 此外，注意到由于 $\mathsf{Q}_{\mathsf{Y}} = \mathsf{P}_{y^n}$ 和 $\mathsf{Q}_{\mathsf{Z}} = \mathsf{P}_{\mathsf{Z},n}$，我们有

$$2^{-n[I(\mathsf{Q}_{\mathsf{YZ}})+\delta]} \leq \mathsf{P}\left\{(y^n, \mathsf{Z}^n) \in \mathcal{T}^{(n)}(\mathsf{Q}_{\mathsf{YZ}})\right\} \leq 2^{-n[I(\mathsf{Q}_{\mathsf{YZ}})-\delta]}. \quad (106)$$

因此，若 $I(\mathsf{Q}_{\mathsf{YZ}}) \leq B - \delta'$，则

$$\lim_{n\to\infty} \mathsf{P}\left\{\bigcup_{w=1}^{2^{nB}} \hat{\mathsf{P}}_{\mathbf{yz}(w)} \in \mathcal{S}_n^{(+)}\right\} = 1. \quad (107)$$

否则，若 $I(\mathsf{Q}_{\mathsf{YZ}}) \geq B + \delta'$，则

$$\lim_{n\to\infty} \mathsf{P}\left\{\bigcup_{w=1}^{2^{nB}} \hat{\mathsf{P}}_{\mathbf{yz}(w)} \in \mathcal{S}_n^{(-)}\right\} = 0. \quad (108)$$

■

因此，以高概率有 $\hat{}\mathsf{P}_{\mathbf{yz}} \in \mathcal{S}_n^{(+)}$。

f) 解码器处错误概率的分析：注意，由于码本生成具有对称性

$$\mathsf{P}\left\{\hat{\mathsf{M}} \neq \mathsf{M}\right\} = \mathsf{P}\left\{\hat{\mathsf{M}} \neq \mathsf{M} | \mathsf{M} = 1\right\}$$

$$\leq \mathsf{P}\left\{I(\hat{\mathsf{P}}_{\mathsf{X}^n(1)\mathsf{Z}^n(W)}) < \theta | \mathsf{M} = 1\right\} \quad (109)$$

$$+ \mathsf{P}\left\{I(\hat{\mathsf{P}}_{\mathsf{X}^n(m)\mathsf{Z}^n(W)}) \geq \theta \text{ for some } m \neq 1 | \mathsf{M} = 1\right\}$$

$$\leq \mathsf{P}\left\{I(\hat{\mathsf{P}}_{\mathsf{X}^n(1)\mathsf{Z}^n(W)}) < \theta | \mathsf{M} = 1\right\}$$

$$+ 2^{nR} \cdot \mathsf{P}\left\{I(\hat{\mathsf{P}}_{\mathsf{X}^n(2)\mathsf{Z}^n(W)}) \geq \theta | \mathsf{M} = 1\right\}. \quad (110)$$

这意味着我们不失一般性地假设发送的是 $\mathsf{M} = 1$。考虑第一项。注意，对于每个 $x^n \in \mathcal{T}_{(n)\,\epsilon''}(\mathsf{P}_{\mathsf{X}})$、$y^n \in \mathcal{T}_{(n)\,\epsilon'}(\mathsf{P}_{\mathsf{Y}|\mathsf{X}})$ 以及 $z^n \in \mathcal{T}_{(n)}(\mathsf{Q}_{\mathsf{Z}|\mathsf{Y}})$，我们有

$$\hat{\mathsf{P}}_{\mathbf{xyz}}(a,b,c) \quad (111)$$

$$= \frac{\sum_{i=1}^n \mathbf{1}\{x_i = a, y_i = b, z_i = c\}}{n} \quad (112)$$

$$= \frac{\sum_{i=1}^n \mathbf{1}\{x_i = a\}}{n} \frac{\sum_{i=1}^n \mathbf{1}\{x_i = a, y_i = b, z_i = c\}}{\sum_{i=1}^n \mathbf{1}\{x_i = a\}} \quad (113)$$

$$= \hat{\mathsf{P}}_{\mathbf{x}}(a)\hat{\mathsf{P}}_{\mathbf{y}|\mathbf{x}}(b|a)\mathsf{Q}_{\mathsf{Z}|\mathsf{Y}}(c|b). \quad (114)$$

ˆˆ 因此，$\mathsf{P}_{x^n(1)z^n(W)} = \mathsf{P}_{x^n(1)y^n} \circ \mathsf{Q}_{\mathsf{Z}|\mathsf{Y}}$，并且

$$\lim_{n\to\infty} I(\hat{\mathsf{P}}_{x^n(1)z^n}) = I(\mathsf{P}_{\mathsf{XY}} \circ \mathsf{Q}_{\mathsf{Z}|\mathsf{Y}}). \quad (115)$$

因此，

$$\lim_{n\to\infty} I(\hat{\mathsf{P}}_{\mathsf{X}^n(1)\mathsf{Z}^n(W)}) \geq \min_{\mathsf{Q}_{\mathsf{Z}|\mathsf{Y}} \in \mathcal{S}_n^{(+)}} I(\mathsf{P}_{\mathsf{XY}} \circ \mathsf{Q}_{\mathsf{Z}|\mathsf{Y}}) > \theta. \quad (116)$$

在这种情况下，$\lim_{n\to\infty} \mathsf{P}\left\{I(\mathsf{P}_{\mathsf{X}^n(1)\mathsf{Z}^n(W)}) < \theta | \mathsf{M} = 1\right\} = 0$。考虑第二项。令

$$\mathcal{A}_n \triangleq \left\{\mathsf{Q}_{\mathsf{XZ}} \in \mathcal{P}^n_{\mathcal{X}\times\mathcal{Z}} \colon I(\mathsf{Q}_{\mathsf{XZ}}) \geq \theta\right\}. \quad (117)$$

由于对于 $m \neq 1$，我们有 P

$\prod_{i=1}^n$ X^n m ^{n} W $= {}_{()\mathsf{Z}()}(x^n, z^n)$ P ( )P ( ), 因此根据 Sanov 定理（定理 11.4）。1] [35,

$$\mathsf{P}\left\{I(\mathsf{P}_{\mathbf{X}(m)\mathbf{Z}(W)}) \geq \theta | \mathsf{M} = 1\right\} \quad (118)$$

$$\leq (n+1)^{|\mathcal{X}|\cdot|\mathcal{Z}|} 2^{-n \min_{\mathsf{Q}_{\mathsf{XZ}} \in \mathcal{A}_n} D(\mathsf{Q}_{\mathsf{XZ}} \| \mathsf{P}_{\mathsf{X}} \mathsf{P}_{\mathsf{Z}})} \quad (119)$$

$$\leq (n+1)^{|\mathcal{X}|\cdot|\mathcal{Z}|} 2^{-n\theta}. \quad (120)$$

因此，如果 $R < \theta$，第二项会在 $n \to \infty$ 时趋于零。定理的证明至此完成。