基于最大流的低速率界限
噪声网络中的信息传播

Yan Hao Ling 和 Jonathan Scarlett

arXiv:2304.02226v1 [cs.IT] 5 Apr 2023

摘要

We study error exponents for the problem of low-rate communication over a directed graph, where each edge in the graph represents a noisy communication channel, and there is a single source and destination. We derive maxflow-based achievability and converse bounds on the error exponent that match when there are two messages and all channels satisfy a symmetry condition called pairwise reversibility. More generally, we show that the upper and lower bounds match to within a factor of 4. We also show that with three messages there are cases where the maxflow-based error exponent is strictly suboptimal, thus showing that our tightness result cannot be extended beyond two messages without further assumptions.

一、引言

近期，噪声网络中的低速率（例如 1 比特）信息传播问题日益受到关注，以下是其中一些典型例子:

- The transmission of a single bit over a tandem of channels (the 2-hop setting) was studied in [1], [2], [3], and the multi-bit variant was studied in [4]. The goal is to characterize the associated error exponent; as we further discuss below, this line of works is the one most relevant to the present paper.
- The transmission of a single bit over a long chain of noisy channels (connected by relays) dates back to the work of Schulman and Rajagopalan [5] and recently regained attention [1], [6], [7] after being posed by Polyanskiy (who termed it the *information velocity* problem) and connected to the 2-hop setting by Huleihel, Polyanskiy, and Shayevitz [1].
- Another line of works studied *broadcasting* problems on various kinds of graphs, such as trees [8], random graphs [9], and grids [10]. Here the goal is to reconstruct the original bit from the information received at the leaves, and the focus has been on simple intermediate nodes performing memoryless operations. (In contrast, the problems described above allow the intermediate/relay nodes to perform complicated coding strategies.)

The authors are with the Department of Computer Science and Department of Mathematics, School of Computing, National University of Singapore (NUS). Jonathan Scarlett is also with the Institute of Data Science, NUS. Emails: lingyh@nus.edu.sg; scarlett@comp.nus.edu.sg
This work was supported by the Singapore National Research Foundation (NRF) under grant number R-252-000-A74-281.

In this paper, we address a notable gap in this list by extending the 2-hop problem to transmission over a fixed but arbitrary directed graph, with a single source and a single destination.[1] An extensive range of channel capacity results are known for such settings [11, Ch. 15], but 1-bit and other low-rate settings have remained less explored.

Specifically, we generalize the 2-hop multi-bit setup to arbitrary networks and show a maxflow-based achievability bound (Theorem 1), which supersedes our main previous multi-bit result for the 2-hop setting [4]. We also provide a converse bound (Theorem 2) based on the cutset idea, which matches the achievability bound in the 1-bit case when every channel satisfies a symmetry condition called pairwise reversibility (see Definition 7 below). While we are unable to find the optimal error exponent more generally, our achievability and converse bounds provide a 4-approximation (Lemma 32).

我们接下来将更详细地概述相关工作，然后正式描述问题设置并陈述我们的主要结果。

A. 相关工作

Tandem of channels (2-hop). The problem of transmitting a *single bit* over a tandem of channels was introduced by Huleihel, Polyanskiy, and Shayevitz [1], as well as Jog and Loh using different motivation/terminology based on teaching and learning in multi-agent problems [2]. Both of these works gave various protocols with varying trade-offs between their simplicity and their associated error exponents. We then showed in [3] that the 1-hop and 2-hop error exponents coincide whenever the two channels have the same transition law. In [4], we generalized this problem to the multi-bit setting and showed that whenever both channels are pairwise reversible, the 2-hop error exponent is simply given by the bottleneck (i.e., the worse of the two 1-hop error exponents).

Information velocity (many-hop). As we outlined above, the information velocity problem seeks to establish the minimum possible ratio of the number of hops to the block length to reliably send a single bit over a long chain of relays connected by binary symmetric channels (or more generally, other channels). This problem also has interesting connections to the 2-hop setting [1].

A general result of Schulman and Rajagopalan [5] translates noiseless protocols (on arbitrary graphs) to noisy ones, and when specialized to the line graph, this proves that positive information velocity is attainable. We gave a simple and efficient protocol achieving positive information velocity in [6], though precisely characterizing the best possible information velocity remains very much open. The information velocity of erasure-type channels was studied in [7], and for this class of channels an exact characterization was attained.

Broadcasting problems. As mentioned above, problems of broadcasting [8], [9], [10] have a similar flavor to the information velocity problem, but with some substantial differences including (i) the consideration of more complicated graphs, (ii) the combination of many nodes' information to form the final estimate, and (iii) the consideration of simple memoryless intermediate nodes without the use of coding.

Capacity results on graphs. Another notable line of works has sought to characterize the channel capacity of noisy networks, as summarized in [12, Sec. 15.10] and [11, Ch. 15]. The study of capacity is largely quite

[1]一般图也适用于 [5], 这一框架，即使不要求其大小相对于分组长度保持固定；但该框架仅关注渐近标度规律，而不关注精确常数或错误指数。

distinct from that of 1-bit (or other low-rate) communication, but we will see some similarities throughout the paper. In particular, our results are based on maxflow and mincut ideas, which are widely used in capacity results. Instead of using 1-hop channel capacities as edge weights for the maxflow problem, we will use 1-hop error exponents. Our protocols and analysis will be substantially different compared to capacity results and will require further assumptions such as pairwise reversibility, but we will use an existing idea of replacing individual edges by multiple parallel edges of smaller weight in order to form edge-disjoint paths (e.g., see [11, Sec. 15.3]).

II. 问题设置

We first describe the classical channel coding setup on a discrete memoryless channel (DMC) P, whose alphabets are denoted by $\mathcal{X}$ and $\mathcal{Y}$ (or sometimes $\mathcal{X}_P$ and $\mathcal{Y}_P$). The source node is given a message $m \in \{1, 2, \ldots, M\}$, and has n uses of the channel P to send information across to the destination node, after which the destination node guesses the value of m. An error is said to have occurred if the destination's guess is different from m. The minimum probability of an error across all protocols with n time steps is denoted by $p_{\text{err}}(n, M, P)$, and the error exponent of the channel with M messages is defined as

$$\mathcal{E}_{M,P} = \liminf_{n\to\infty} -\frac{1}{n} p_{\text{err}}(n, M, P), \tag{1}$$

其中，极限是在 M 保持固定时令 $n \to \infty$ 趋于极限。只要上下文中省略的参数明确，我们就使用 $\mathcal{E}$、$\mathcal{E}_M$ 或 $\mathcal{E}_P$。

In this paper, we consider a communication over a directed graph with edges representing noisy channels. Specifically, we are given a directed graph G (not necessarily acyclic), with two special nodes called source and destination. Without loss of generality, we assume that there exists at least one path from the source to the destination. For every edge e, there is a DMC P_e. We note that for a given pair of nodes, say (i, j), we allow for the possibility of multiple parallel edges connecting i to j (though this could equivalently be viewed as a single edge whose associated channel is a product channel).

At time 0, the source node receives a message $m \in \{1, 2, \ldots, M\}$. For each time step, every node gets one use of the channel associated with its outgoing edge. This process occurs for a total of n time steps. After n time steps, the destination node forms an estimate of the message; let $p_{\text{err}}(n, M, G)$ be the resulting error probability.[2] We are interested in the error exponent of the network, given by

$$\mathcal{E}_{M,G} = \liminf_{n\to\infty} -\frac{1}{n} \log p_{\text{err}}(n, M, G). \tag{2}$$

注意，我们取 $n \to \infty$ 的极限，同时 M 和 G 保持不变。上下文明确时，我们将简单地记作 $\mathcal{E}_G$。

For an edge e, we let $\mathcal{E}_e$ be short for $\mathcal{E}_{P_e}$ (i.e., the 1-hop error exponent of the associated channel). For a fixed (M, G) pair, we can create a network flow problem (see Section III-A) that assigns a capacity[3] of $\mathcal{E}_e$ on every edge

[2]Our results are unaffected by whether this error probability is averaged over all m or taken with respect to the worst-case m, as is the case for the vast majority of error exponent studies (e.g, [13], [14]).

[3]This is "capacity" in a generic sense, and not the Shannon capacity.

e, and consider the maximum flow across this network. We will refer to this as the maxflow for the (M, G) pair, and denote it by $|f_{M,G}|$. We will also shorten this to $|f|$ when the context is clear.

Our main result is the following achievability result for channels that are *pairwise reversible*; this is a classical notion of symmetry [14] that notably includes the binary symmetric channel (BSC), binary erasure channel (BEC), and their generalizations to non-binary inputs. Briefly, the property is that the quantity $\sum_y P(y|x)^{1-s}P(y|x')^s$ attains its minimum at $s = \frac{1}{2}$ for all (x, x'), which implies that the Chernoff distance between two inputs simplifies to the Bhattacharyya distance. See Definition 7 below for further details, and [14] for a more comprehensive treatment with numerous examples.

定理 1。若 G 中的每个信道均为两两可逆，则错误指数 $|f_{M,\ G}|$ 可实现，即，

$$\mathcal{E}_{M,G} \geq |f_{M,G}|. \tag{3}$$

证明。我们在第四节给出串联情形（即线图）的证明，并以此为基础在第五节证明一般情形。

□

我们注意到，这是一个信息论可达性结果，我们并不试图建立相应的计算高效协议（尽管这样做会很有意义）。相比之下，我们关于 2 跳设置的相关工作包含了基于高效和低效协议的各种结果 [3], [4]。

我们还证明了以下反向界：

定理 2。对于任意一对 (M, G)，我们有

$$\mathcal{E}_{M,G} \leq |f_{2,G}|. \tag{4}$$

证明。见第 VI 节。 □

结合定理 1和定理 2，我们证明，当每个信道都是两两可逆且 $M = 2$ 时等式成立：

推论 3. 如果 G 中的每个信道都是两两可逆的，那么

$$\mathcal{E}_{2,G} = |f_{2,G}|. \tag{5}$$

We are also interested in the zero-rate error exponent. Momentarily returning to classical 1-hop communication, for $R > 0$, we define the error exponent of the channel P at rate R by

$$E_P(R) = \liminf_{n\to\infty} \left\{ -\frac{1}{n} \log p_{\text{err}}(n, e^{nR}, P) \right\}, \tag{6}$$

and we extend this function to $R = 0$ via $E_P(0) = \lim_{R\to 0^+} E_P(R)$. Analogously, in the setting of the present paper, we let $E_G(R)$ and $E_G(0)$ denote the zero-rate error exponent of the network. We can again consider the network flow problem which assigns a capacity of $E_e(0)$ on every edge e, which we will denote by $|f_G(0)|$ or sometimes simply $|f|$. The following result extends Theorem 1 to zero-rate error exponents, even without requiring the pairwise reversible assumption:

定理 4. 对于任意 G，其信道不必两两可逆，我们有

$$E_G(0) \geq |f_G(0)|. \tag{7}$$

证明。见第 VII-B 节。 □

其他结果见第 VII 节和第 VIII 节，简要概述如下：

- 引理 27 给出了信道不一定具有成对可逆性时的可达性结果；其思路是在图中的每条边上使用较弱的“对称化”错误指数，而非最优的单跳指数。
- 在定理 29 中，我们给出了若干条件（尽管有些限制性），在这些条件下，可以为满足 $M > 2$ 的成对可逆信道得到相应的逆定理。
- 在引理 32 中，我们证明对于一般的 M 和 G，最优指数的误差至多为 4 倍；在引理 33 中，我们给出将该误差改进至 2 倍的条件。
- In Theorem 35, we show that in general there exist scenarios in which the maxflow-based error exponent is strictly suboptimal when $M = 3$, even under the constraint of an acyclic graph and pairwise reversible channels. Thus, we place a strong restriction on the extent to which our tightness result for $M = 2$ can be generalized.

III. 预备知识

在本节中，我们引入一些额外的记号与定义，并给出若干有用的辅助结果，这些结果将用于证明我们的主要结论。

A. 流与割

我们暂时偏离上述设定，考虑图上的一般概念。考虑一个有向图$G = (V, E)$，其中有两个特殊节点，称为源点 v_s 和汇点 v_t。假设每条边 e上都有一个非负实值容量，记为 c_e。

令 $f = \{f_e\}_{e\in E}$ 是一个从 E 到 $\mathbb{R}_{\geq 0}$的函数。如果满足以下条件，我们称 f 是一个流。 :

- (容量约束) 对于每个 e， $0 \leq f_e \leq c_e$。
- (Flow conservation) For each $v \in V$, except for the source and destination, we have

$$\sum_{u:(u,v)\in E} f_{u,v} = \sum_{v:(v,w)\in E} f_{v,w}. \tag{8}$$

Define the size of a flow f as

$$|f| = \sum_{u:(u,v_s)\in E} f_{u,v_s}. \tag{9}$$

流 f是极大的，当对于任意其他流 f'， $|f| \geq |f'|$。

与流密切相关的是割的概念。割 c 是将 V 划分为两个集合 (A, B)，使得 $v_s \in A$且 $v_t \in B$。割 c 的大小定义为

$$|c| = \sum_{u\in A, v\in B} f_{u,v}, \tag{10}$$

其中我们强调，这只计算从 A 到 B 的边，而不计算反向边。若对于任意割 c'、 $|c| \leq |c'|$，割 c 是极小的。现在我们陈述著名的最大流-最小割对偶定理:

定理 5.（最大流-最小割对偶定理，例如 [15]）最大流的大小等于最小割的大小。

以下结果同样众所周知，但我们为感兴趣的读者提供一个简短的证明。

Theorem 6. (Flow decomposition theorem, e.g., [15]) *Let f be any flow. There exist finitely many simple paths*[4] *from v_s to v_t, denoted by $p_1, p_2, \ldots, p_k$, with associated flows $f_1, f_2, \ldots, f_k$, such that*

- 每个 f_i 仅沿路径 p_i流动。换言之，对于所有 $e \in p_i$ ，$(f_i)_e = |f_i|$ ，否则为 $(f_i)_e = 0$ 。
- 所有 f_i 的总和等于 f。换句话说，对于所有 e ，我们有

$$\sum_i (f_i)_e = f_e. \tag{11}$$

Proof. We describe a procedure for iteratively building up a list of (p_i, f_i) pairs. Given f, let p_i be any path that has non-zero flow across every edge. Let $|f_i|$ be the smallest among these (non-zero) values, and let f_i be the flow formed by pushing $|f_i|$ units of flow along the path p_i. We then append (p_i, f_i) to the list of pairs, subtract $|f_i|$ along every edge on p_i, and repeat the preceding steps until no flow remains. Since the number of active edges (i.e., those with non-zero flow) strictly decreases on every step of this process, it must terminate and produce a finite number of paths, as desired. □

B. Chernoff 散度

Let $P(y|x)$ be a generic discrete memoryless channel. To lighten notation, we let $P_x(\cdot)$ denote the conditional distribution $P(\cdot|x)$. An (M, ℓ)-*codebook* is defined to be a collection of M codewords each having length ℓ, and when utilizing such a codebook, we will use the notation $(x^{(1)}, \ldots, x^{(M)})$ for the associated codewords.

对于某个有限集合 $\mathcal{X}$ 上的两个概率分布 Q、Q'，巴氏距离定义为

$$d_{\mathrm{B}}(Q, Q') = -\log \sum_{x\in\mathcal{X}} \sqrt{Q(x)Q'(x)}. \tag{12}$$

[4]如果一条路径从不重复访问同一节点，则称其为简单路径。

对于随机变量 X_1、X_2，我们也用 d_{B} 表示其分布 P_{X_1} 和 P_{X_2} 之间的巴氏亚距离，即。

$$d_{\mathrm{B}}(X_1, X_2) = d_{\mathrm{B}}(P_{X_1}, P_{X_2}), \tag{13}$$

and 对于 x、$x' \in \mathcal{X}_P$，我们定义与两个信道输入 a 相关的巴氏距离 s

$$d_{\mathrm{B}}(x, x', P) = d_{\mathrm{B}}(P_x, P_{x'}) \tag{14}$$

略微滥用记号。

Gen推广 Bhattacharyya 距离，参数为 s 的 Chernoff 散度定义为

$$d_{\mathrm{C}}(Q, Q', s) = -\log \sum_{x \in \mathcal{X}} Q(x)^{1-s} Q'(x)^s, \tag{15}$$

而 Chernoff 散度（对 s 进行优化后）为

$$d_{\mathrm{C}}(Q, Q') = \max_{0 \le s \le 1} d_{\mathrm{C}}(Q, Q', s). \tag{16}$$

类似于 (13)–(14)，我们还写出

$$d_{\mathrm{C}}(X_1, X_2) = d_{\mathrm{C}}(P_{X_1}, P_{X_2}), \tag{17}$$

$$d_{\mathrm{C}}(x, x', P) = d_{\mathrm{C}}(P_x, P_{x'}). \tag{18}$$

对于某一类信道，量 d_{C} 和 d_{B} 相同，从而得到如下定义：

定义7。 [14] 对于所有 x、$x' \in \mathcal{X}_P$，离散无记忆信道是成对可逆的，

$$d_{\mathrm{B}}(x, x', P) = d_{\mathrm{C}}(x, x', P), \tag{19}$$

这发生在该量

$$\sum_{y \in \mathcal{Y}_P} P(y|x)^{1-s} P(y|x')^s \tag{20}$$

在 $s = \frac{1}{2}$ 处达到其最小值。

The pairwise reversibility assumption allows for a number of useful results and constructions that are not available for general channels. In particular, Lemma 12 below gives an explicit expression for the 1-hop error exponent, and Lemma 14 below is based on an explicit construction that attains that exponent. We again refer the reader to [14] for a detailed discussion on pairwise reversibility with several examples.

对于任意 p 正整数 k，我们令 P^k 表示 P 的 k 重积，其概率质量函数

$$P^k(\vec{y}|\vec{x}) = \prod_{i=1}^{k} P(y_i|x_i). \tag{21}$$

For two sequences $\vec{x}, \vec{x}'$ of length k, we also use the notation $d_{\mathrm{B}}(\vec{x}, \vec{x}', P^k)$ and $d_{\mathrm{C}}(\vec{x}, \vec{x}', P^k)$ similarly to (14) and

(18), with the understanding that $\vec{x}, \vec{x}'$ are treated as inputs to P^k. We will use a well-known tensorization property of d_{C} (and hence also d_{B}), stated as follows.

引理 8. (　例如，参见 [4, 引理 8]) 对于任意序列 $\vec{x} = (x_1, \ldots, x_k)$ 和 $\vec{x}' = (x'_1, \ldots, x'_k)$，我们有　*e*

$$d_{\mathrm{C}}(\vec{x}, \vec{x}', P^k, s) = \sum_{i=1}^{k} d_{\mathrm{C}}(x_i, x'_i, P, s), \tag{22}$$

and

$$d_{\mathrm{C}}(\vec{x}, \vec{x}', P^k) = \max_{0 \le s \le 1} \sum_{i=1}^{k} d_{\mathrm{C}}(x_i, x'_i, P, s). \tag{23}$$

C. 1 跳误差指数

我们将使用 [14] 中关于 1 跳错误指数的若干结果。首先陈述两个密切相关的结果，它们隐含在 [14, 定理 2] 及其证明中。

引理9。（隐含于 [14, 定理 2]）令 $\vec{y} \in \mathcal{Y}^n$ 为信道 P 使用 n 次所接收的序列，对于任意固定的 M 以及任何由 n 索引的码本序列，

$$\liminf_{n\to\infty} -\frac{1}{n} \log p_{\mathrm{err}} \ge \liminf_{n\to\infty} \min_{m_1 \ne m_2} d_{\mathrm{C}}((\vec{y}|m = m_1), (\vec{y}|m = m_2)). \tag{24}$$

Theorem 10. (Implicit in [14, Thm. 2]) *For any $\mathcal{E}^\dagger < \mathcal{E}_M^{(1)}$, it holds for all sufficiently large ℓ that there exists an (M, ℓ)-codebook $\mathcal{C}$ such that for all pairs of codewords $(\vec{x}, \vec{x}')$ in $\mathcal{C}$, we have*

$$d_{\mathrm{C}}(\vec{x}, \vec{x}', P^\ell) \ge \ell \cdot \mathcal{E}^\dagger. \tag{25}$$

接下来，我们分别给出零速率和固定-M 设置下错误指数的显式公式/界。

定理 11. [14, 定理 4] 对于任意 P，零速率错误指数为

$$E_P(0) = \max_{q \in \mathcal{P}(\mathcal{X})} \sum_{x, x' \in \mathcal{X}} q_x q_{x'} d_{\mathrm{B}}(x, x', P). \tag{26}$$

引理 12. [14, 定理 3] 对于任意 (M, P)，我们有

$$\mathcal{E}_{M,P} \ge \frac{2}{M(M-1)} \max_{(x_1, \ldots, x_M) \in \mathcal{X}^M} \sum_{1 \le m_1 < m_2 \le M} d_{\mathrm{B}}(x_{m_1}, x_{m_2}, P), \tag{27}$$

当 P 两两可逆时取等号。

In accordance with Lemma 12, we will use the following definition:

Definition 13. *Define the Bhattacharyya distance based error exponent by*

$$\tilde{\mathcal{E}}_{M,P} = \frac{2}{M(M-1)} \max_{(x_1,\ldots,x_M)\in\mathcal{X}^M} \sum_{1\le m_1<m_2\le M} d_{\mathrm{B}}(x_{m_1}, x_{m_2}, P), \tag{28}$$

˜并注意到 $\mathcal{E}_{M,P} = \mathcal{E}_{M,P}$ 对于成对可逆信道。

在引理 12, [14] 的证明中还证明了以下内容:

引理 14. 对于任意离散无记忆信道 P，存在一个(M,ℓ)-码本 $(\vec{x}_1,\ldots,\vec{x}_M)$，满足 $\ell = M!$，使得对于任意一对$m_1 \neq m_2$，都有

$$d_{\mathrm{B}}(\vec{x}_{m_1}, \vec{x}_{m_2}, P^\ell) = \ell \cdot \tilde{\mathcal{E}}_{M,P}. \tag{29}$$

由于我们将使用乘积信道上的码，下面的引理很有用:

Lemma 15. *Let ℓ be a fixed integer and $\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_M$ be length-ℓ codewords, and let Q be the restriction of P^ℓ to these inputs. Then $\mathcal{E}_P \ge \frac{1}{\ell}\mathcal{E}_Q$. Furthermore, if P is pairwise reversible, then Q is also pairwise reversible.*

Proof. Let $\vec{w}_1, \ldots, \vec{w}_M$ be codewords achieving an error exponent arbitrarily close to $\mathcal{E}_Q$. Since each character in $\vec{w}_i$ is a string of ℓ characters over X, we can expand the strings $\vec{w}_1, \ldots, \vec{w}_M$ to form $\vec{w}'_1, \ldots, \vec{w}'_M$, achieving the same error probability when used as codewords for P. The $\frac{1}{\ell}$ scaling factor comes from the factor of ℓ difference in the block length.

对于第二个论断，我们应用引理 8 得到

$$d_{\mathrm{C}}(\vec{x}_{m_1}, \vec{x}_{m_2}, Q) = \sum_{i=1}^{\ell} d_{\mathrm{C}}(\vec{x}_{m_1}(i), \vec{x}_{m_2}(i), P), \tag{30}$$

由于 P 是两两可逆的，右端在 $s = 1/2$处取得其最小值，这意味着 Q 也是两两可逆的。

□

下一个引理使我们能够在 $M = 2$ 时，将乘积信道的错误指数分解为各个组成部分。（请注意，此结果无法推广至 $M > 2$。）

引理 16。设 $P^{(1)}, P_{(2)} \ldots, P_{(k)}$ 为任意信道，并设 $\hat{}P = P^{(1)} \times P_{(2)} \times \ldots \times P_{(k)}$ 为乘积信道。则

$$\mathcal{E}_{2,\hat{P}} \le \sum_i \mathcal{E}_{2,P^{(i)}}. \tag{31}$$

证明。对于任意离散无记忆信道， $M = 2$ 的错误指数如下所示 [14, 式（1.19）] ：

$$\mathcal{E}_{2,P} = \max_{x_1,x_2\in\mathcal{X}_P} d_{\mathrm{C}}(x_1, x_2, P). \tag{32}$$

因此，我们选择使.....中的值达到最大值的 $\vec{x}_1, \vec{x}_2 \in \prod_i \mathcal{X}_{P^{(i)}}$。

$$\mathcal{E}_{2,\hat{P}} = \max_{\vec{x}_1,\vec{x}_2} d_{\mathrm{C}}(\vec{x}_1, \vec{x}_2, \hat{P}), \tag{33}$$

并令 $\vec{x}_1(i)$ 和 $\vec{x}_2(i)$ 分别表示 $\vec{x}_1(i)$ 和 $\vec{x}_2(i)$ 的第 i 个分量。然后，

$$\begin{aligned} d_{\mathrm{C}}(\vec{x}_1, \vec{x}_2, P) &= \max_{0\le s\le 1} d_{\mathrm{C}}(\vec{x}_1, \vec{x}_2, P, s) && (34)\\ &\overset{Lem.\ 8}{=} \max_{0\le s\le 1} \sum_i d_{\mathrm{C}}(\vec{x}_1(i), \vec{x}_2(i), P^{(i)}, s) && (35)\\ &\le \sum_i \max_{0\le s\le 1} d_{\mathrm{C}}(\vec{x}_1(i), \vec{x}_2(i), P^{(i)}, s) && (36)\\ &\overset{(32)}{\le} \sum_i \mathcal{E}_{2,P^{(i)}} && (37) \end{aligned}$$

按要求。 □

以下结果表明，如果将所有 $P^{(i)}$ 限制为两两可逆，则引理 16 中的等号成立：

引理 17. 设 $P_{(1)}\, P_{(2)} \ldots, P_{(k)}$ 为两两可逆信道，且 $\hat{} P = P^{(1)} \times P^{(2)} \times \ldots \times P_{(k)}$ 为其乘积信道。则

$$\mathcal{E}_{M,\hat{P}} = \sum_i \mathcal{E}_{M,P^{(i)}}. \tag{38}$$

$\hat{}$证明。我们首先注意到，根据引理 15 的第二部分， P 是成对可逆的。因此，由引理 12，存在某个 $(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_M)$，使得

$$\mathcal{E}_{\hat{P}} = \frac{2}{M(M-1)} \sum_{1\le m_1<m_2\le M} d_{\mathrm{B}}(\vec{x}_{m_1}, \vec{x}_{m_2}, \hat{P}). \tag{39}$$

此外，b 根据引理 8，我们可以将其分解为各个信道的 Chernoff 指数 s:

$$\mathcal{E}_{\hat{P}} = \frac{2}{M(M-1)} \sum_{1\le m_1<m_2\le M} \sum_i d_{\mathrm{B}}(\vec{x}_{m_1}(i), \vec{x}_{m_2}(i), P^{(i)}) \overset{Lem.\ 12}{\le} \sum_i \mathcal{E}_{P^{(i)}}. \tag{40}$$

为了证明另一个方向的不等式，我们只需将整个论证反过来。对于每个信道 $P^{(i)}$，选择 $(x_1^{(i)}, ..., x_M^{(i)})$使其在 (27) 中取等号。定义 $\vec{x}_m = (x_m^{(i)})_{i=1}^k$，即可得到

$$\mathcal{E}_{\hat{P}} \ge \frac{2}{M(M-1)} \sum_{1\le m_1\le m_2\le M} d_{\mathrm{B}}(\vec{x}_{m_1}, \vec{x}_{m_2}, \hat{P}) = \sum_i \mathcal{E}_{P^{(i)}}, \tag{41}$$

其中 s $\hat{}$teps基于 (27) 取等号成立（先对 P，然后对各个 $\mathcal{E}_{P^{(i)}}$） . □

D. 反馈误差指数

In the well-studied problem of 1-hop communication with feedback, the sender is given access to all of the previous channel outputs before sending the next channel input. For a discrete memoryless channel P, we define

the minimum probability of error with feedback $p_{\mathrm{err}}^{f}(n, M, P)$, and define the error exponent of the channel with M messages via

$$\mathcal{E}_{M,P}^{f} = \liminf_{T\to\infty} -\frac{1}{n} p_{\mathrm{err}}^{f}(n, M, P). \tag{42}$$

显然，$\mathcal{E}_{M,\ P}^{f} \geq \mathcal{E}_{M,\ P}$。一般而言，即使对于二元对称信道 [16]，有反馈误差指数也可能严格高于无反馈误差指数。然而，当 $M = 2$ 时，对于任意 DMC，两个误差指数相同：

定理 18. [16], [17] 对于任意 DMC P，当 $M = 2$ 时反馈不会改善错误指数，即

$$\mathcal{E}_{2,P}^{f} = \mathcal{E}_{2,P}. \tag{43}$$

E. Chernoff 散度的性质

以下引理讨论复合信道的 Chernoff 散度：

引理 19. [4, 引理 7] 设 P_1, P_2 为满足 $\mathcal{Y}_{P_1} \subseteq \mathcal{X}_{P_2}$ 的信道。对于所有 $x, x' \in \mathcal{X}_{P_1}$，我们有

$$d_{\mathrm{C}}(x, x', P_2 \circ P_1, s) \geq \min_{y, y' \in \mathcal{Y}_{P_1}} \left\{ d_{\mathrm{C}}(y, y', P_2, s) - (1-s)\log P_1(y|x) - s \log P_1(y'|x') \right\} - 2\log|\mathcal{Y}_{P_1}|. \tag{44}$$

下一个结果将似然比检验的错误概率与 Chernoff 散度联系起来。我们预计这一结果是标准的，但为完整起见，仍给出一个简短的证明。

引理 20. 设 P 和 Q 是定义在同一字母表 $\mathcal{X}$ 上的两个分布。对每个 $x \in \mathcal{X}$，将似然比定义为 $L(x) = \frac{P(x)}{Q(x)}$。那么

$$-\log \mathbb{P}_{x\sim P}(L(x) \leq L_0) \geq d_{\mathrm{B}}(P, Q) - \frac{1}{2}\log L_0. \tag{45}$$

证明。考虑集合

$$S = \{x \,|\, L(x) \leq L_0\}. \tag{46}$$

对于所有 $x \in S$, $P(x) \leq L_0 \cdot Q(x)$，因此

$$P(x) = \sqrt{P(x)} \cdot \sqrt{P(x)} \leq \sqrt{P(x)} \cdot \sqrt{L_0 \cdot Q(x)}, \tag{47}$$

从而得到

$$\mathbb{P}_{x\sim P}(L(x) \leq L_0) = \sum_{x\in S} P(x) \leq \sum_{x\in S} \sqrt{L_0 \cdot P(x) \cdot Q(x)} \leq \exp(-d_{\mathrm{B}}(P, Q)) \cdot \sqrt{L_0}. \tag{48}$$

□

IV. 定理1的级数情形

In this section, we handle the special case in which the nodes are arranged in a series (i.e., a line graph). That is, the nodes are labeled $\{0, 1, \ldots, |E|\}$, and the edges are $\{(j, j+1) : 0 \le j < |E|\}$. We refer to the corresponding channels as $P^{(1)}, P^{(2)} \ \ldots P^{(|E|)}$, with corresponding 1-hop error exponents $\mathcal{E}^{(1)}, \mathcal{E}^{(2)}, \ldots \mathcal{E}^{(|E|)}$. All channels are assumed to be pairwise reversible in this section.

This series setup directly generalizes previous work on the 2-hop setting in [1], [2], [3], [4], particularly our work on the the multi-bit setting [4]. Specifically, Theorem 1 for the series setup generalizes the main result therein by extending it from 2 hops to any constant number of hops. The protocol that we adopt has some significant differences to [4], which we discuss in Section IV-D. On the other hand, a key similarity is that we still adopt a block-structured strategy.

A. 顺序分块传输

We momentarily turn our attention to protocols in which nodes transmit blocks to each other *one after the other* without any parallel transmission. This idea would be highly wasteful if used in a standalone manner, but the idea (in the next subsection) will be to chain many instances of this approach together in parallel so that a negligible fraction of the total block length n is ultimately wasted.

对于一个串联图，将块大小为 B 的顺序分块传输定义如下：

对于每条边 $e = (u, v)$， u 通过 B 条信道跨越 e向 v 传输；

- 此过程按顺序进行；必须先完成较早边（即更靠近源点的边）上的传输，才能开始下一条边上的传输。

在第 IV-C 节中，我们将证明此设置下的以下关键引理：

Lemma 21. *Given a series graph, consider a sequential block transmission with block size B. Let $\vec{y}$ be the length-B sequence received by the final node $|E|$, and let $|f|$ be the maximum flow (since we are in the series setup, this is equal to $\min_j \mathcal{E}^{(j)}$). Then there exists a protocol such that for all $m_1 \neq m_2$,*

$$d_{\mathrm{B}}\big((\vec{y}|m = m_1), (\vec{y}|m = m_2)\big) \ge B(|f| - o_B(1)), \tag{49}$$

where $o_B(1)$是一个随着 B 增加而趋近于零，同时 G 和 M 保持不变的项。

B. 对我们主要问题设定的启示

回到我们的主要问题设定，即不要求节点一个接一个地传输，我们得到引理 21 的以下推论。

推论 22. 给定一个具有 n 个时间步的串联图，存在一种协议，使得对于任意 $m_1 \neq m_2$，节点 $|E|$ 接收到的长度为 n 的序列 $\vec{y}_{\mathrm{all}}$ 满足

$$d_{\mathrm{B}}\big((\vec{y}_{\mathrm{all}}|m = m_1), (\vec{y}_{\mathrm{all}}|m = m_2)\big) \ge (n - |E|B) \cdot (|f| - o_B(1)), \tag{50}$$

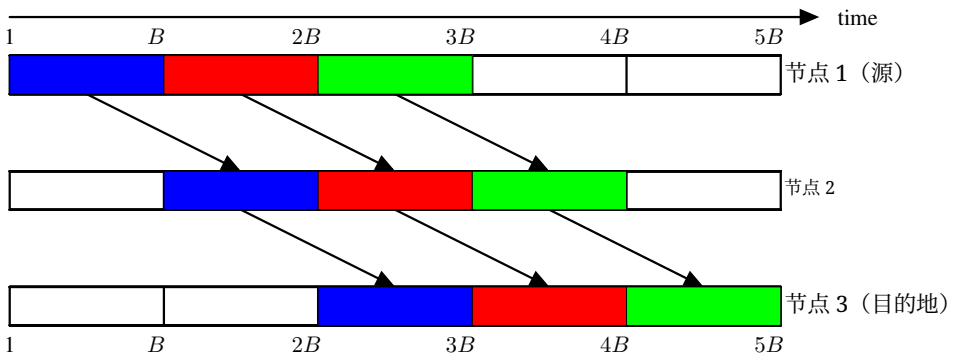


图 1. 示意图展示了如何将总计 n 个时间步划分为 $n/B - |E|$ 次顺序块传输。在此图中，$n = 5B$ 和 $|E| = 2$。

其中， o_B(1)是一项，随着 B 增加而趋近于零，同时 G 和 M 保持不变（且不依赖于 n）。

Proof. Consider the process of chaining several sequential block transmissions in parallel, as depicted in Figure 1. If there are t sequential block transmissions of length B, then they can clearly be executed in $(t + |E|) \cdot B$ time steps. In other words, with a block length of n, we can perform $n/B - |E|$ sequential block transmissions (up to asymptotically negligible rounding issues).

For each such sequential block transmission, we use the protocol that achieves Lemma 21. Since memory is not stored between the blocks, the blocks are conditionally independent given the source message. Let $\vec{y}(t)$ represent the t-th block received by the final node, and let $\vec{y}_{\text{all}} = (\vec{y}(1)\ \vec{y}(2)\ \ldots, \vec{y}(n/B - |E|))$. Then, using tensorization (Lemma 8) and Lemma 21, we have for all $m_1 \neq m_2$ that

$$\begin{aligned}
& d_{\text{B}}\big((\vec{y}_{\text{all}}|m = m_1), (\vec{y}_{\text{all}}|m = m_2)\big) \\
= & \sum_{t=1}^{n/B-|E|} d_{\text{B}}\big((\vec{y}(t)|m = m_1), (\vec{y}(t)|m = m_2)\big) && (51)\\
\geq & (n/B - |E|)B(|f| - o_B(1)) && (52)\\
= & (n - |E|B) \cdot (|f| - o_B(1)), && (53)
\end{aligned}$$

按要求。 □

于是，错误指数被以下数值所界定

$$\begin{aligned}
& \liminf_{n\to\infty} -\frac{1}{n} \log p_{\text{err}} && (54)\\
\geq & \liminf_{n\to\infty} \min_{m_1 \neq m_2} d_{\text{B}}\big((\vec{y}_{\text{all}}|m = m_1), (\vec{y}_{\text{all}}|m = m_2)\big) && (55)\\
= & \liminf_{n\to\infty} \frac{n - |E|B}{n}(|f| - o_B(1)) && (56)\\
= & |f| - o_B(1), && (57)
\end{aligned}$$

其中，(55)由引理 9 得出，(56)由推论 22 得出。这证明了任意接近 $|f|$ 的错误指数都是可实现的，从而证明了定理 1 在级数情形下成立。

C. 引理 21 的证明

我们首先证明，只需在一个便于处理的附加假设下证明引理 21：

Lemma 23. *Suppose that Lemma 21 holds under the additional restriction that each channel $P^{(j)}$ has M inputs $\{1, 2, \ldots, M\}$ such that $d_{\mathrm{B}}(m_1, m_2, P^{(j)}) \geq |f|$ for all $m_1 \neq m_2$. Then the general form of Lemma 21 as stated also holds.*

证明。考虑具有任意 $P^{(j)}$ 的引理21的一般设置。根据引理 14，对于每个 j，存在长度为 $\ell = M!$ 的码字 $\vec{x}_1^{(j)}, \vec{x}_2^{(j)}, \ldots, \vec{x}_M^{(j)}$，使得

$$d_{\mathrm{B}}(\vec{x}_{m_1}^{(j)}, \vec{x}_{m_2}^{(j)}, (P^{(j)})^{\ell}) \geq \ell \cdot \mathcal{E}_{M,P^{(j)}}. \tag{58}$$

令 $Q^{(j)}$ 为 $(P^{(j)})^{\ell}$ 在码字 $\vec{x}_1^{(j)}, \vec{x}_2^{(j)}, \ldots, \vec{x}_M^{(j)}$ 上的限制。我们首先证明信道 $Q^{(j)}$ 满足引理 23 中所述的额外要求。

By 引理 15， $Q^{(j)}$ 是成对可逆的。此外，对于每个 $m_1 \neq m_2$，由 (58) 可得 th at

$$d_{\mathrm{B}}(\vec{x}_{m_1}^{(j)}, \vec{x}_{m_2}^{(j)}, Q^{(j)}) = d_{\mathrm{B}}(\vec{x}_{m_1}^{(j)}, \vec{x}_{m_2}^{(j)}, (P^{(j)})^{\ell}) \geq \ell \cdot \mathcal{E}_{M,P^{(j)}}. \tag{59}$$

Then, we deduce from Lemma 12 (with the M inputs of $Q^{(j)}$ being substituted for $(x_1, x_2, \ldots, x_M)$) that $\mathcal{E}_{M,Q^{(j)}} \geq \ell \cdot \mathcal{E}_{M,P^{(j)}}$. Since the maxflow is $|f| = \min_j \mathcal{E}_{M,P^{(j)}}$ in the series setup, this gives $\mathcal{E}_{M,Q^{(j)}} \geq \ell \cdot |f|$. Then, let $|f'|$ be the maximum flow for the sequence of channels $Q^{(1)}, \ldots, Q^{(|E|)}$. Since $\mathcal{E}_{M,Q^{(j)}} \geq \ell \cdot |f|$ for all j, the new maximum flow satisfies $|f'| \geq \ell \cdot |f|$. On the other hand, by Lemma 15, we have $\mathcal{E}_{Q^{(j)}} \leq \ell \cdot \mathcal{E}_{P^{(j)}}$, which gives

$$|f'| = \min_j \mathcal{E}_{Q^{(j)}} \leq \min_j \ell \cdot \mathcal{E}_{P^{(j)}} \leq \ell \cdot |f|. \tag{60}$$

结合对 $|f'|$ 的这些上界和下界 ，我们得出 $|f'| = \ell \cdot |f|$ ，并由 (59) 我们推出 channels $Q^{(j)}$ satisfy the additional requirement imposed in the statement of Lemma 23 (with $|f|$ renamed to $|f'|$).

Now suppose that Lemma 21 holds for the new sequence of channels $Q^{(1)}, \ldots, Q^{(|E|)}$, and let $\vec{y}'$ denote the corresponding length-B sequence received by the destination node. This means that there exists a sequential block transmission protocol of block size B such that for any $m_1 \neq m_2$,

$$d_{\mathrm{B}}\big((\vec{y}'|m = m_1), (\vec{y}'|m = m_2)\big) \geq B(|f'| - o_B(1)) \geq B(\ell|f| - o_B(1)). \tag{61}$$

Any sequential block transmission with block size B using the new channels $Q^{(1)}, \ldots, Q^{(|E|)}$ can easily be converted to a sequential block transmission with block size $B \cdot \ell$ over the original channels $P^{(1)}, \ldots, P^{(|E|)}$ by simply expanding instances of the new letter m by the codeword $\vec{x}_M$. Therefore, there exists a sequential block

块大小为 $B \cdot \ell$ 的传输，使得对于任意 $m_1 \neq m_2$,

$$d_{\mathrm{B}}\big((\vec{y}|m=m_1),(\vec{y}|m=m_2)\big) \geq B(\ell|f| - o_B(1)). \tag{62}$$

由于 ℓ 是常数，我们可以进行适当的缩放，将上面的 $B\ell$ 重新命名为 B （整体顺序分块传输长度），由此得出引理 21 对原始信道 $P^{(1)}, \ldots, P^{(|E|)}$ 成立。 □

本节余下部分中，我们在引理 23 中陈述的额外假设下进行讨论。我们还假设 B 为偶数，以避免舍入问题。

1) 转发协议描述：我们考虑以下针对长度为 B 的单个顺序分组传输的协议：

- Every node j ($0 \leq j \leq |E|$) is assigned a state consisting of two integers $(m^{(j)}\ \ell^{(j)})$ with $1 \leq m^{(j)} \leq M$ and $0 \leq \ell^{(j)} \leq B/2$, which we will refer to as (m, ℓ) when there is no ambiguity. A state of $(m^{(j)}\ \ell^{(j)})$ roughly indicates that the node believes $m^{(j)}$ to be the most likely value of m, where larger values of $\ell^{(j)}$ indicate higher confidence, whereas a smaller value of $\ell^{(j)}$ indicates that there is another competing value $m' \neq m^{(j)}$ that is also relatively likely (we make this more precise below).
- 节点 0 被分配了 $m_{(0)} = m$ 和 $\ell_{(0)} = B/2$，其中 m 是实际消息。
- If node j is assigned a state (m, ℓ), then this node sends out a $B/2 + \ell$ copies of m followed by $B/2 - \ell$ copies of $m + 1$ (unless $m = M$, in which case we wrap around and replace $m + 1$ by 1 here). We refer to the corresponding string as $\vec{x}_{m,\ell}$, or $\vec{x}^{(j)}_{m,\ell}$ when we want to emphasize which node is being considered.
- Node j ($j \geq 1$), upon receiving $\vec{y}^{(j)}$ from node $j-1$, computes the likelihoods $\mathbb{P}(\vec{y}^{(j)}|m^{(0)} = m)$ for each possible message $1 \leq m \leq M$. Let m^*, m^{**} be the resulting most likely and second most likely value of m respectively (with arbitrary tie-breaking).
- 节点 j 现已被分配状态为

$$m^{(j)} = m_1, \ell^{(j)} = \min\left(\frac{B}{2}, \left\lfloor \frac{1}{4|f|} \cdot \log \frac{\mathbb{P}(\vec{y}^{(j)}|m^*)}{\mathbb{P}(\vec{y}^{(j)}|m^{**})} \right\rfloor\right). \tag{63}$$

选择这些方程，使得 $m^{(j)}$ 反映最可能的源消息，而 $\ell^{(j)}$ 恰当地反映置信度，即最可能的消息 $m^{(j)}$ 大约比 m 的下一个取值更有可能，概率约为其 $\exp(4\ell^{(j)}|f|)$ 倍。

2) 策略分析：我们首先在状态上定义一个伪度量，方法是

$$d((m_1,\ell_1),(m_2,\ell_2)) = \begin{cases} |f||\ell_1 - \ell_2| & m_1 = m_2 \\ |f||\ell_1 + \ell_2| & m_1 \neq m_2. \end{cases} \tag{64}$$

To see that d defined above is a pseudometric, note that d is equivalent to the ℓ_1-norm after applying the mapping $(m, \ell) \to (0, \ldots, 0, \ell|f|, 0, \ldots, 0)$ (with $\ell|f|$ occurring at position m). The reason for being a pseudometric rather than a metric is that $d((m_1, 0), (m_2, 0)) = 0$.

Intuitively, d is a measure of how far apart the beliefs (m_1, ℓ_1) and (m_2, ℓ_2) are. If $m_1 = m_2$, then these two beliefs are close when ℓ_1 and ℓ_2 are close. If $m_1 \neq m_2$, the beliefs are close only when m_1 and m_2 are both comparatively likely, which corresponds to the case where both ℓ_1 and ℓ_2 are small.

我们使用 d 来界定与两个码字相关的巴氏距离：

$$\begin{aligned} d_{\mathrm{B}}(\vec{x}^{(j)}_{m_1,\ell_1}, \vec{x}^{(j)}_{m_2,\ell_2}, (P^{(j)})^B) &= \sum_{i=1}^{B} d_{\mathrm{B}}(\vec{x}^{(j)}_{m_1,\ell_1}(i), \vec{x}^{(j)}_{m_2,\ell_2}(i), P^{(j)}) && (65)\\ &\geq |f| d_H(\vec{x}^{(j)}_{m_1,\ell_1}(i), \vec{x}^{(j)}_{m_2,\ell_2}(i)) && (66)\\ &\geq d((m_1,\ell_1),(m_2,\ell_2)), && (67)\end{aligned}$$

其中：

- (65) 由张量化（引理 8）可得。
- (66) comes from the fact that the number of indices in which $\vec{x}^{(j)}_{m_1,\ell_1}(i)$ and $\vec{x}^{(j)}_{m_2,\ell_2}(i)$ differ is simply their Hamming distance, and each of these corresponding terms are at least $|f|$ due to the extra assumption in Lemma 23.
- 为说明（67）成立的原因，我们使用 $\vec{x}_{(j)\ m,\ell}$ 的定义，并分别处理 $m_1 = m_2$ 和 $m_1 \neq m_2$ 两种情况。若 $m_1 = m_2$，则

$$d_H(\vec{x}^{(j)}_{m_1,\ell_1}, \vec{x}^{(j)}_{m_2,\ell_2}) = |\ell_1 - \ell_2|, \tag{68}$$

这与（64）相符。如果 $m_1 \neq m_2$，那么两个码字的前 $B/2 + \min(\ell_1, \ell_2)$ 个符号分别仅由 m_1 和 m_2 组成，因此

$$d_H(\vec{x}^{(j)}_{m_1,\ell_1}, \vec{x}^{(j)}_{m_2,\ell_2}) \geq B/2 + \min(\ell_1,\ell_2) \geq \max(\ell_1,\ell_2) + \min(\ell_1,\ell_2) = \ell_1 + \ell_2, \tag{69}$$

这再次与（64）相符。

Observe that the sequence of states taken by the nodes forms a Markov process. We establish a bound on the probability of achieving any intermediate state given the beginning state; intuitively, the following result states that given the original state $(m^{(0)}, \ell^{(0)}) = (m, B/2)$, the probability of seeing another state $(m^{(j)}, \ell^{(j)})$ is exponentially small in the distance between the two states defined by (64).

引理 24. 对所有 j、所有 $m_1 \neq m_2$ 以及所有 (m', ℓ')，我们有

$$-\log \mathbb{P}\big(m^{(j)} = m', \ell^{(j)} = \ell' | m = m_1\big) \geq 2d\big((m_1, B/2), (m', \ell')\big) - 2j\log(M(B+1)) - 2j|f| - j\log(M-1) \tag{70}$$

and

$$d_{\mathrm{B}}\big((\vec{y}^{(j+1)}|m = m_1), (\vec{y}^{(j+1)}|m = m_2)\big) \geq B|f| - 2(j+1)\log(M(B+1)) - 2j|f| - j\log(M-1). \tag{71}$$

证明。我们使用归纳论证，基例是当 $j = 0$ 时，(70) 显然成立：二者之一

左侧因概率为零而趋于无穷大，或者因概率为一而为零（即 $(m',\ell') = (m_1, B/2)$），在这种情况下，右侧的第一项为零。

我们对 j 进行归纳证明，并将归纳论证分解为以下两个断言：

- 若(70)对于 j成立，则(71)对于 j成立
- 如果(71) 对 j 成立，则 (70) 对 $j+1$ 成立。

Start by assuming that (70) holds for some j. Observe that $m \to (m^{(j)}, \ell^{(j)}) \to \vec{y}^{(j+1)}$ forms a Markov chain. Let P_1 be the channel $m \to (m^{(j)}, \ell^{(j)})$ and P_2 be the channel $(m^{(j)}, \ell^{(j)}) \to \vec{y}^{(j+1)}$; applying Lemma 19 with $s = 1/2$ then gives for any $m_1 \neq m_2$ that

$$\begin{aligned} & d_{\mathrm{B}}\left(\big(\vec{y}^{(j+1)}|m^{(0)}=m_1\big), \big(\vec{y}^{(j+1)}|m^{(0)}=m_2\big)\right) \\ \geq & \min_{m',m'',\ell',\ell''} \left\{ d_{\mathrm{B}}(\vec{x}_{m',\ell'}, \vec{x}_{m'',\ell''}, (P^{(j)})^B) - \frac{1}{2}\log \mathbb{P}(m',\ell'|m^{(0)}=m_1) - \frac{1}{2}\log \mathbb{P}(m'',\ell''|m^{(0)}=m_2) \right\} \\ & -2\log(M(B+1)). \end{aligned} \tag{72}$$

由于（70）对 j （以及所有（m'， ℓ'）和 m_1）成立，我们可以将其与（67）结合，以简化最小化条件中的表达式：

$$\begin{aligned} & d_{\mathrm{B}}(\vec{x}_{m',\ell'}, \vec{x}_{m'',\ell''}, (P^{(j)})^B) - \frac{1}{2}\log \mathbb{P}(m',\ell'|m=m_1) - \frac{1}{2}\mathbb{P}(m'',\ell''|m=m_2) \\ \geq & \; d\big((m',\ell'),(m'',\ell'')\big) + d\big((m_1,B/2),(m',\ell')\big) + d\big((m_2,B/2),(m',\ell')\big) \qquad (73) \\ & -2j\log(M(B+1)) - 2j|f| - j\log(M-1) \\ \geq & \; d\big((m_1,B/2),(m_2,B/2)\big) - 2j\log(M(B+1)) - 2j|f| - j\log(M-1) \qquad (74) \\ = & \; B|f| - 2j\log(M(B+1)) - 2j|f| - j\log(M-1), \qquad (75) \end{aligned}$$

where (74) follows from non-negativity (first d term) and the triangle inequality (second and third d terms) for the pseudometric d, and the last step uses the definition of d with $m_1 \neq m_2$. By substituting (75) into (72), we conclude that (71) holds for j.

We now turn to the second dot point mentioned at the start of the proof. We will split this into two cases. We note that in (70) we are interested in the event that $(m^{(j)}, \ell^{(j)}) = (m', \ell')$, and for this to happen, m' must be the value assigned to m^* used in (63), i.e., m' must be the most likely message given $\vec{y}^{(j)}$. We will use this fact in both cases.

Case 1 ($m' = m_1$)**.** If $\ell' = B/2$, then $d((m_1, B/2), (m', \ell')) = 0$, so that the RHS of (70) is upper bounded by 0, and therefore (70) trivially holds. Therefore, we assume that $\ell' < B/2$.

对于每个 $\vec{y}^{(j+1)}$ 和每个 $m'' \neq m_1$，定义

$$L_{m''}(\vec{y}^{(j+1)}) = \frac{\mathbb{P}(\vec{y}^{(j+1)}|m^{(0)}=m_1)}{\mathbb{P}(\vec{y}^{(j+1)}|m^{(0)}=m'')}. \tag{76}$$

根据 (63) 以及假设 设 $\ell^{(j+1)} < B/2$，存在 m'' （其可以选择为第二个最可能的 y

在给定 $\ell^{(j+1)}$ 的情况下， $m^{(0)}$ 的值（注意最可能的值是 $m'=m_1$），使得

$$\frac{1}{4|f|}\log L_{m''}(\vec{y}^{(j+1)}) < \ell^{(j+1)}+1. \tag{77}$$

结合对所有 m'' 的并集界，我们得到

$$\begin{aligned}
&\mathbb{P}(m^{(j+1)}=m', \ell^{(j+1)}=\ell' | m^{(0)}=m_1) && (78)\\
\leq\ & (M-1)\max_{m''\neq m_1}\mathbb{P}\left(L_{m''}(\vec{y}^{(j+1)}) < \exp(4|f|(\ell'+1)) \,\middle|\, m^{(0)}=m_1\right). && (79)
\end{aligned}$$

现在我们可以得出结论：

$$\begin{aligned}
&-\log\mathbb{P}(m^{(j+1)}=m', \ell^{(j+1)}=\ell' | m^{(0)}=m_1)\\
\geq\ & -\log\left(\max_{m''\neq m_1}\mathbb{P}\left(L_{m''}(\vec{y}^{(j+1)}) < \exp(4|f|(\ell'+1)) \,\middle|\, m^{(0)}=m_1\right)\right) - \log(M-1) && (80)\\
\geq\ & \min_{m''\neq m_1}\left(d_{\mathrm{B}}\left((\vec{y}^{(j+1)}|m^{(0)}=m_1), (\vec{y}^{(j+1)}|m^{(0)}=m'')\right)\right) - 2|f|(\ell'+1) - \log(M-1) && (81)\\
\geq\ & B|f| - 2(j+1)\log(M(B+1)) - 2j|f| - j\log(M-1) - 2|f|(\ell'+1) - \log(M-1) && (82)\\
=\ & (B-2\ell')|f| - 2(j+1)\log(M(B+1)) - 2(j+1)|f| - (j+1)\log(M-1) && (83)\\
=\ & 2d\big((m_1, B/2), (m', \ell')\big) - 2(j+1)\log(M(B+1)) - 2(j+1)|f| - (j+1)\log(M-1). && (84)
\end{aligned}$$

其中，(81) 由引理 20 得出，(82) 由归纳假设 (71) 得出，而 (84) 由 d 与 $m'=m_1$的定义得出。因此，如所需，(70) 对于 $j+1$ 成立。

情况 2 ($m'\neq m_1$)。对于每个 $\vec{y}^{(j+1)}$，定义

$$L(\vec{y}^{(j+1)}) = \frac{\mathbb{P}(\vec{y}^{(j+1)}|m^{(0)}=m_1)}{\mathbb{P}(\vec{y}^{(j+1)}|m^{(0)}=m')}. \tag{85}$$

令 m'' （其可能取决于 $\vec{y}^{(j+1)}$）为 $m^{(0)}$ 中概率第二高的取值（注意， m' 是概率最高的取值）。因此我们有

$$\frac{1}{4|f|}\log L(\vec{y}^{(j+1)}) \leq \frac{1}{4|f|}\log\frac{\mathbb{P}(\vec{y}^{(j+1)}|m^{(0)}=m'')}{\mathbb{P}(\vec{y}^{(j+1)}|m^{(0)}=m')} \leq -\ell^{(j+1)}, \tag{86}$$

其中，第一个不等式利用了 m_1 的似然不大于 m'' 的似然这一事实；第二个不等式则是在将其中的最小值下界为第二项后，由式（63）得出。

我们现在以与情况 1 类似的方式继续：

$$
\begin{aligned}
& -\log \mathbb{P}(m^{(j+1)}=m', \ell^{(j+1)}=\ell' | m^{(0)}=m_1) && (87)\\
\geq\ & -\log \mathbb{P}\left(L(\vec{y}^{(j+1)}) \leq \exp(-4|f|\ell') \,\middle|\, m^{(0)}=m_1\right) \\
\geq\ & d_{\mathrm{B}}\left((\vec{y}^{(j+1)}|m^{(0)}=m'), (\vec{y}^{(j+1)}|m^{(0)}=m_1)\right) + 2|f|\ell' \\
\geq\ & B|f| - 2(j+1)\log(M(B+1)) - 2j|f| - j\log(M-1) + 2|f|\ell' \\
=\ & (B+2\ell')|f| - 2(j+1)\log(M(B+1)) - 2j|f| - j\log(M-1) \\
\geq\ & 2d\big((m',\ell'),(m_1,B/2)\big) - 2(j+1)\log(M(B+1)) - 2(j+1)|f| - (j+1)\log(M-1), && (88)
\end{aligned}
$$

其中，每一步均与（80）–（84）中的对应步骤类似。由此，我们证明了（70）对 $j+1$ 成立，从而完成了引理 24 的证明。 □

引理 21 的关键结果现在可由将（71）应用于 $j=|E|-1$ 并回顾该引理是针对不断增大的 B 、固定的 $|E|$ 和 M 而陈述的，立即得到。

D. 与先前工作的比较

在我们之前的论文 [4], 中，我们证明了引理 21 的 2 跳版本。在 [4], 中，问题被约化为情形$|\mathcal{X}_P| = M$（与引理 23 类似），然后按如下方式处理：

- 中继节点计算所接收符号的经验分布与每条消息对应的输出分布之间的 KL 散度。

- 随后，中继发送一个排序后的序列，其中与 KL 散度最小的符号相关的出现次数更多，而其余符号的出现次数相同。

本文采用的方法则基于似然比（见 (63)）；计算该比值后，发送一个至多包含两个不同取值的有序序列。

At first glance, one may have expected that using likelihood ratios would be difficult, as there can be $\binom{M}{2}$ likelihood ratios of interest. However, our approach here shows that using the likelihood ratios between only the two most likely values is sufficient. Overall, our approach here is more similar to the one used in [1] (for general binary-input channels rather than the BSC), which directly computes the likelihood ratio but only handles the 2-hop setting and requires $M = 2$.

V. 定理 1 的证明（主要可实现性界）

In this section, we prove Theorem 1 in the general case, using the series case from Section IV as a stepping stone. We crucially make use of the flow decomposition theorem (Theorem 6) together with its associated paths $p_1, p_2, \ldots, p_k$ and flows $f_1, f_2, \ldots, f_k$. The overall idea is to send conditionally independent messages along different paths and only accumulate them at the end.

A. 边不相交路径

当路径 $p_1, p_2, .. ., p_k$ 两两边不相交时，协议会更简单，因此我们先讨论这种情况，再转向一般情况。

The key idea is to transmit information across each path independently and aggregate them only at the end. Although we assume that paths are edge-disjoint, they may still share nodes. However, each node is designed to act independently on every path it is involved in, so that information entering a node in one path does not affect other outgoing paths.

Let $\vec{y}(1) \ldots, \vec{y}(k)$ be the sequences received by the destination node along the k different paths after n time steps, and let $\vec{y}_{\text{all}} = (\vec{y}(1) \ldots, \vec{y}(k))$. By Corollary 22, for all i and all $m_1 \neq m_2$, there exists a (single-path) protocol such that

$$d_{\text{B}}\big((\vec{y}(i)|m = m_1), (\vec{y}(i)|m = m_2)\big) \geq (n - |V|B)(|f_i| - o_B(1)), \tag{89}$$

其中， $|V|$ 项源于简单路径最多只能包含 $|V|$ 条边。因此，再令 $B \to \infty$，我们有

$$\begin{aligned} & d_{\text{B}}\big((\vec{y}_{\text{all}}|m = m_1), (\vec{y}_{\text{all}}|m = m_2)\big) \\ = & \sum_{i=1}^{k} d_{\text{B}}\big(\vec{y}(i)|m = m_1, \vec{y}(i)|m = m_2\big) \\ \geq & \sum_{i=1}^{k} (n - |V|B)(|f_i| - o_B(1)) \qquad (90) \\ = & (n - |V|B)(|f| - o_B(1)). \qquad (91) \end{aligned}$$

注意，此处将 k 视为常数，因为 k 是图 G 的性质，因此不依赖于 B。

应用引理 9，我们推得总错误指数的下界为

$$\mathcal{E}_G \geq \liminf_{n\to\infty} -\frac{1}{n}(n - |V|B)(|f| - o_B(1)) = |f| - o_B(1), \tag{92}$$

这表明可以达到任意接近 $|f|$ 的误差指数。

B. 一般情形

For the general case, we will still rely on the edge decomposition as before. Our approach is to reduce to the case where paths are edge-disjoint by considering the use of parallel edges, which are allowed to exist in all of our analysis and results. This idea has also been used previously when studying channel capacity results for general networks, e.g., see [11, Sec. 15.3]

我们首先注意到两点：

- Suppose that there is some edge e such that P_e is a product channel, say $P_e = Q_1 \times Q_2$. We can delete e and replace it by two parallel edges e_1 and e_2, and put the channels Q_1 and Q_2 respectively and preserve the error exponent. This is because any symbol that can be sent across $Q_1 \times Q_2$ can be separately sent across the two channels in an identical manner.

• 对于任意图 G，设 G' 为通过将每个 P_e 替换为其 ℓ 重积 P_e^ℓ 而得到的结果图。那么$\mathcal{E}_{G'} \le \ell\mathcal{E}_G$。为此，注意到，在 G' 上执行的任何具有 n 个时间步的协议，都可以在 G上用 $n \cdot \ell$ 个时间步执行。

这两个观察使我们能够进行如下操作：

• 选取一个较大的分组长度 B，并将每个信道 P_e 替换为$(B+k)$重乘积信道 P_e^{B+k}

• 对于每条边 e，令

$$S_e = \{i : e \in p_i\} \tag{93}$$

表示经过 e 的路径的索引。对于每个 $i \in S_e$，定义

$$B_i = \left\lceil \frac{f_i}{\sum_{j\in S_e} f_j} \cdot B \right\rceil. \tag{94}$$

由于只有 k 条路径，我们有 $|S_e| \le k$。此外，我们有

$$\sum_{i\in S_e} B_i = \sum_{i\in S_e} \left\lceil \frac{f_i}{\sum_{j\in S_e} f_j} \cdot B \right\rceil \le \sum_{i\in S_e} \left(\frac{f_i}{\sum_{j\in S_e} f_j} \cdot B + 1 \right) = B + |S_e| \le B + k. \tag{95}$$

For each edge e with corresponding channel P_e^{B+k}, replace this by $|S_e|$ parallel edges, where the corresponding channels are $P_e^{B_i}$ for each $i \in S_e$. We will refer to each of these new edges by e_i'. This is possible since $\sum_{i\in S_e} B_i \le B + k$. If $\sum_{i\in S_e} B_i < B + k$, then we simply add a dummy (unused) channel to make up for the difference. Let G' denote this new graph.

注意，经过变换后，相应平行边的容量为

$$\left\lceil \frac{f_i}{\sum_{j\in S_e} f_j} \cdot B \right\rceil \cdot \mathcal{E}_e \ge B \cdot f_i, \tag{96}$$

注意到 $\sum_{j\in S_e} f_j = f_e \le \mathcal{E}_e$ 由流约束可得。

Let p_i' be the path in G' formed by joining up edges of the form e_i' for each $e \in p_i$. Since p_i is a path and e_i' joins the same pair of vertices as e, p_i' is also a path. By (96), each of these edges e_i' has error exponent at least $B \cdot f_i$, so that the capacity on path p_i' is at least $B \cdot f_i$. Thus, there exist edge-disjoint paths $p_1', p_2', \ldots, p_k'$ on G' such that the capacity on each path p_i' is at least $B \cdot f_i$.

现在，我们已经将问题归结为上一小节所研究的边不相交情形，并得出结论：我们可以实现任意接近 $B|f|$的错误指数。由于时间被缩放了 $B + k$ 倍，我们得出 $\mathcal{E}_{G'} \le (B + k)\mathcal{E}_G$，因此

$$\mathcal{E}_G \ge \frac{1}{(B+k)}\mathcal{E}_{G'} = \frac{B|f|}{(B+k)} \tag{97}$$

取 $B \to \infty$，我们得出 $\mathcal{E}_G \ge |f|$。

VI. Proof of Theorem 2 (Main Converse Bound)

我们反向界的证明通过最大流最小割对偶定理（定理 5）构造一个值与 $|f|$ 相同的割。此外，我们还将使用以下引理，其中回顾一下，$\mathcal{E}^f_{M,\ P}$ 表示

具有完全反馈的信道的1跳错误指数:

Lemma 25. *Consider any partition (A, B) of the nodes in G with the source node in A and the destination node in B. Let $P^{(1)}$ $P^{(2)}$ $\ldots, P^{(k)}$ be the channels associated with edges from A to B, and let $\hat{P} = P^{(1)} \times P^{(2)} \times \ldots \times P^{(k)}$ be the product channel. Then for all M,*

$$\mathcal{E}_{M,G} \leq \mathcal{E}^{f}_{M,\hat{P}}. \tag{98}$$

证明。我们将 A 压缩为单个节点，方法是允许 A 内部不同节点之间进行无限信息传输，并对 B 采取相同处理。显然，这一修改不会降低总体错误指数。

At first glance, it may appear that the net result of doing this is simply allowing A to send information to B via $P^{(1)}, P^{(2)}$ $\ldots, P^{(k)}$, each a total of n times, which is the same as using $\hat{P}$ for n times. However, we must also account for the possibility of back-edges, which amounts to having a certain degree of feedback. In view of this, since any protocol that can be executed under the original setup can also be executed under the new setup with n uses of $\hat{P}$ *with full feedback*, we conclude that $\mathcal{E}_{M,G} \leq \mathcal{E}^{f}_{M,\hat{P}}$. □

由于$\mathcal{E}^{f}_{M,P}$是 M的减函数， $\mathcal{E}^{f}_{M,P} \leq \mathcal{E}^{f}_{2,P}$。此外，当 $M = 2$ （定理18）时，反馈不会改善误差指数，因此引理 25 立即给出以下推论:

ˆ 推论 26。在引理 25 中对 P 的定义下，我们有

$$\mathcal{E}_{M,G} \leq \mathcal{E}_{2,\hat{P}}. \tag{99}$$

现在我们可以利用最大流-最小割对偶性（定理 5），并考虑一个割（A， B），使得 A 包含源点， B 包含汇点，并且

$$\sum_{i=1}^{k} \mathcal{E}_{2,P^{(i)}} = |f|, \tag{100}$$

where $P^{(1)}, P^{(2)}$ $\ldots, P^{(k)}$ are the channels associated with edges that begin in A and end in B. Letting $\hat{P} = P^{(1)} \times P^{(2)} \times \ldots \times P^{(k)}$ be the product channel, Lemma 16 gives

$$\mathcal{E}_{2,\hat{P}} \leq \sum_{i} \mathcal{E}_{2,P^{(i)}} \tag{101}$$

这意味着

$$\mathcal{E}_{M,G} \leq \mathcal{E}_{2,\hat{P}} \leq \sum_{i} \mathcal{E}_{2,P^{(i)}} = |f| \tag{102}$$

按要求。

VII. Achievability Bounds for General Channels

本节研究信道不必两两可逆情况下的可达性界。

A. 固定数量的消息

When the pairwise reversible assumption is dropped, Theorem 1 can fail to hold even in the 2-hop setting corresponding to a graph with 3 nodes and 2 edges [4, Thm. 3]. However, it is possible to prove some weaker achievability bounds. Recalling $\tilde{\mathcal{E}}$ from Definition 13, we state the following achievability bound for general channels:

˜引理 27。给定任意图 G，令 $|f|$ 为将每条边 e 的容量设置˜为 $\mathcal{E}_{P_e}$ 时所形成的最大流。那么

$$\mathcal{E}_G \geq |\tilde{f}| \tag{103}$$

Proof. We use the same proof as the pairwise reversible case. The only use of the pairwise reversibility condition is in Lemma 23 when constructing the codewords $\vec{x}_1^{(j)}, \vec{x}_2^{(j)}, \ldots, \vec{x}_M^{(j)}$ of length $\ell = M!$ such that $d_{\mathrm{B}}(\vec{x}_{m_1}^{(j)}, \vec{x}_{m_2}^{(j)}, (P^{(j)})^\ell) \geq \ell \cdot \mathcal{E}_{M,P^{(i)}}$. However, such codewords of length $\ell = M!$ with $d_{\mathrm{B}}(\vec{x}_{m_1}^{(j)}, \vec{x}_{m_2}^{(j)}, (P^{(j)})^\ell) \geq \ell \cdot \tilde{\mathcal{E}}_{M,P^{(i)}}$ (i.e., with $\tilde{\mathcal{E}}$ in place of $\mathcal{E}$) are still guaranteed to exist by Lemma 14 even when the pairwise reversibility condition is dropped. Outside of Lemma 23, the proof follows with $\mathcal{E}_{M,P^{(i)}}$ replaced by $\tilde{\mathcal{E}}_{M,P^{(i)}}$ where appropriate. □

B. 零速率误差指数

在此 ˜小节，我们证明定理 4。我们首先将$\mathcal{E}_{M,P}$ 与零速率错误指数联系起来 :

引理 28。对于任意 $M \geq 2$，我们有

$$\tilde{\mathcal{E}}_{M,P} \geq E_P(0) \geq \frac{M-1}{M}\tilde{\mathcal{E}}_{M,P}. \tag{104}$$

在两边取 $M \to \infty$ 后，我们也有

$$\lim_{M\to\infty} \tilde{\mathcal{E}}_{M,P} = E_P(0). \tag{105}$$

证明。我们使用定理 11 中 $E_P(0)$ 的表达式。令 q 使 (26) 中的表达式达到最大值。按照分布 q独立地绘制 $(x_1, x_2, \ldots, x_M)$˜，并注意到定义 13 中 $\mathcal{E}$ 的定义给出

$$\tilde{\mathcal{E}}_{M,P} \geq \mathbb{E}\left(\frac{2}{M(M-1)} \sum_{1\leq m_1<m_2\leq M} d_{\mathrm{B}}(x_{m_1}, x_{m_2}, P)\right) = E_P(0). \tag{106}$$

为证明其中第二个 引理陈述中的等号，令 q为 $(x_1, x_2, \ldots, x_M)$ 的经验分布（类型）)

从而得到 (28)。将 q 代入定理 11，可得

$$\begin{aligned} E_P(0) &\geq \sum_{x,x'} q_x q_{x'} d_{\mathrm{B}}(x, x', P) && (107)\\ &= \frac{1}{M^2} \sum_{1\leq m_1, m_2 \leq M} d_{\mathrm{B}}(x_{m_1}, x_{m_2}, P) && (108)\\ &= \frac{2}{M^2} \sum_{1\leq m_1 < m_2 \leq M} d_{\mathrm{B}}(x_{m_1}, x_{m_2}, P) && (109)\\ &\overset{(28)}{=} \frac{M-1}{M} \tilde{\mathcal{E}}_{M,P}. && (110) \end{aligned}$$

□

现在我们通过设计一种协议来证明定理2，该协议的错误指数可以任意接近 $|f|$，其中 f 中的边权分别是相应的 1 跳零速率错误指数（即边 e的 $E_{P_e}(0)$）。

Let M be a (large) constant, and let $|\tilde{f}_M|$ be the error exponent formed by setting the capacity of each edge e to $\tilde{\mathcal{E}}_{M,P_e}$. Since $\tilde{\mathcal{E}}_{M,P_e} \geq E_{P_e}(0)$ for all P, we have $|\tilde{f}_M| \geq |f|$.

Building on Sections IV and V, choose a (large) block size B, and generate a sequential block transmission protocol of length B with M messages. We can treat a single sequential block transmission as a discrete memoryless channel $Q_{B,M}$. More precisely, the input to $Q_{B,M}$ is the message that the source node receives, and the output of $Q_{B,M}$ is what the destination receives from a single block. The analysis in Sections IV and V shows that given any $\epsilon > 0$, there is some sufficiently large B such that

$$d_{\mathrm{B}}(m_1, m_2, Q_{B,M}) \geq B(|\tilde{f}_M| - \epsilon) \tag{111}$$

对于所有 $m_1 \neq m_2$。代入 (13)，可得

$$\tilde{\mathcal{E}}_{M,Q_{B,M}} \geq B(|\tilde{f}_M| - \epsilon). \tag{112}$$

我们现在继续界定 Q_B 的零速率错误指数。根据引理 28,

$$E_{Q_{B,M}}(0) \geq \frac{M-1}{M} \tilde{\mathcal{E}}_{M,Q_B} \geq \frac{M-1}{M} B(|\tilde{f}_M| - \epsilon) \geq \frac{M-1}{M} B(|f| - \epsilon) \tag{113}$$

Similar to the proof of Corollary 22, we can run $n/B - |V|$ conditionally independent copies of the block protocol in n time steps, which is equivalent to using channel Q_B for a total of $n/B - |V|$ times. By adopting any coding strategy for Q_B that attains its zero-rate error exponent,[5] we deduce that

$$E_G(0) \geq \lim_{n\to\infty} -\frac{1}{n} \cdot \frac{n - |V|B}{B} \cdot \frac{M-1}{M} \cdot B(|f| - \epsilon) = \frac{M-1}{M} \cdot (|f| - \epsilon). \tag{114}$$

由于 ϵ 任意小且 M 任意大，我们得出结论:

$$E_G(0) \geq |f| \tag{115}$$

[5]这样的策略 egy可能非常复杂，但我们要提醒的是，我们关注的是信息论极限，而不是实用协议 ls.

按要求。

VIII. 进一步结果

本节中，我们给出一些补充结果，说明在何种情况下可以（或不可以）放宽主要结果中的假设。

A. $M > 2$ 的逆界

Recall that $|f_{M,G}|$ denotes the maxflow for the graph G with M messages. One might hope that the proof of Theorem 2 can be modified to prove a converse bound of $\mathcal{E}_{M,G} \leq |f_{M,G}|$ (instead of $|f_{2,G}|$). However, there are at least two problems in doing so:

- 当 $M > 2$ 时，我们无法保证单跳反馈与无反馈的错误指数相等；
- 当 $M = 2$ 时，可能不成立所需的引理 16 的对应结论。

在本小节中，我们通过施加额外假设，部分地解决这些问题。具体而言，我们证明以下结论：

定理 29. 假设下列条件成立：

- G 具有一个没有反向边的最小割（即，所有边都从源侧指向汇侧）；
- G 中的每个信道都是成对可逆的。

因此，我们有 $\mathcal{E}_{M,\ G} = |f_{M,\ G}|$。

我们继续证明。尽管引理 25 中的界使用了反馈错误指数，但如果划分中没有反向边，我们就无需考虑反馈。由此立即得到：

Lemma 30. *Consider any partition (A, B) of the nodes in G with the source node in A and the destination node in B. Let $P^{(1)}\ P^{(2)}\ \ldots, P^{(k)}$ be the channels associated with edges from A to B, and let $\hat{P} = P^{(1)} \times P^{(2)} \times \ldots \times P^{(k)}$ be the product channel. If there are no back-edges, then for all M, we have*

$$\mathcal{E}_{M,G} \leq \mathcal{E}_{M,\hat{P}}. \tag{116}$$

从一个没有反向边的最小割出发，令 $P^{(1)}, P_{(2)} \ldots, P_{(k)}$ 为穿过该割的信道。由于我们的成对可逆假设，我们可以应用引理 17 得到

$$\mathcal{E}_{M,\hat{P}} = \sum_i \mathcal{E}_{M,P^{(i)}} = |f_{M,G}|, \tag{117}$$

其中，最后一步由最大流-最小割对偶性成立。这就完成了定理 29 的逆向证明部分。可达性部分可由定理 1 直接推出。

接下来，对于零速率错误指数，我们证明可以舍弃成对可逆性条件：

引理 31. 假设 G 有一个不含反向边的最小割。那么，我们有 $E_G(0) = |f|$，其中 $|f|$ 是通过将每条边 e 的容量设为 $E_{P_e}(0)$ 而形成的最大流。

证明。与 (117) 类似，引理 17 的证明表明 国产成人

$$\tilde{\mathcal{E}}_{M,\hat{P}} = \sum_i \tilde{\mathcal{E}}_{M,P^{(i)}}. \tag{118}$$

取 $M \to \infty$ 并应用 (105)，我们得到

$$E_{\hat{P}}(0) = \sum_i E_{P^{(i)}}(0). \tag{119}$$

在相同的 e 条件与引理 30 相同，我们可以类似地得出 $E_G(0) \leq E_{\hat{}P}(0)$，因此 re

$$E_G(0) \leq \sum_i E_{P^{(i)}}(0), \tag{120}$$

其中 co 完成了逆向部分。可达性部分是定理 4 的直接推论 . □

B. 近似保证

引理 27 和定理 2 表明，对于任意网络 G 和任意 M，我们有

$$|\tilde{f}_{M,G}| \leq \mathcal{E}_{M,G} \leq |f_{2,G}|, \tag{121}$$

˜其中我们回顾一下，$f_{M,G}$ 在最大流计算中使用式（28）作为边权重。接下来我们将证明，这实际上是一个 4 倍近似：

引理 32. 对于任意 M 和 G，有

$$|f_{2,G}| \leq 4|\tilde{f}_{M,G}|. \tag{122}$$

˜ ˜证明。只需证明 $\mathcal{E}_{2,P} \leq 4\mathcal{E}_{M,P}$，这意味着 $f_{2,G}$ 和 $f_{M,G}$ 是基于彼此相差至多 4 倍的边权定义的。我们有

$$\begin{aligned}
\mathcal{E}_{2,P} &= \max_{x,x'} d_{\mathrm{C}}(x, x', P) && (123)\\
&= \max_{x,x'} \max_{s\in[0,1]} d_{\mathrm{C}}(x, x', P, s) && (124)\\
&\leq 2 \max_{x,x'} d_{\mathrm{C}}(x, x', P, 1/2) && (125)\\
&= 2\tilde{\mathcal{E}}_{2,P}, && (126)
\end{aligned}$$

where (123) follows from (32), (125) follows by using the concavity of d_{C} in s [13, Thm. 5] and considering an equal combination of the values at s and $1-s$ (along with $d_{\mathrm{C}} \geq 0$ for the latter), and (126) follows from (28).

接下来，令 (x, x') 取值以最大化 $d_{\mathrm{B}}(x, x', P, 1/2)$。通过考虑将质量 $1/2$ 分配给 x 、将质量 $1/2$ 分配给 x' 的分布，我们有

$$\tilde{\mathcal{E}}_{M,P} \geq E_P(0) \geq \frac{1}{2} d_{\mathrm{B}}(x, x', P, 1/2) = \frac{1}{2} \tilde{\mathcal{E}}_{2,P}, \tag{127}$$

其中，我们首先应用引理 28，接着应用式 (26)，并将最大值替换为上述分布，然后应用式 (28)，其中 $M = 2$。结合式 (126) 和式 (127) 即得引理 32。 □

下一个引理给出了可以实现更紧近似的条件：

引理 33. 假设以下条件之一成立：

- $M = 2$；或
- G 中的每个信道都是两两可逆的。

˜因此，我们在 (121) 中得到一个 2-近似，即 $|f_{2,\ G}| \leq 2|f_{M,\ G}|$。

证明。当 $M = 2$ 时，结论立即由（126）得出。此外，如果每个信道都是两两可逆的，那么 $\mathcal{E}_{2,\ P} = \mathcal{E}_{2,\ P}$，因此 2-近似结果由（127）得出。 □

我们还发现了另一种我们的界恰好紧致的情形。考虑输入为 $\{1, 2, \ldots, K\}$、输出为 $\{1, 2, \ldots, K\}$ 的 K 元对称信道，以及

$$P(y|x) = \begin{cases} 1 - (K-1)p & y = x \\ p & y \neq x \end{cases} \tag{128}$$

对于某些 $p \in \left(0, \frac{1}{K-1}\right)$。

引理34。假设每个信道对于某个 $K \geq M$ 都是 K 元对称信道（不同信道可以˜具有不同的 K 值以及不同的噪声水平）。那么 $|f_{M,\ G}| = |f_{2,\ G}|$，且第五节给出的协议提供了可达性部分。

˜证明。与本小节前面的结果类似，我们只需证明对每个 P_e都有$\mathcal{E}_{M,P_e} = \mathcal{E}_{2,P_e}$。˜回顾 $\mathcal{E}$的定义：

$$\tilde{\mathcal{E}}_{M,P_e} = \frac{2}{M(M-1)} \max_{(x_1,\ldots,x_M)} \sum_{1 \leq m_1 \leq m_2 \leq K} d_{\mathrm{B}}(x_{m_1}, x_{m_2}, P_e). \tag{129}$$

For the K-ary symmetric channel, $d_{\mathrm{B}}(x, x', P_e)$ ie either 0 (if $x = x'$) or a fixed positive value (if $x \neq x'$). Since $K \geq M$, we can achieve the maximum in (129) by letting $x_1, \ldots, x_M$ be distinct, yielding

$$\tilde{\mathcal{E}}_{M,P_e} = -\log(2\sqrt{(1-(M-1)p) \cdot p} + (M-2)p). \tag{130}$$

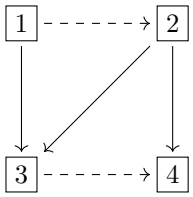


图 2. 一个唯一最小割（即 $A=\{1,3\}$ 和 $B=\{2,4\}$）包含一条反向边的示例。实线箭头的容量为无穷大，虚线箭头的容量有限。

对于$\mathcal{E}_{2,\ P_e}$ 的计算也很简单；由于 P_e 是成对可逆的，我们只需选择任意两个不同的输入即可得到

$$\mathcal{E}_{2,P_e}=\max_{x_1,x_2} d_{\mathrm{B}}(x_1,x_2,P_e)=-\log(2\sqrt{(1-(M-1)p)\cdot p}+(M-2)p). \tag{131}$$

由于 $\mathcal{E}_{M,P_e}=\mathcal{E}_{2,P_e}$ 对所有边都成立，因此可得 $|f_{M,G}|=|f_{2,G}|$。 □

C. 一个自然猜想的反例

在定理 29中，我们假设 G 存在一个不含反向边的最小割。人们可能会倾向于认为，只要 G 不含有向环，情况就总是如此。然而，事实并非如此，正如图 2 所示，其中的最小割在图注中有所描述。

利用图 2 中的例子，我们证明，即使每个信道都是成对可逆的，且图是无环的，误差指数仍可能大于 $|f_{M,G}|$ 当 $M=3$:

Theorem 35. *Even if we assume that G is acyclic and every channel is pairwise reversible, there exist scenarios in which $\mathcal{E}_{M,G}>|f_{M,G}|$ when $M=3$. Moreover, the same statement holds even when the 1-hop error exponents defining $f_{M,G}$ are replaced by their counterparts with full feedback.*

证明。考虑图 2。设定 $M=3$，并考虑如下针对某些 $p\in\left(0,\ \frac{1}{3}\right)$ 所描述的设置:

- 在实线边上，允许进行无限的无噪声通信。
- On the edge $1\to 2$, we consider a ternary symmetric channel with inputs $\{0,1,2\}$, outputs $\{0,1,2\}$, and with $\mathbb{P}(y|x)=1-2p$ if $y=x$, and $\mathbb{P}(y|x)=p$ otherwise.
- 在边 $3\to 4$ 上，我们考虑一个噪声参数为 p 的二元对称信道。

我们将证明，当 p 充分小时，可以超过最大流速率。本节其余部分的渐近记号均指 $p\to 0$ 时的极限。

Let us first compute the error exponents of all the channels. Since all channels are pairwise reversible, we may compute the error exponents using $\mathcal{E}_{M,P}=\tilde{\mathcal{E}}_{M,P}$ as per (28). For the ternary symmetric channel, (28) is maximized when x_1,x_2,x_3 are all distinct, so the error exponent is given by

$$\mathcal{E}=-\log(2\sqrt{p(1-2p)}+p)=\frac{1}{2}\log\frac{1}{p}+\mathcal{O}(1). \tag{132}$$

对于二元对称信道， x_1、 x_2、 x_3 中最多只有两个可以不同，因此错误指数为

$$\mathcal{E} = -\frac{2}{3}\log(2\sqrt{p(1-p)}) = \frac{1}{3}\log\frac{1}{p} + \mathcal{O}(1). \tag{133}$$

因此，最大流就是两个错误指数之和，即$_{5\ 6}$ $\log\ \frac{1}{p} + \mathcal{O}(1)$。

我们将证明 $\log\ \frac{1}{p} + \mathcal{O}(1)$ 是一个可实现的错误指数，从而推翻关于最大流是错误指数上界这一合理猜想。与第 IV 节类似，将一次“顺序传输”定义如下：

- 节点 1将真实消息 m 输入其连接到节点 2 的信道，并且还向节点 3 发送 m ；
- 接着，节点 2 将其收到的内容告知节点 3 和节点 4；
- 节点 3 随后会查看来自节点 1 的真实值和节点 2 转发的值。如果两者相同，就向节点 4 发送消息 0；否则，向节点 4 发送消息 1。

A total of $n-3$ such transmissions can occur in the n time steps by chaining them one after the other (similar to Figure 1). Moreover, we can view a single sequential transmission as a discrete memoryless channel Q, where the input is m, and the output consists of what node 4 receives from both nodes 2 and 3.

我们可以将每次顺序传输的结果分类为三种可能的结果：

(i) Node 4 receives message 1 from node 3. This means that either node 2 received the wrong message (and therefore node 3 sent 1), or the message from node 3 to node 4 got flipped. Each event happens with probability $\mathcal{O}(p)$, so the total probability of this is $\mathcal{O}(p)$.

(ii) 节点 4 从节点 3 接收到消息 0，并从节点 2 接收到错误的消息。在这种情况下，从 1 到 2 以及从 3 到 4 的消息都必须被破坏，因此这种情况发生的概率为 $\mathcal{O}(p^2)$

(iii) 节点 4 从节点 3 接收到消息 0，并从节点 2 接收到正确消息。发生此情况的概率为 $1-\mathcal{O}(p)$。

□

我们在图 3 中总结了端到端转移概率：

		output			
		(0,0)	(1,0)	(2,0)	(*,1)
	0	$1-\mathcal{O}(p)$	$\mathcal{O}(p^2)$	$\mathcal{O}(p^2)$	$\mathcal{O}(p)$
input	1	$\mathcal{O}(p^2)$	$1-\mathcal{O}(p)$	$\mathcal{O}(p^2)$	$\mathcal{O}(p)$
	2	$\mathcal{O}(p^2)$	$\mathcal{O}(p^2)$	$1-\mathcal{O}(p)$	$\mathcal{O}(p)$

图 3。输入表示节点 0 收到的消息。输出 (a,b) 表示 a 是从节点 2 收到的， b 是收到的从节点 3 接收。（$*,1$）表示从节点 3 接收 1，同时节点 2 传来的信息任意（且被忽略）。

由于我们可以在 n 个时间步内使用 Q $n-3$ 次，因此有 $\mathcal{E}_{3,\ G} \geq \mathcal{E}_{3,\ Q}$。现在我们继续对 $\mathcal{E}_{3,\ Q}$ 进行界定；我们有

$$d_{\mathrm{B}}(0,1,Q) = -\log\sum_y \sqrt{Q_0(y)Q_1(y)} = -\log\sqrt{\mathcal{O}(p^2)} = \log\frac{1}{p} + \mathcal{O}(1) \tag{134}$$

类似地，当 (0, 1) 被任意两个不同的符号替换时也成立。利用引理 12，可得

$$\mathcal{E}_{3,Q} \geq \log \frac{1}{p} + O(1), \tag{135}$$

表明可以达到 log $\frac{1}{p} + O(1)$ 的错误指数。

This same counterexample in fact also establishes the second part of Theorem 35 in which the mincut is defined with respect to edge capacities $\mathcal{E}_{P_e}^f$ instead of $\mathcal{E}_{P_e}$. To see this, first note that for the ternary symmetric channel, we have $\mathcal{E}_3 = \mathcal{E}_2$; this follows from the proof of Lemma 34, with the ternary symmetric channel being pairwise reversible. On the other hand, $\mathcal{E}_2 = \mathcal{E}_2^f$ (Theorem 18), $\mathcal{E}_2^f \geq \mathcal{E}_3^f$ (error exponent is decreasing in M) and $\mathcal{E}_3^f \geq \mathcal{E}_3$ (feedback cannot decrease error exponents). Therefore, the ternary symmetric channel has $\mathcal{E}_2 = \mathcal{E}_2^f \geq \mathcal{E}_3^f \geq \mathcal{E}_3 = \mathcal{E}_2$; this means that all the inequalities must hold with equality, yielding

$$\mathcal{E}_3^f = \mathcal{E}_3 = \mathcal{E}_2 = -\log(2\sqrt{p(1-2p)} + \sqrt{p \cdot p}) = \frac{1}{2}\log\frac{1}{p} + \mathcal{O}(1). \tag{136}$$

对于二元　对于 y 对称信道，　$M = 3$ 的反馈错误指数如下给出 [16, o。64　　]:

$$\mathcal{E}_3^f = -\log(p^{1/3}(1-p)^{2/3} + p^{2/3}(1-p)^{1/3}) = \frac{1}{3}\log\frac{1}{p} + \mathcal{O}(1), \tag{137}$$

这给出了最大流 5 6 log $\frac{1}{p}$ + $\mathcal{O}(1)$。再次，由于当p足够小时，该值严格小于 log $\frac{1}{p} + \mathcal{O}(1)$，我们便得到所需结果。

参考文献

[1] W. Huleihel, Y. Polyanskiy, and O. Shayevitz, "Relaying one bit across a tandem of binary-symmetric channels," *IEEE International Symposium on Information Theory (ISIT)*, 2019.

[2] V. Jog and P. L. Loh, "Teaching and learning in uncertainty," *IEEE Transactions on Information Theory*, vol. 67, no. 1, pp. 598–615, 2021.

[3] Y. H. Ling and J. Scarlett, "Optimal rates of teaching and learning under uncertainty," *IEEE Transactions on Information Theory*, vol. 61, no. 11, pp. 7067–7080, 2021.

[4] ——, "Multi-bit relaying over a tandem of channels," *IEEE Transactions on Information Theory (to appear)*, 2023.

[5] S. Rajagopalan and L. Schulman, "A coding theorem for distributed computation," *ACM Symposium on Theory of Computing*, 1994.

[6] Y. H. Ling and J. Scarlett, "Simple coding techniques for many-hop relaying," *IEEE Transactions on Information Theory*, vol. 68, no. 11, pp. 7043–7053, 2022.

[7] E. Domanovitz, T. Philosof, and A. Khina, "The information velocity of packet-erasure links," in *IEEE Conference on Computer Communications*, 2022.

[8] W. Evans, C. Kenyon, Y. Peres, and L. J. Schulman, "Broadcasting on trees and the Ising model," *Annals of Applied Probability*, vol. 10, no. 2, pp. 410–433, 2000.

[9] A. Makur, E. Mossel, and Y. Polyanskiy, "Broadcasting on random directed acyclic graphs," *IEEE Transactions on Information Theory*, vol. 66, no. 2, pp. 780–812, 2020.

[10] ——, "Broadcasting on two-dimensional regular grids," *IEEE Transactions on Information Theory*, vol. 68, no. 10, pp. 6297–6334, 2022.

[11] A. El Gamal and Y.-H. Kim, *Network information theory*. Cambridge University Press, 2011.

[12] T. M. Cover and J. A. Thomas, *Elements of information theory*. John Wiley & Sons, Inc., 2006.

[13] C. Shannon, R. Gallager, and E. Berlekamp, "Lower bounds to error probability for coding on discrete memoryless channels. i," *Information and Control*, vol. 10, no. 1, p. 65–103, 1967.

[14] ——, "Lower bounds to error probability for coding on discrete memoryless channels. ii," *Information and Control*, vol. 10, no. 5, p. 522–552, 1967.

[15] L. Trevisan, “Stanford CS261 lecture notes,” 2011, https://theory.stanford.edu/ trevisan/cs261/lecture11.pdf.

[16] E. R. Berlekamp, “Block coding with noiseless feedback,” *Ph.D. thesis, Massachusetts Institute of Technology, Department of Electrical Engineering*, 1964.

[17] B. Nakiboglu, “Exponential bounds on error probability with feedback,” *Ph.D. thesis, Massachusetts Institute of Technology, Department of Electrical Engineering*, 2011.