The 大输出函数的通信复杂度

Lila Fontes* Sophie Laplante† Mathieu Lauri`ere‡ Alexandre Nolin§

arXiv:2304.00391v1 [cs.CC] 1 Apr 2023

摘要

我们研究大输出函数的双向通信复杂度，并表明通信复杂度会因所考虑的输出模型不同而有很大差异。我们研究了多种输出模型，从开放模型（外部观察者可以计算结果）到异或模型（协议的结果应为参与者各自局部输出的按位异或）。该模型的灵感源自异或博弈，这类博弈在双参与者量子博弈研究中受到广泛关注。

我们着重探讨这些新型输出模型中的错误缩减问题。对于输出大小为 k 的函数，在异或模型中应用标准错误缩减技术会引入一个与 k 成线性关系的额外代价。我们证明，无须依赖 k 。类似地，标准随机性消除技术在异或模型中会带来 2^k 的乘法代价。我们展示了如何将这一因子降至 $O(k)$ 。

此外，我们还证明了其他模型中类似的错误缩减和随机性消除结果，区分了彼此之间的所有模型，并表明，当输入的汉明重量有界时，一些自然问题（包括集合交集和查找第一个差异）可以区分这些模型。最后，我们展示了如何将秩下界技术用于我们的弱输出模型。

1 引言

关于通信复杂度这一主题的大部分文献都聚焦于布尔函数。通常的定义规定，在协议结束时，其中一位参与者知道函数的值。在基于矩形的下界中，假设稍强一些：协议结束时，通信记录确定一个组合矩形，其中所有输入都得到相同的结果。这意味着，给定通信记录（以及公币随机设置下的公共硬币），外部观察者可以确定输出。对于布尔函数而言，这一假设并无实质影响，因为知道函数值的参与者可以在协议的最后一条消息中发送该值，额外代价至多为一比特。然而，当函数的输出很大时，将函数值作为通信记录的一部分发送，所需代价可能超过此前所有通信的总和。遇到这种情况时，问题的“真正”通信复杂度应如何定义？

研究输出规模较大的函数时，会出现若干基本问题。哪些下界技术可以推广到非布尔函数？在组合输出规模较大的协议时，让双方参与者都知道中间函数的取值可能并无用处，而且在每个中间步骤传递结果所产生的累计成本，可能会超过组合问题本身的复杂度。这些问题同样适用于信息复杂度，其中成本以信息论术语衡量，而非以通信比特数衡量。要求协议将结果作为通信记录的一部分公开，可能会成为寻找信息量极低的协议的障碍。这还引出了以下问题：当输出规模较大时，如何放大成功概率？放大方案通常通过重复

*斯沃斯莫尔学院 – fontes@cs.swarthmore.edu fontes@cs.swarthmore.edu

†IRIF，巴黎西岱大学´e 巴黎西岱´e– laplante@irif.fr laplante@irif.fr

‡纽约大学上海校区—纽约大学华东师范大学数学科学研究所 – ml5197 5197@nyu.edu ml@nyu.edu§CISPA 亥姆霍兹信息安全中心 – alexandre.nolin@cispa.de alexandre.nolin@cispa.de

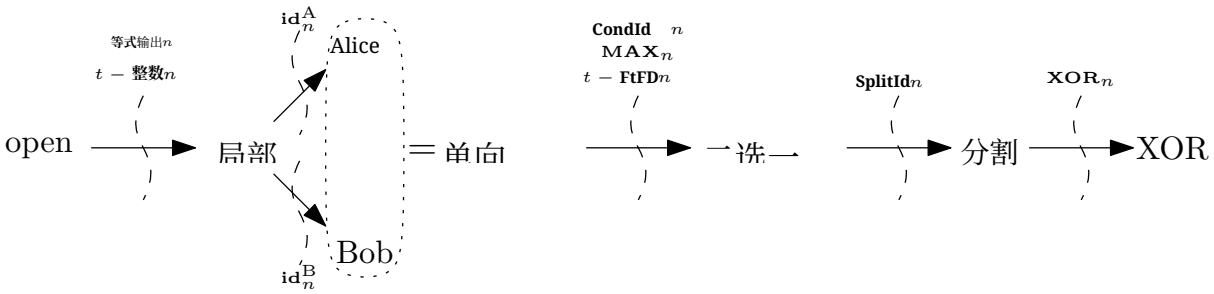


图1：大输出函数的各种通信模型及用于区分这些模型的问题。从 A 指向 B 的箭头表示，针对 A 类任务的通信协议也适用于 B 类任务。关于这些分离的详细信息见附录A附录A 和 HH。

执行协议并取多数结果来实现，但天真地求出该多数结果会产生取决于输出长度的成本。我们探讨这些问题，并提出新的模型和放大方案。

经过深入研究的大输出函数示例包括非对称博弈，例如 NBA问题 [Orl9090, Orl Orl91] Orl（另见 [KN9797, KN 示例 4.53，第 64] 页），以及许多输出本质上与输入大小相同的问题（例如，计算两个集合的交集 [BCK+14BCK , BCK+16]BCK）。函数的大输出形式对应的判定问题，其通信复杂度可能相近（例如，集合不相交性 [KS9292, KS Raz9292, Raz BYJKS04]BYJKS），也可能截然不同（例如，判断双方的数字之和是否大于给定常数 [Nis9393, Nis Vio15]Vio）。研究同一函数的多个实例是否具有规模经济时，也会出现大输出函数；这类问题称为直接和问题，其变体包括一致性和消除 [Aar0505, Aar ABG+01ABG, BDKW14]BDKW。在这些问题及其他问题中，计算输出的一个比特，可能与计算完整输出同样困难，也可能容易得多，这取决于函数及所采用的模型。最后，模拟协议的输出是另一个协议的通信记录，它们在压缩 [BR1414, BR Bra1515, Bra Kol1616, Kol She1818, She BK18] BK 以及结构性结果 [BBK+16BBK, GKR1616, GKR GKR2121,GKR RS18]RS 的研究中发挥了关键作用。查找第一个差异问题一直是压缩协议研究中的重要工具。当输出条件较弱时，已知存在更优的协议 [BBCR1313, BBCR BMY15]BMY。

1.1 输出模型

我们提出了几种自然的替代模型，以取代这样一种模型：通信记录和公开随机性会揭示函数的值（即使没有明确包含该值）（我们称之为开放模型）。在局部模型中，两个参与者都能在局部确定函数的值（但外部观察者未必能做到这一点——这与开放模型不同）。在单向模型中，一名参与者总能获知答案。在二选一模型中，获知答案的参与者可以有所不同。在分割模型中，输出的比特以双方都知道的任意方式分配给两个参与者。最后，在异或模型中，每名参与者都输出一个字符串，结果是这些输出按位异或的结果。这些模型构成一个层级，如图 11图 所示。形式化定义见附录 A附录 A。

在协议的语境下，我们区分参与者的输出与协议计算出的结果。例如，在异或模型中，参与者输出字符串 a 和 b，而协议的结果是 $a \oplus b$。我们用“输出”指参与者在协议结束时给出的内容，用“结果”或“结果值”指协议的结果（它应当——以概率方式或确定性地——等于函数的值或输出）。同样，我们用“协议”指产生结果的完整机制，用“通信协议”指协议中参与者交换消息的交互部分，不包括输出机制。

在我们提出的所有模型中，异或模型或许最有趣。这个模型部分受到（量子）异或博弈的启发，在这类博弈中，参与者之间不会交换任何

消息（例如 [Bel64]，[BelPV16]，[PVBCMdW10]）。异或模型的一个有趣性质是，例如，每个参与者单独来看，其输出可能服从均匀分布[1]，，既不透露任何输入的信息，甚至也不透露函数作为黑盒运行时的取值。

此外，在通信复杂度领域，常见的做法是研究由某个应用于输入的“小工具”组合而成的布尔函数的复杂度。例如，对于布尔函数 f，我们可以询问 $F(u, v) = f(u \oplus v)$ 的通信复杂度是多少，其中按位异或被用作应用于输入的小工具。异或模型可以看作是将异或小工具应用于输出而非输入：参与者输出 (a, b)，并且我们要求$F(u, v) = a \oplus b$，才能保证计算正确。

1.2 我们的贡献

我们聚焦于异或模型：每位参与者都输出一个字符串，协议的结果是这些字符串的按位异或。

错误缩减。我们在第 5 节第 5 节中研究错误缩减问题。错误缩减通常很简单：将计算重复足够多次，然后取多数结果。然而，在异或模型中，两位参与者都不知道任何一次计算的结果，因此若不进行额外通信，双方都无法计算多数结果。将所有结果发送给其中一位参与者以便其计算多数结果，会增加一个代价高昂的 $\Theta(k)$ 项，其中 k 是输出的长度。不过，可以消除对 k 的依赖，而要做到这一点，需要相当复杂的协议，这些协议凸显了异或模型固有的局限性（**定理 5.3**。）。

我们进一步改进了 ϵ 直接和问题（**定理 5.8**中误差参数的依赖关系），方法是结合摊销相等性 [FKNN95] FKNN和查找第一个差异 [FRPU94], FRPU以及间隙汉明距离 [IW03、IWCR12、CRVid12、VidShe12]She。

确定性与随机化复杂度。在第6节6中，第我们重新探讨了经典结果：对于任意布尔函数，其确定性通信复杂度至多为私有硬币随机化复杂度的指数级。当输出大小为 k时，直接将现有方案应用于我们较弱的模型，会带来 2^k的乘法开销。我们证明，依赖因子 k 就足够了（**定理 6.4**。）。

异或与间隙多数的复合。为了证明异或模型下的结果，我们考虑非布尔的间隙多数问题与异或小工具的复合。在标准多数问题中，输入是一组元素，目标是找出出现频率最高的元素。间隙多数问题额外承诺多数元素出现的频率至少达到某个固定比例（超过一半）。与异或小工具复合后，该问题就转化为一个通信复杂度问题（见第 5 节第5节和附录 G附录 G）。我们证明，该问题的通信复杂度与异或模型下的降低错误和去随机化问题密切相关。

其他模型与分离结果。我们定义了几种通信模型，并给出能够最大程度区分这些模型的问题（附录 A附录 A）。我们在其他模型中重新考察错误缩减和随机性消除（附录 D附录 D 和 FF）。二选一模型的去随机化方案巧妙地运用了 NBA问题，作为协调两位参与者的多数候选项过程的一部分。我们将对 k 的依赖降低到一个因子

[1]该模型中的任何协议都可以转换为具有相同复杂度且满足以下性质的协议：参与者选择一个长度与输出相同的共享随机字符串 r，并输出 $a \oplus r\,(b \oplus r)$，其中 a，b 是原始协议的输出。

log(k)，并在误差参数 ϵ 有界于 $1/3$(定理 F.44 定理 F.）时彻底消除这种依赖。

最后，我们研究了几个其他问题，它们在我们的各种通信模型之间呈现出差距。特别是，若输入的汉明重量有界，若干常见问题会呈现出差距（附录 H附录 H）。

秩下界。 我们通过重新考察与协议树叶节点相关的单色矩形概念，展示如何将下界技术应用于弱输出模型。我们聚焦于确定性通信的秩下界，并证明它可用于我们所有的模型，包括异或模型。（第 7 节第 7 节）

需要注意的是，我们的结果大多不适用于大输出关系 （例如直和、消除和一致性的各种变体），因为我们的许多证明都关键地依赖于只有一个正确答案这一事实。

2 相关工作

以往的研究探讨了大输出函数的输出模型问题。Braverman 等人 [BRWY13] BRWY区分了协议的“模拟”和“强模拟”。在强模拟中，外部观察者无需了解任何输入即可确定结果。在他们关于压缩为内部信息的论文中[BMY15], BMYBauer 等人强调，将信息压缩为内部信息时，压缩本身不必向外部观察者泄露信息。他们考虑了两种输出模型，分别称为内部计算和外部计算。在外部计算中（我们称之为开放模型），外部观察者可以确定协议的结果；而在内部计算中（我们称之为局部模型），参与者双方都会在协议结束时确定结果。[2] 他们指出，在确定性情形下，对于全函数，这两种模型是一致的；但在分布式情形下，两者可能有所不同。他们研究了一个关键问题：当每位参与者各持有两个字符串中的一个时，查找这两个字符串首个不同的比特。该问题用于协调协议，以找出通信记录出现差异的首个位置。Feige 等人 [FRPU94] FRPU以外部（开放）方式解决查找第一个差异问题，复杂度为 $O(\log(\frac{n}{\epsilon}))$，Viola [Vio15] Vio证明了这一界是紧的。Bauer 等人 [BMY15] BMY给出了一种复杂度更低的内部（局部）协议，其改进程度取决于输入分布的熵。

3 预备知识

通信复杂度简介可见 Kushilevitz 和 Nisan 的 [KN97],KN以及 Rao 和 Yehudayoff 的 [RY20]RY教材。

我们用 $\mathcal{X}$（分别用 $\mathcal{Y}$）表示 Alice（分别为 Bob）的输入集合， $\mathcal{R}_A$ 表示其私有随机性（$\mathcal{R}_B$ Bob 的私有随机性）， $\mathcal{R}^{pub}$ 表示两个参与者都可以访问的公共随机性。当 $|\mathcal{X}| = |\mathcal{Y}|$ 时，我们用 n 表示输入的大小（因此 $n = \lceil \log(|\mathcal{X}|) \rceil$）。计算函数时，我们用 k 表示输出的长度，用 $\mathcal{Z}$ 表示函数的像，并用 $k = \lceil \log(|\mathcal{Z}|) \rceil$表示。此外，我们有时会考虑一个额外的输出符号 $\top$。

我们将完整协议定义为通信协议与输出机制的组合（详见附录 A附录 A）。我们将一个（双参与者）通信协议定义为 Π一棵满二叉树，其中每个非叶节点 v 都分配给一个 $\mathcal{P}^v$ 参与者 A（lice）或 B（ob），并有一个映射 $\mathcal{N}^v$ 至 $\{0,1\}$，其输入空间取决于该节点分配给哪位参与者。当 $\mathcal{P}^v = A$（分别为 B）时， $\mathcal{N}^v$的输入空间为 $\mathcal{X} \times \mathcal{R}_A \times \mathcal{R}^{pub}$（分别为 $\mathcal{Y} \times \mathcal{R}_B \times \mathcal{R}^{pub}$）。请注意，树及每个节点的所有者都是固定的，不依赖于输入。在通信协议的执行过程中，两位参与者共同沿树向下移动，从根节点开始

[2]为避免混淆计算与 *internal* 以及信息与 *external* 这两个概念，我们倾向于使用 *open* 和 *local* 这两个术语。

直到到达叶节点。每一步都由拥有当前节点的参与者执行其对应的映射 $\mathcal{N}^v$，并将结果发送给另一位参与者。若结果为 0，参与者就将当前节点替换为其左子节点；否则替换为右子节点。协议 Π 的通信成本CC(Π)是最坏情况下输入所需交换的比特总数。

由于通信协议 Π 的一次执行完全由参与者的输入（$(x,y) \in \mathcal{X} \times \mathcal{Y}$）和随机性（参与者的私有随机性 $r_{\mathsf{A}} \in \mathcal{R}_{\mathsf{A}}$和 $r_{\mathsf{B}} \in \mathcal{R}_{\mathsf{B}}$，以及公共随机性 $r \in \mathcal{R}^{\mathsf{pub}}$）所确定，我们也将通信协议视为一个函数Π： $\mathcal{X} \times \mathcal{Y} \times \mathcal{R}_{\mathsf{A}} \times \mathcal{R}_{\mathsf{B}} \times \mathcal{R}^{\mathsf{pub}} \to \{0,1\}^*$，其取值称为 Π的通信记录。本文中，我们不将公共随机性纳入通信记录。对于给定的协议 Π，我们用 $T_\pi = \Pi(X, Y, R_{\mathsf{A}}, R_{\mathsf{B}}, R)$表示由作为随机变量的 X、 Y、 R_{A}、 R_{B}和 R自然导出的、取值为该协议通信记录的随机变量。我们用 $\mathcal{T}_\pi$表示分布 T_π的支撑集。我们用 x、 y、 z、 r_{A}、 r_{B}、 r、 t_π分别表示集合$\mathcal{X}$、 $\mathcal{Y}$、 $\mathcal{Z}$、 $\mathcal{R}_{\mathsf{A}}$、 $\mathcal{R}_{\mathsf{B}}$、 $\mathcal{R}^{\mathsf{pub}}$、 $\mathcal{T}_\pi$中的元素；这些集合依次是随机变量X、 Y、 Z、 R_{A}、 R_{B}、 R、 T_π的支撑集。

我们回顾本文将用到的函数定义及已知界。对于所有这些问题，请注意，无论是要求两位参与者都知道输出，还是只要求其中一位知道输出，其通信复杂度的数量级都相同，因为输出的大小不超过让一位参与者知道输出所需的通信量。在本节余下部分中，我们用 $R_\epsilon(f)$ 表示计算函数 f时错误至多为 ϵ且由 Bob 输出结果的随机化协议的最小通信成本。 $D(f) = R_0(f)$ 表示确定性通信复杂度。

定义 3.1 （查找第一个差异问题）。$\mathbf{FtFD}_n$： $\{0,1\}^n \times \{0,1\}^n \to \{0,\ldots,n\}$ 定义为$\mathbf{FtFD}_n(x, y) = \min(\{i : x_i \neq y_i\} \cup \{n\})$。

命题 3.2。对于任意$0 < \epsilon < \frac{1}{2}$， $R_\epsilon(\mathbf{FtFD}_n) \in \Theta(\log(n) + \log(1/\epsilon))$ [FRPU94, FRPUVio15]Vio。

上界利用树上的游走，其中每一步都依据哈希函数的结果来决定。下界则来自大于函数 $\mathbf{GT}_n$的下界，而该函数可归约到 $\mathbf{FtFD}_n$。关于上界的精彩阐述，参见[BBCR13]BBCR中的附录 C。

定义 3.3 (间隙汉明距离问题)。 设 n、 L、 U为满足 $0 \leq L < U \leq n$的整数。$\mathbf{GHD}_n^{L,U}$： $\{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ 是一个承诺问题，其输入满足以下承诺：输入 x、 y 之间的汉明距离为 $\geq U$ 或 $\leq L$。那么 $\mathbf{GHD}_n^{L,U}(x,y) = 1$ 在第一种情况下取值，而在第二种情况下取值 0 。

间隙汉明距离的界因参数而异。本文采用的线性上界在我们所需的参数范围内基本是紧的。对于其他参数范围，已有许多其他界 [Koz1515, KozCR1212, CRVid1212, VidShe1212, SheBCW9898, BCWWat18]Wat。

定义 *3.4*（相等性问题）。函数 EQ_n： $\{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ 定义为 $EQ_n(x,y) = \mathbf{1}_{x=y}$。 k重相等性问题为 $EQ_n^{\otimes k}((x_1,\ldots,x_k),(y_1,\ldots,y_k)) = (EQ_n(x_1,y_1)\ldots,EQ_n(x_k,y_k))$，其中对于所有 i，$(x_i, y_i) \in \{0,1\}^n$。

命题 3.5。对于 $0 < \epsilon < \frac{1}{2}$， $R_\epsilon(\mathbf{EQ}_n^{\otimes k}) \in \Theta(k + \log(1/\epsilon))$。

来自 [HPZZ21] HPZZ的算法通过哈希实现最优通信，与处理单个实例的算法一样。与 k 连续使用错误率为 ϵ/k 的相等性协议相比，它让参与者同时对所有 k 个实例进行哈希、交换结果并重复这一过程，从而节省通信量，因为需要通信的信息越来越少。直观地说，随着算法运行，需要找出的不相等实例数量应当逐渐减少。一旦通过不相等的哈希值确定实例（x_i, y_i）满足 $x_i \neq y_i$，参与者就无需再就该实例进行通信。不相等的实例不太可能经受住多轮测试，因此在算法后期，参与者可以交换

哈希值，因为其中大多数应该相同。此前改进平凡算法的算法 [FKNN95] FKNN中也采用了这一思路。下界直接来自：发送 k比特的信息至少需要 Ω(k 比特的通信量，即使错误率为 ϵ 。

除非另有说明，我们的协议都使用私有硬币和公共硬币。当协议和映射无法使用公共随机性时，我们用上标‘priv’表示。

4 异或模型

在异或模型中，每位参与者都输出一个字符串，函数的值是两个输出的按位异或（定义 44.11 定义 .）。该模型的灵感来自异或博弈；这类博弈在量子非定域性和唯一博弈的背景下都得到了广泛研究。

定义 4.1 (异或计算)。 考虑一个输出集合为 $\mathcal{Z}=\{0,1\}^k$ 的函数 f 。如果存在两个映射 $\mathcal{O}_{\mathsf{A}}$ 和 $\mathcal{O}_{\mathsf{B}}$ ，满足 $\mathcal{O}_{\mathsf{A}}$： $\mathcal{T}_\pi\times\mathcal{R}^{\mathsf{pub}}\times\mathcal{R}_{\mathsf{A}}\times\mathcal{X}\to\{0,1\}^k$，并且同样有 $\mathcal{O}_{\mathsf{B}}$： $\mathcal{T}_\pi\times\mathcal{R}^{\mathsf{pub}}\times\mathcal{R}_{\mathsf{B}}\times\mathcal{Y}\to\{0,1\}^k$，使得对所有 $(x、y)\in\mathcal{X}\times\mathcal{Y}$，则称协议 Π 以 ϵ 错误 *XOR* 计算 f 。

$$\Pr_{r,r_{\mathsf{A}},r_{\mathsf{B}}}[\mathcal{O}_{\mathsf{A}}(t_\pi,r,r_{\mathsf{A}},x)\oplus\mathcal{O}_{\mathsf{B}}(t_\pi,r,r_{\mathsf{B}},y)=f(x,y)]\geq 1-\epsilon.$$

我们将 $D^{\mathsf{xor}}(f)$（分别为 $R^{\mathsf{xor}}_\epsilon(f)$）定义为：在异或模型中，以错误率 $\epsilon=0$ （分别地，对于 $0<\epsilon<\frac{1}{2}$，错误率至多为 ϵ）计算 f 的任意协议所需的最佳通信成本。（对于其他模型, –带上标 open、loc、A、B、uni、spl、1of2 的记号也作类似定义。）

5 错误缩减与间隙多数问题

我们研究在较弱的通信模型中，如何降低通信协议的错误率及其成本。在这些模型中，协议的结果并非双方参与者都知晓。正文主要关注更有意思的异或模型情形，其他模型的结果见附录 D.22附录 D。。

标准错误缩减方案通过多次重复执行协议，计算并输出所有执行结果中出现频率最高的值。只要重复执行足够多次，出现频率最高的输出就能以高概率保证正确。利用霍夫丁不等式，可以推导出所需迭代次数的上界。

引理 5.1 (霍夫丁不等式)。 设 $(V_i)^{i\in[N]}$ 是 N 相互独立、期望值为 p 的伯努利试验。我们有 $\Pr\left[\left|\frac{1}{N}\sum_{i=1}^N V_i-p\right|\geq\delta\right]\leq 2\cdot\exp\left(-\frac{\delta^2 N}{2p\ 1-p}()\right)$.

在 Bob 于协议结束时输出函数值的设定下，以下结论成立。

定理 5.2。 *(民间结果，见 [KN97]KN)* 设 $0<\epsilon'<\epsilon<\frac{1}{2}$，且 $C_{\epsilon,\ \epsilon'}=\frac{2\epsilon}{(1-\epsilon)}\left(\frac{1}{2}-\epsilon\right)^{2}\ln\left(\frac{2}{\epsilon'}\right)$。对于所有函数 f： $\mathcal{X}\times\mathcal{Y}\to\mathcal{Z}$， $R^{\mathsf{B}}_{\epsilon'}(f)\leq C_{\epsilon,\ \epsilon'}\cdot R^{\mathsf{B}}_\epsilon(f)$。

请注意，这里 f 是一个函数而非关系，这一点很重要，因为这样才有唯一正确的输出，参与者才能计算多数。

在异或模型中，从若干次 T 次运行中找出多数结果，比在标准模型中困难得多，因为双方都无法确定合理的候选项作为多数答案。交换所有 T k 比特的输出，会得到一个 $R^{\mathsf{xor}}_{\epsilon'}(f)$ $\leq C_{\epsilon,\ \epsilon'}(R^{\mathsf{xor}}_\epsilon(f)\ +k)$。我们证明，这种对 k 的依赖是不必要的。

定理 5.3。 令 $0<\epsilon'<\epsilon<\frac{1}{2}$，$C_{\epsilon,\ \epsilon'}=8\epsilon\left(\frac{1}{2}-\epsilon\right)^{-2}\ln\left(\frac{8}{\epsilon'}\right)$。对于所有 f： $\mathcal{X}\times\mathcal{Y}\to\{0,1\}^k$，$R^{\mathsf{xor}}_{\epsilon'}(f)\leq C_{\epsilon,\ \epsilon'}\cdot R^{\mathsf{xor}}_\epsilon(f)+O\left(C_{\epsilon,\ \epsilon'}\right)$.

为了证明这一结果，我们引入间隙多数（**GapMAJ**）问题，说明定理 55.33 定理 .可归约为求解 **GapMAJ∘异或** (引理 55.55引理 .)，然后给出求解 **GapMAJ∘异或** (定理 55.66定理 .) 的上界。

偏函数 $\mathbf{GapMAJ}_{N,\ k,\ \epsilon,\ \mu}$ 的输入是 N 个长度为 k 的字符串，并承诺存在一个长度为 k 的字符串 z，它在这些 N个字符串中的权重至少为 $(1-\epsilon)$，其中 μ 是 $[N]$中索引上的一个分布。

定义 5.4 （间隙多数）。 在间隙多数问题 $\mathbf{GapMAJ}_{N,\ k,\ \epsilon,\ \mu}$： $\left(\{0,1\}^k\right)^N \to \{0,1\}^k$ 输入为 $(Z_1,\ \ldots,\ Z_N)$，且 μ 是索引 $[N]$上的固定分布。若未另行说明， μ 默认为均匀分布。承诺是存在 $\exists z \in \{0,1\}^k$，使得 $\mu\left(\{i \in [N]:\ \ Z_i = z\}\right) \ \geq \ (1-\epsilon)$。那么

$$\mathbf{GapMAJ}_{N,k,\epsilon,\mu}((Z_i)_{i\in[N]}) = z \quad s.t. \quad \mu(\{i : Z_i = z\}) \geq (1-\epsilon).$$

在 **GapMAJ∘异或**中，参与者获得 N 长度为 k 的字符串，他们的目标是在**GapMAJ** 承诺成立时，计算参与者输入按位异或后的结果。 （注 $k = 1$意，当此时，这等价于间隙汉明距离问题（定义 33.33定义 .），参数为 $L = \epsilon N$， $U = \ (1-\epsilon)\ N$。）

对于输入 $(X_1,\ldots,X_N),(Y_1,\ldots,Y_N)$，我们将 (X_i,Y_i) 称为一行，并称其为 X_i Alice 的第 i 行，以及 Y_i Bob 的第 i 行。作为热身练习，我们证明，错误缩减可归约为求解一个**GapMAJ∘异或**实例。

引理 5.5。设 $0 < \epsilon' < \epsilon < \frac{1}{2}$ 和 $C_{\epsilon,\epsilon'} = 2\epsilon\left(\frac{1}{2}-\epsilon\right)^{-2}\ln\left(\frac{4}{\epsilon'}\right)$。对于每个 $f:\ \ \mathcal{X}\times\mathcal{Y}\to\{0,1\}^k$, $R^{\mathsf{xor}}_{\epsilon'}(f) \leq C_{\epsilon,\epsilon'}\cdot R^{\mathsf{xor}}_{\epsilon}(f) + R^{\mathsf{xor}}_{\epsilon'/2}\left(\mathbf{GapMAJ}^{C_{\epsilon,\epsilon'},k,\frac{1}{4}+\frac{\epsilon}{2}}\circ\mathbf{XOR}\right)$.

引理 55.55 的证明。令 π 是一个协议，它以 ϵ 错误率对 $f(x,y)$ 进行异或计算；令 π' 是一个在异或模型中计算 $\mathrm{GapMAJ}^{C_{\epsilon,\epsilon'},k,\frac{1}{4}+\frac{\epsilon}{2}}\circ\mathrm{XOR}$ 的协议，其错误率为 $\epsilon'/2$。我们考虑以下协议，记为 $\hat{}\pi$：首先运行 π $C_{\epsilon,\epsilon'}$ 次；然后，将此次计算产生的输出作为 π' 的输入，运行后一个协议，并输出其结果。我们如下分析新协议 $\hat{}\pi$ 。第一步产生的输出是 Alice 一侧的字符串 $X_1,\cdots,X_{C_{\epsilon,\epsilon'}}$，以及 Bob 一侧的字符串 $Y_1,\cdots,Y_{C_{\epsilon,\epsilon'}}$。当且仅当 $X_i\oplus Y_i = f(x,y)$ 时， π 的一次运行才是正确的。根据 Hoeffding 界（引理 55.11），取 $N = C_{\epsilon,\epsilon'}$, $V_i = 1$（若 $X_i\oplus Y_i \neq f(x,y)$）和 $V_i = 0$（其他情况）作为 $i = 1,\ldots,N$, $p = \mathbb{E}[V_i] \leq \epsilon$，并令 $\delta = \frac{1}{2}(\frac{1}{2}-\epsilon)$，可得以至少 $1-2e^{-\delta^2 N/(2p(1-p))} \geq 1-\epsilon'/2$ 的概率，出现错误的计算所占比例为 $p+\delta \leq (\frac{1}{2}+\epsilon)/2$，这 N 次计算中亦是如此。换言之，上述字符串未能满足 $\mathrm{GapMAJ}^{C_{\epsilon,\epsilon'},k,\frac{1}{4}+\frac{\epsilon}{2}}\circ\mathrm{XOR}$ 定义中的承诺的概率至多为 $\epsilon'/2$。在未发生这种情况的条件下（即承诺得到满足时）， π' （因此 $\hat{}\pi$）出错的概率至多为 $\epsilon'/2$。总错误率至多为 ϵ'。 □

要利用引理 55.55 推导错误缩减的一般上界， 引理 .只需给出 $R^{\mathsf{xor}}_{\epsilon'}(\mathbf{GapMAJ}_{N,k,\epsilon}\circ\mathbf{异或})$ 的上界。当误差参数较大时（$\epsilon \leq \epsilon'$），**GapMAJ∘异或** 在异或模型中很简单：参与者只需采样一个公共行，并根据该行输出。然而，引理 55.55 Lemma.要求求解一个误差较小的 **GapMAJ∘异或**实例 $\epsilon'/2$，这又让我们回到了原点：寻找一种可应用于 **GapMAJ∘异或** 的错误缩减方案。

在本节余下部分，我们先给出一个 **GapMAJ∘异或** 协议（第 55.11Section. 节），随后介绍一种针对直和函数的错误缩减方案（第 55.22Section . 节）。这两种情况下，我们都利用异或函数的结构以及一个针对行对的相等性协议来找出多数结果。针对直和函数的错误缩减方案是引理 55.55 Lemma .的改进，适用于初始错误非常接近 $\frac{1}{2}$，且计算一比特输出的成本远低于计算完整输出的成本的情形。

5.1 求解 GapMAJ∘异或

给定一个 $\mathbf{GapMAJ}_{N,\ k,\ \epsilon}\circ$**异或**实例，如果 Alice 和 Bob 选取一行，并输出他们在该行上的值，那么得到正确输出的概率为 $\geq 1-\epsilon$。回顾一下，我们希望在不依赖参数 k 的情况下达到错误率 $\epsilon' < \epsilon$；在我们的错误缩减应用中，该参数对应于输出长度。我们证明这是可行的。

定理 5.6. 设 $0 < \epsilon' < \epsilon < \frac{1}{2}$，$R^{\text{xor}}_{\epsilon'}(\mathbf{GapMAJ}_{N,\ k,\ \epsilon}\circ\textbf{异或}) \leq O\left(N + \text{对数}\left(\frac{1}{\epsilon'}\right)\right)$.

证明思路。我们利用这样一个事实：$a \oplus b = a' \oplus b'$ 当且仅当 $a \oplus a' = b \oplus b'$。因此，参与者可以通过求解相等性问题的实例，找出异或后得到相同字符串的行。仅凭这一思路，就足以构造一个 $\mathbf{GapMAJ}_{N,\ k,\ \epsilon}\circ$**异或** 复杂度为 $O\left(N^2 + \text{对数}\left(\frac{1}{\epsilon'}\right)\right)$ 的协议：对所有 $\binom{N}{2}$ 行对计算相等性，以确定多数结果。我们通过使用 Erd˝os-R´ényi 随机图（引理 5.7 引理 .）减少所计算的相等性实例数量，从而改进这一方法。

引理 5.7（式（9.18）在 [ER60]ER中的变体。设 $G(n, p(n))$ 为以下图上的分布：n 个顶点的图中，每条边都以相互独立的概率 $p(n)$被采样。设 $L_1(G)$ 为 G的最大连通分量的大小。那么：

$$\forall \alpha \in [0,1], c \in \mathbb{R}^+, \qquad \Pr[L_1(G(n, c/n)) < (1-\alpha)n] \leq e^{\left(\ln(2) - \frac{\alpha}{2}\left(1-\frac{\alpha}{2}\right)c\right)n}.$$

特别是，当 n 趋于无穷且 $\alpha c > 4\ln(2)$ 时，这个概率趋于0 。

为完整起见，证明见附录 D附录 D。

定理 *5* 的证明*5.6*6定理 。。考虑$\mathbf{GapMAJ}\circ$**异或** 实例，将其表示为一个 $N \times k$ 矩阵，其中 $(X_i)_{i\in[N]}$ 是 Alice 的行，而 $(Y_i)_{i\in[N]}$ 是 Bob 的行。根据$\mathbf{GapMAJ}\circ$**异或** 问题的承诺，我们知道存在一个 $z \in \{0,1\}^k$ 使得 $\{i : X_i \oplus Y_i = z\} \geq (1-\epsilon)N$。现在的目标是让 Alice 和 Bob 找到属于这个大集合的某一行，使得它们异或后得到相同的 k比特串。

设 i 和 j 为两行的索引。两行异或后得到相同字符串这一事件表示为 $X_i \oplus Y_i = X_j \oplus Y_j$，这等价于 $X_i \oplus X_j = Y_i \oplus Y_j$。这意味着，我们可以使用相等性协议来测试任意两行异或后是否得到相同的比特串。

该协议包括以下步骤：

1. 参与者随机选取若干行，选取的行数要足够多，使得以高概率有常数比例的行异或后得到多数元素 z。2. 参与者求解相等性的实例，以找出异或后得到相同字符串的大集合。在每个这样的行集合中，他们选取单独一行。这样，他们便得到常数数量的候选行，这些候选行异或后可能得到多数元素 z。3. 参与者将这些候选行分别与所有行进行比较，以在候选项中作出决定。其中有一行候选行异或后得到的字符串与大多数行相同；该行异或后得到多数元素 z。

步骤 1. 使用公共随机性，Alice 和 Bob 现在从他们所有的行中选取一个大小为 $|S| = T_{\epsilon'} = 50\ln\binom{10}{\epsilon'}$ 的多重集 S。其中的每个元素都以均匀且相互独立的方式选取。根据霍夫丁不等式（引理 5.1 1引理 .），这些执行中有超过 $\frac{2}{5}$ 的概率会异或得到多数元素 z。

步骤 2. 现在我们将 S 视为随机图 $G = G(V,\ E)$ 的顶点 V，其中每条边都以概率 $\frac{c}{|V|}$ 被选中，且 $c > 0$。考虑由对应于异或结果为多数元素z的执行的顶点 $V' \subseteq V$ 所诱导的 G 的子图 G'。根据上一步，我们知道 $|V'| \geq \frac{2}{5}T_{\epsilon'} = 20\ln\binom{10}{\epsilon'}$。子图 G' 是

一个随机图，其中每条边以相同的概率被选中 $\frac{c}{|V|} = \frac{c'}{|V'|}$ 其中$c' = c\frac{|V'|}{|V|} \geq \frac{2}{5}c$。根据引理 5.7，引理。该子图 G' 以概率 $\geq 1 - 2^{-|V'|} \geq 1 - \frac{\epsilon'}{5}$ 包含大小为 $\geq (1 - \frac{1}{12})$ $|V'| \geq \frac{11}{30}|V|$ 的连通分量 $c \geq \frac{720}{143}\ln(2) \approx 3.49$ 当 $|V'| \geq 20\ln\left(\frac{10}{\epsilon'}\right) \geq \log\left(\frac{5}{\epsilon'}\right)$。

此时，Alice（分别地，Bob）计算 G 中对应于一条边的所有执行对的按位异或：$(X_i \oplus X_j)_{(i,j)\in E,\ i<j}$（分别地，$(Y_i \oplus Y_j)_{(i,j)\in E,\ i<j}$）。当 ϵ' 足够小时，根据霍夫丁不等式，边集以高概率（$\geq 1 - \frac{\epsilon'}{5}$）小于 $2c \cdot T_{\epsilon'}$（否则参与者可以中止协议）。然后，Alice 和 Bob 通过求解 $\leq 2c \cdot T_{\epsilon'}$ 个相等性实例（总）错误率为 $\leq \frac{\epsilon'}{5}$，找出一个较大的行集合，使这些行异或后得到相同的比特串。结合霍夫丁论证和随机图引理，我们现在得到若干行组，并且知道其中至少有一组代表了 S 的行中超过 $\frac{11}{30}$ 的部分。

现在，对于 S 中每个异或结果为同一比特串、且占 S 所有行中超过 $\frac{11}{30}$ 的行子多重集，任意选取其中一行。如果这样的子多重集只有一个，Alice 和 Bob 可以在此结束协议，并输出所选子多重集中该行的内容。如果有两个这样的子多重集，则令 i_1 和 i_2 分别为从每个子多重集中选出的索引。

步骤 3. 为了从两个候选项中作出选择，Alice 和 Bob 针对所有 $j \in [N]$，在 $X_{i_1} \oplus X_j$ 和 $Y_{i_1} \oplus Y_j$ 之间求解 N 个相等性实例，错误率为 $\leq \frac{\epsilon'}{5}$。如果超过一半的 N 行与 i_1^{th} 行异或后得到相同的字符串，Alice 和 Bob 就输出他们的 i_1^{th} 行；否则，他们输出另一个候选行 i_2。

计算 $\mathbf{GapMAJ}_{N,\ k,\ \epsilon}\circ$**异或** 的复杂度在错误率为 $\epsilon' < \epsilon$ 时满足

$$R_{\epsilon'}^{\mathsf{xor}}\left(\mathbf{GapMAJ}_{N,k,\epsilon}\circ\mathbf{XOR}\right) \leq R_{\epsilon'/5}\left(\mathbf{EQ}_k^{\otimes 2cT_{\epsilon'}}\right) + R_{\epsilon'/5}\left(\mathbf{EQ}_k^N\right).$$

总之，我们应用一个用于相等性（命题 3.5 命题 .）的摊销协议。 □

结合引理 5.5 引理 .和定理 5.6 定理 .即可完成定理 5.3 定理 .的证明。我们将在附录 G 中回到 **GapMAJ**$\circ$**异或** 问题，并在附录 G附录 G 中给出各种模型下的上界（推论 G.2 推论 G.）。

5.2 直和函数的异或错误缩减

定理 5.3 定理 .中的协议首先生成一个完整的 **GapMAJ**$\circ$**异或**实例，然后求解该实例。生成此实例可能会隐式地依赖输出长度 k，在 f非常接近 ϵ 的情况下，这种依赖可能难以接受。我们给出另一种协议，其中参与者无需完整生成这些中间结果。

对于大输出函数，生成输出的一个比特，所需成本可能远低于生成全部 k，例如，当 f 是函数 g的 k 个实例的直和时。我们针对布尔函数的直接和问题陈述了更强的放大定理，但我们指出，该协议也可用于其他问题，在这些问题中，计算输出的一个比特比计算整个输出的成本低。

定理 5.8。 令 $0 < \epsilon' < \epsilon < \frac{1}{2}$和 $C_{\epsilon,\epsilon'} = 8\epsilon\left(\frac{1}{2} - \epsilon\right)^{-2}\ln\left(\frac{12}{\epsilon'}\right)$. 对于任意 $g\colon \mathcal{X} \times \mathcal{Y} \to \{0,1\}$ 和 $f = g^{\otimes k}$,

$$R_{\epsilon'}^{\mathsf{xor}}(f) \leq 50\ln\left(\tfrac{12}{\epsilon'}\right) \cdot R_{\epsilon}^{\mathsf{xor}}(f) + C_{\epsilon,\epsilon'} \cdot R_{\epsilon}^{\mathsf{xor}}(g) + O\left(C_{\epsilon,\epsilon'} + \log(k)\right).$$

注意，$C_{\epsilon,\ \epsilon'}$ 因子（随 $\left(\frac{1}{2} - \epsilon\right)^{-1}$ 变化）适用于 g的复杂度，而非 f的复杂度。

证明思路。 我们不会将基本协议迭代$C_{\epsilon,\ \epsilon'}$ 次，而是先迭代较少次数；这个次数不依赖于 ϵ，只依赖于log($\frac{1}{\epsilon'}$)。这样的迭代次数足以保证出现频率最高的结果占据超过行的 $1/3$比例。如果没有其他结果占据行的较大比例，我们就根据属于这一大比例的某一行输出。否则，仍然至多只有两个结果能各自占据超过行的 $1/3$ 比例。我们找出输出函数的一个“关键索引”，以帮助确定两个候选结果中的多数结果。为此，我们针对该关键索引求解一个间隙汉明距离实例。在剩余的$C_{\epsilon,\ \epsilon'}$ 次运行中，我们只需要输出中的 k 比特之一。

证明的详细内容见附录 E附录 E。

6 确定性复杂度与随机化复杂度

现在我们来讨论如何从私有硬币协议中去随机化。

从私有硬币协议推导出确定性协议的标准方案[3] 如下进行 [KN9797，KN引理 3.8，第 31]。参与者交换消息，以估计每个通信记录出现的概率。他们利用这样一个事实：通信记录的概率可以分解为两部分，分别由两位参与者计算。其中一位参与者将自己的所有因子以一定精度发送给另一位，随后第二位参与者便能估计每个通信记录的概率。每个通信记录都决定一个输出，因此，根据对通信记录概率的估计，该参与者可以推导出每个输出概率的估计，并输出多数答案。

定理 6.1 (KN[KN97], KN第 31 页的引理 3.8)。 对于任意函数 $f\colon\ \mathcal{X}\times\mathcal{Y}\to\mathcal{Z}$ 和 $0<\epsilon<\frac{1}{2}$， 令 $R=R_{\epsilon}^{\mathsf{priv}}(f)$。则 $D(f)\leq 2^{R}\left(R+\log\left(\frac{1}{\frac{1}{2}-\epsilon}\right)+1\right)$.

将这一著名结果用于我们的输出模型（首先向原始协议添加 k 比特的通信，得到一个在单边模型中运行的协议，通信成本为 R ）会使复杂度增加 $2^{R}R\cdot 2^{k}$比特。对于异或模型，我们将这一依赖性降至 $O(2^{R}k)$ 项。在附录 F附录 F中，我们展示了其他模型中对 k 的更低依赖性。

我们将这个问题形式化，并称其为通信记录分布估计。令 $\Delta(\mu,\ \nu)=\frac{1}{2}\sum_{u\in\mathcal{U}}|\mu(u)-\nu(u)|$ 表示全集 $\mathcal{U}$ 上两个概率分布 μ和ν之间的全变差距离。对于协议 Π，令 $\mathcal{T}_{\pi}$表示 Π 的通信记录集合；对于 $(x,\ y)\in\mathcal{X}\times\mathcal{Y}$，我们将运行 Π于 $(x,\ y)$ 时观察到的 $\mathcal{T}_{\pi}$上的分布记为 $T^{x,\ y}_{\pi}$。

定理 66.11 定理 。证明的关键步骤是针对以下问题的一个（在标准模型中的）协议。

定义 6.2 (通信记录分布估计问题)。 对于任意协议 Π 和 $\delta<\frac{1}{2}$，~~~

多数结果变得不可能。我们证明了异或模型中确定性通信的以下界限。

定理 6.4。设 $0 < \epsilon < 1/2$ 和 $f\colon \mathcal{X} \times \mathcal{Y} \to \mathcal{Z} = \{0,1\}^k$。设 $R = R^{\mathsf{xor}}_{,\epsilon}{}^{\mathsf{priv}}(f)$、$M = 16 \cdot \left(\frac{1}{2} - \epsilon\right)^{-2} \cdot 2^R$ 以及 $\epsilon' = \frac{5}{8} - \frac{\epsilon}{4}$。则

$$\begin{aligned} D^{\mathsf{xor}}(f) &\le D(\mathbf{TDE}_{\Pi_f, \epsilon' - \frac{1}{2}}) + D^{\mathsf{xor}}(\mathbf{GapMAJ}_{M,k,\epsilon',\mu} \circ \mathbf{XOR}) \\ &\le \left(2^{R+1}\right) \cdot \left(R + \log\left(\tfrac{8}{\frac{1}{2} - \epsilon}\right) + 1\right) + k \cdot \left(\frac{5 - 2\epsilon}{4} M + 1\right). \end{aligned}$$

其中，μ 是定义在 $[M]$上的未指定分布。

证明思路。我们通过将输出的概率离散化，把求多数结果的问题归约为规模小得多的 **GapMAJ∘异或** 实例。这样一来，对输出规模的依赖就可以减少到仅为 $k = \log(|\mathcal{Z}|)$的一个因子（而不是$2^{2k} = |\mathcal{Z}|^2$的一个因子）。

定理 6*6.4*4定理 。。设 Π 是针对 f的最优私有硬币异或协议。参与者开始运行 $\mathbf{TDE}_{\Pi,\ \delta}$ 引理 F.1 的协议1 引理 F.（引理 66.33 引理 。针对局部模型进行调整，详见附录 F.11 附录 F.），其中 $\delta = \frac{1}{4}\left(\frac{1}{2} - \epsilon\right)$，从而以统计距离 $\bar{}\delta$学习该协议产生的叶节点上的概率分布。

设 $o_{\mathsf{A}}(. \mid w,\ x)$ 和 $o_{\mathsf{B}}(. \mid w,\ y)$ 是两个相互独立的概率分布，分别描述 Alice 和 Bob 在接收到输入x 和 y并到达叶节点 w 后的输出。为了将问题归约为 **GapMAJ∘异或**，他们将 o_{A} 和 o_{B} 离散化为 $\lceil \delta^{-1} \rceil$ 个事件。令 $\dot{o}_{\mathsf{A}}$ 表示 o_{A} 的离散化，它对 Alice 满足以下性质（Bob 的情况类似，令 $\dot{o}_{\mathsf{B}}$）：

$$\forall z, w: \qquad \dot{o}_{\mathsf{A}}(z \mid w, x) \cdot \lceil \delta^{-1} \rceil \in \mathbb{N} \quad \text{and} \quad |o_{\mathsf{A}}(z \mid w, x) - \dot{o}_{\mathsf{A}}(z \mid w, x)| \le \frac{1}{\lceil \delta^{-1} \rceil}.$$

一种简单的离散化贪心方法如下：

1. Replace all $o_{\mathsf{A}}(z \mid w, x)$ by $\dot{o}_{\mathsf{A}}(z \mid w, x) = \frac{1}{\lceil \delta^{-1} \rceil} \left\lfloor \lceil \delta^{-1} \rceil o_{\mathsf{A}}(z \mid w, x) \right\rfloor$.

2. 当 $\dot{o}_{\mathsf{A}}$ 的概率之和小于 1 时，选取一个使 $o_{\mathsf{A}}(z \mid w,\ x) - \dot{o}_{\mathsf{A}}(z \mid w,\ x)$ 最大的 z。对于该 z，令 $\dot{o}_{\mathsf{A}}(z \mid w,\ x) = \frac{1}{\lceil \delta^{-1} \rceil} \left\lceil \lceil \delta^{-1} \rceil o_{\mathsf{A}}(z \mid \overline{w,\ x}) \right\rceil$。

随后，参与者以如下方式构造一个具有 M 行的分布式 **GapMAJ∘异或** 实例：

- 对于每个叶节点 w，参与者定义 $\lceil \delta^{-1} \rceil^2$ 行。行以 $(i,\ j) \in \left[\lceil \delta^{-1} \rceil\right] \times \left[\lceil \delta^{-1} \rceil\right]$ 为索引，并满足：

  – 对于每个 z，恰好有 $\lceil \delta^{-1} \rceil \dot{o}_{\mathsf{A}}(z \mid w,\ x)$ 个索引 $i_z \in \left[\lceil \delta^{-1} \rceil\right]$，使得 Alice 在所有形式为 $(i_z,\ j)$ $\forall j$ 的行上都输出 z。– 对于每个 z，恰好有 $\lceil \delta^{-1} \rceil \dot{o}_{\mathsf{B}}(z \mid w,\ y)$ 个索引 $j_z \in \left[\lceil \delta^{-1} \rceil\right]$，使得 Bob 在所有形式为 $(i,\ j_z)$ $\forall i$ 的行上都输出 z。

- 在分布 i下，与叶节点 j关联的行（w，μ）的概率取为 $p_{\mathsf{If}}(w \mid x,\ y) \cdot \lceil \delta^{-1} \rceil^{-2}$，其中 $p_{\mathsf{If}}(w \mid x,\ y)$ 是原始协议中最终到达叶节点 w的概率 Π。（μ 是定理 6.4 陈述中 $[M]$ 上未指定的分布。）

随后，参与者求解**GapMAJ∘异或** 实例并输出结果。显然，上述过程具有先前所声称的通信复杂度。接下来需要证明，参与者构造出了一个有效的**GapMAJ∘异或** 实例，其结果为 $f(x,\ y)$；也就是说，根据 μ 从这个**GapMAJ∘异或** 实例中随机选取一行，会得到 Alice 和 Bob 两侧的输出 z_{A} 和 z_{B}，使得 $z_{\mathsf{A}} \oplus z_{\mathsf{B}} = f(x,\ y)$ 的概率为 $> \frac{1}{2}$。 _

1. 在原始协议 Π 中，令 $p^{\mathsf{out}}(z \mid x,\ y)$ 表示计算 z （异或之后）的概率，$p^{\mathsf{out}}(z \mid w,\ x,\ y)$ 表示协议以叶节点 w 结束这一条件下的相同概率；对于所有 w，令 $o_{\mathsf{A}}(. \mid w,\ x)$（分别令 $o_{\mathsf{B}}(. \mid w,\ y)$）表示 Alice（分别为 Bob）在叶节点 w 输出一次时所依据的分布。那么 $p^{\mathsf{out}}(z \mid x,\ y)$ 可表示为：

$$\begin{aligned} p^{\mathsf{out}}(z \mid x,y) &= \sum_{w} p^{\mathsf{lf}}(w \mid x,y) \cdot p^{\mathsf{out}}(z \mid w,x,y) \\ &= \sum_{w} p^{\mathsf{lf}}(w \mid x,y) \cdot \sum_{\substack{z_{\mathsf{A}},z_{\mathsf{B}} \\ z_{\mathsf{A}} \oplus z_{\mathsf{B}} = z}} o_{\mathsf{A}}(z_{\mathsf{A}} \mid w,x) \cdot o_{\mathsf{B}}(z_{\mathsf{B}} \mid w,x). \end{aligned}$$

根据协议的正确性，$p^{\mathsf{out}}(f(x,\ y) \mid x,\ y) \geq 1-\epsilon$。

2. 将构造我们的**GapMAJ∘异或**实例时遇到的上述各量的近似值分别记为 $p'^{\mathsf{lf}}(. \mid x,\ y)$、$p'^{\mathsf{out}}(. \mid x,\ y)$、$p'^{\mathsf{out}}(. \mid w,\ x,\ y)$、$\dot{o}_{\mathsf{A}}(. \mid w,\ x)$ 和 $\dot{o}_{\mathsf{B}}(. \mid w,\ y)$。在我们的加权**GapMAJ∘异或**实例中，随机一行对应于给定 z 的概率 $p'^{\mathsf{out}}(z \mid x,\ y)$ 为：

$$p'^{\mathsf{out}}(z \mid x,y) = \sum_{w} p'^{\mathsf{lf}}(w \mid x,y) \cdot \sum_{\substack{z_{\mathsf{A}},z_{\mathsf{B}} \\ z_{\mathsf{A}} \oplus z_{\mathsf{B}} = z}} \dot{o}_{\mathsf{A}}(z_{\mathsf{A}} \mid w,x) \cdot \dot{o}_{\mathsf{B}}(z_{\mathsf{B}} \mid w,x).$$

3. $p'^{\mathsf{lf}}(. \mid x,\ y)$ 与 $p^{\mathsf{lf}}(. \mid x,\ y)$ 的统计距离至多为 δ。$\dot{o}_{\mathsf{A}}(. \mid w,\ x)$ 与 $o_{\mathsf{A}}(. \mid w,\ x)$ 逐点相差至多 δ（$\dot{o}_{\mathsf{B}}$ 和 o_{B} 也类似）。

考虑由 $o_{\mathsf{A}} \cdot o_{\mathsf{B}}(z) = \sum_{z'} o_{\mathsf{A}}(z' \mid w,x) \cdot o_{\mathsf{B}}(z \oplus z' \mid w,y)$ 定义在 $z \in \{0,1\}^k$ 上的分布 $o_{\mathsf{A}} \cdot o_{\mathsf{B}}$。类似地定义 $o_{\mathsf{A}} \cdot \dot{o}_{\mathsf{B}}$ 和 $\dot{o}_{\mathsf{A}} \cdot \dot{o}_{\mathsf{B}}$。由上面的第 33 点可知，$\dot{o}_{\mathsf{A}} \cdot \dot{o}_{\mathsf{B}}$ 与 $o_{\mathsf{A}} \cdot \dot{o}_{\mathsf{B}}$ 逐点相差至多 δ，而 $o_{\mathsf{A}} \cdot \dot{o}_{\mathsf{B}}$ 本身与 $o_{\mathsf{A}} \cdot o_{\mathsf{B}}$ 逐点相差至多 δ。可以验证，$\dot{o}_{\mathsf{A}} \cdot \dot{o}_{\mathsf{B}}$ 与 $o_{\mathsf{A}} \cdot o_{\mathsf{B}}$ 逐点相差至多 2δ。

使用引理 F.22（附录中的引理 F.）以及 $V \sim p^{\mathsf{out}}$、$V' \sim p'^{\mathsf{out}}$、$U \sim p^{\mathsf{lf}}$、$U' \sim p'^{\mathsf{lf}}$、$V_u \sim o_{\mathsf{A}} \cdot o_{\mathsf{B}}$ 和 $V'_u \sim \dot{o}_{\mathsf{A}} \cdot \dot{o}_{\mathsf{B}}$，可得 p 与 p' 逐点相差至多 3δ。由于 δ 被取为 $\frac{1}{4}\left(\frac{1}{2}-\epsilon\right)$，GapMAJ∘异或实例的随机行对应于 $f(x,y)$ 的概率为：$p'^{\mathsf{out}}(f(x,y)) \geq p^{\mathsf{out}}(f(x,y)) - 3\delta \geq (1-\epsilon) - \frac{3}{4}\left(\frac{1}{2}-\epsilon\right) = \frac{1}{2} + \frac{1}{4}\left(\frac{1}{2}-\epsilon\right) > \frac{1}{2}$. □

7 弱输出模型的秩下界

由于新模型的输出要求较弱，标准的下界技术可能不再适用。我们将标准的秩下界推广到所有输出模型中（定理 7.11 定理 .）。虽然这一结果没有证明任何新的下界，但本节的主要贡献在于展示如何将已有下界推广到新的通信复杂度模型中。我们的技术也可以用类似方式应用于其他下界技术。

重新审视单色矩形 令 $f\colon \mathcal{X} \times \mathcal{Y} \to \mathbb{F}$，其中函数的值被视为域 $\mathbb{F}$ 的一个元素。与 f 相关联的通信矩阵是一个矩阵，其行由 $\mathcal{X}$ 的元素索引，列由 $\mathcal{Y}$ 的元素索引，并定义为 $M_f = (f(x,y))_{x \in \mathcal{X}, y \in \mathcal{Y}}$。

在开放模型中，由于存在从叶节点到输出的映射，通信协议会将通信矩阵划分为单色矩形。其他计算模型则并非如此。在单边模型和二选一模型中，叶节点处的矩形会呈现水平或垂直的“条纹”（示意见图 2图 2），因为参与者可以根据自己的输入改变答案。在单边模型中，所有矩形的条纹方向始终相同；而在二选一模型中，条纹方向则可能不同，具体取决于由哪位参与者产生输出。

局部模型更为微妙：两位参与者可以根据各自的局部信息（输入和随机性）决定输出 $\mathbb{F}$ 中的不同元素。只要两位参与者输出的内容不同，结果就是错误的，因此这些矩形看起来类似置换矩阵。

在本节余下部分，我们将使用“叶矩形”一词来指代与协议树叶节点对应的矩形。

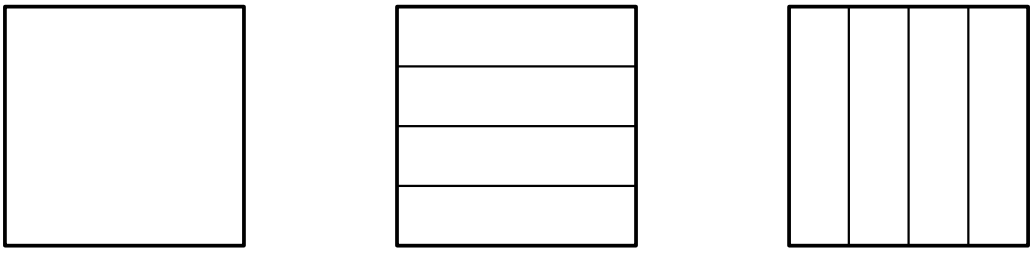


图 2：在开放模型中，协议结束时与叶节点对应的矩形是单色的；而在单向模型和二选一模型中，这些矩形具有单色“条纹”。这里的意思是，可以根据负责输出函数值的参与者所输出的内容，对行进行划分（形成水平条纹）或对列进行划分（形成垂直条纹），从而将矩形进一步划分为单色子矩形。条纹的方向（水平或垂直）取决于哪位参与者输出函数值。

分割模型和异或模型的情况有所不同，因为它们的叶矩形具有更复杂的结构。在异或模型中，异或协议生成的叶矩形与异或函数的通信矩阵相似**异或**$_k$。

秩下界 为了推导我们这些模型的秩下界，我们研究叶矩形的秩。各种模型中叶矩形的秩导出了以下定理。

定理 7.1. 设 f 是一个全函数。那么

$$
\begin{aligned}
D^{\mathsf{open}}(f) = D^{\mathsf{loc}}(f) \geq D^{\mathsf{uni}}(f) \geq D^{\mathsf{1of2}}(f) \geq \log \operatorname{rank}(M_f) \\
D^{\mathsf{spl}}(f) \geq \log \operatorname{rank}(M_f) - 1 \\
D^{\mathsf{xor}}(f) \geq \log \operatorname{rank}(M_f) - \log(k+1)
\end{aligned}
$$

证明。 我们称 M_f 的一个矩形的秩为将 M_f 限制在该矩形上所得到的 M_f 子矩阵的秩。如果存在一种将 M_f 划分为 C 个矩形的划分，使得每个矩形的秩都不超过 R，那么 $\operatorname{rank}(M_f) \leq C \times R$。由于对于每个模型$\mathcal{M}$，$M_f$ 至多被 $2^{D^{\mathcal{M}}(f)}$ 个 $\mathcal{M}$类型的矩形覆盖，因此我们只需为每个模型 $\mathcal{M}$界定 $\mathcal{M}$ 类型矩形的秩。

开放、局部、单向和二选一叶矩形。 这些类型的叶矩形的秩至多为 1，这是因为它们具有条纹结构。另请注意，对于确定性情形下的全函数，开放叶矩形和局部叶矩形是相似的。

分割叶矩形。这种类型的叶矩形的秩至多为 2。直观地说，这是因为该模型中的叶矩形具有以下形式：存在若干数a_1，……，a_s和 b_1，……，b_t，使得大小为 $s \times t$的矩形中单元格（i，j）的值为$a_i + b_j$。因此，该矩形是以下两个秩为 2 的矩阵的乘积：第一个 $s \times 2$ 矩阵的第一列包含从 a_1 到 a_s 的值，第二列的所有单元格均为 1；第二个 $2 \times t$ 矩阵的第一行仅包含值 1，第二行包含从 b_1 到 b_t 的值，如图 3 所示3 图。

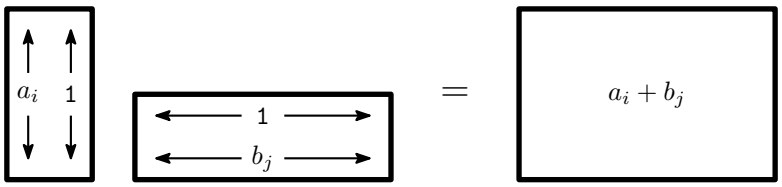


图 3：一个矩阵，其单元格 $M_{i,\ j}$ 可表示为第一个向量的第 i 个分量与另一个向量的第 j 个分量之和，因此其秩至多为 2。分割矩形遵循这一模式。

更正式地说：考虑输出的 k 比特如何在两个参与者之间分配：我们来看比特串 $(s_i)_{1 \le i \le k}$，当且仅当 Alice 输出该输出的第 i^{th} 比特时，满足 $s_i = 1$。

现在定义矩阵 1×1 $S_0 = \begin{bmatrix} 0 \end{bmatrix}$，以及由下式定义的矩阵变换 H_c 和 V_c：

- $H_c(A) = \begin{bmatrix} A & A + c \cdot J \end{bmatrix}$
- $V_c(A) = \begin{bmatrix} A \\ A + c \cdot J \end{bmatrix}$

现在定义三个矩阵序列 $S_1 \dots S_k$、$U_0 \dots U_k$ 和 $V_0 \dots V_k$，使得对所有 i 都有 $S_i = U_i \times V_i$，由此可证明 S_k 的秩至多为 2：

- Let $S_{i+1} = \begin{cases} H_{2^i}(S_i) & \text{if } s_i = 0 \\ V_{2^i}(S_i) & \text{if } s_i = 1 \end{cases}$.
- Let $U_0 = \begin{bmatrix} 0 & 1 \end{bmatrix}$ and $V_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.
- Let $U_{i+1} = \begin{cases} U_i & \text{if } s_i = 0 \\ \begin{bmatrix} & U_i & \\ U_i + \begin{bmatrix} 2^i & 0 \\ \vdots & \vdots \\ 2^i & 0 \end{bmatrix} \end{bmatrix} & \text{if } s_i = 1 \end{cases}$
- Let $V_{i+1} = \begin{cases} \begin{bmatrix} V_i & V_i + \begin{bmatrix} 0 & \dots & 0 \\ 2^i & \dots & 2^i \end{bmatrix} \end{bmatrix} & \text{if } s_i = 0 \\ V_i & \text{if } s_i = 1 \end{cases}$

要证明性质 $S_i = U_i \times V_i$ 对所有 $i \in [k]$ 都成立，请注意，U_i 的第二列和 V_i 的首行只包含 1，因为这对 $i = 0$ 成立，而且随着 i 增大，该性质会保持不变。在 V_i 第二行的后半部分加上一个常数 c 时，这个常数会乘以 U_i 的第二列，而该列只包含 1。最终，我们在矩阵的一半中加上一个 $c \cdot J$ 矩阵，这正是我们想要的。

最后，注意到 S_k 是一个矩阵，其中包含在给定特定分割的情况下，Alice 和 Bob 在分割模型中能够输出的所有内容。分割模型中的叶矩形是这种形式的矩阵的一个子矩阵，其中某些行和列可能经过置换或重复。因此，分割模型中的叶矩形的秩至多为 2。

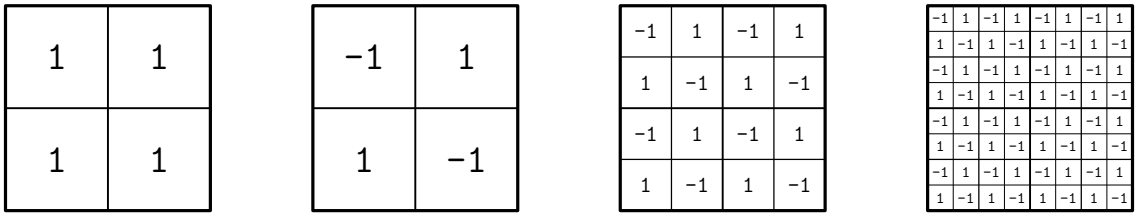
XOR 叶矩形 我们证明，由异或协议产生的叶矩形的秩至多为 $(k + 1)$。

考虑该**异或**$_k$ 函数的通信矩阵。异或叶矩形可以通过对该通信矩阵的某个子矩阵进行行列置换或复制得到。因此，只需证明 $M_{\mathbf{XOR}_k}$ 的秩为 $k + 1$。我们直接给出 $M_{\mathbf{XOR}_k}$ 的一个秩为 $k + 1$ 的分解来证明这一点。考虑以下 $2^k \times 1$ 向量：

- v^k 是全一向量。
- 对于 $0 \le i < k$, $u^{k,\ i}$ 满足 $u^{k,\ i}_{\ j} = (-1)^{1+j_i}$（对于 $0 \le j < 2^k$）。这样的向量有时称为哈达玛向量。

令 S_k 为以下 $2^k \times (k+1)$ 矩阵：

$$S_k = \left[\sqrt{2^{k-1} - 2^{-1}} \cdot v \quad \sqrt{2^{-1}} \cdot u^{k,0} \quad \dots \quad \sqrt{2^{k-2}} \cdot u^{k,k-1}\right]$$



(a) (b) (c) (d)

图 4： $M_{\mathbf{XOR}_3}$ 通信矩阵可以由这些矩阵的线性组合得到。

我们有 $S_k\,{}^tS_k = M_{\mathbf{XOR}_k}$。图 44 图 直观展示了如何得到 $M_{\mathbf{XOR}_k}$ 矩阵。

□

8 结论与开放问题

我们介绍了专为非布尔函数设计的输出模型。我们希望这些模型能有广泛的应用，包括扩展到信息复杂度、更好地理解直接和问题、模拟协议，以及为这些模型量身定制的新下界，等等。

间隙多数与异或复合问题（定义 55.44定义 .）与间隙汉明距离密切相关，后者扩展到了大字母表，但增加了一项承诺，因此针对**GHD** 的下界并不适用。我们猜想，它的确定性通信复杂度为 $\Omega(\epsilon Nk)$，与平凡上界相匹配。若此猜想成立，则表明我们的去随机化方案（定理 66.44定理 .）已接近最优。

9 致谢

我们感谢 J´er´emie Roland 和 Sagnik Mukhopadhyay 与我们进行了有益的讨论，也感谢匿名审稿人提出许多建议，帮助改进论文的表述。本研究部分由法国国家研究署资助，项目为 FLITTLA ANR-21-CE48-0023。

参考文献

[Aar05] Scott Aaronson。一致性的复杂度。载于第三十七届 ACM 计算理论研讨会论文集，第 634–643 页。ACM，2005 年。[ABG+01] Andris Ambainis 、 Harry Buhrman 、 William Gasarch、Bala Kalyanasundaram 和 Leen Torenvliet。枚举、消除与选择的通信复杂度。《计算机与系统科学杂志》，63(2):148–185，2001年。[BBCR13] Boaz Barak、Mark Braverman、Xi Chen 和 Anup Rao。如何压缩交互式通信。《SIAM计算杂志》，42(3):1327—1363，2013年。[BBK+16] Joshua Brody、Harry Buhrman、Michal Kouck´y、Bruno Loff、Florian Speelman 和 Nikolai K. Vereshchagin。迈向交互式信息复杂度中的反向 Newman 定理。《算法学》，76(3):749–781，2016年。[BCK+14] Joshua Brody、Amit Chakrabarti、Ranganath Kondapally、David P. Woodruff 和 Grigory Yaroslavtsev。超越集合不相交性：寻找交集的通信复杂度。载于 ACM 分布式计算原理研讨会论文集，PODC ’14，第 106–113 页，2014年。[BCK+16] Joshua Brody、Amit Chakrabarti、Ranganath Kondapally、David P. Woodruff 和 Grigory Yaroslavtsev。通过有限交互验证相等性。《算法学》，76(3):796–845，2016年。[BCMdW10] Harry Buhrman、Richard Cleve、Serge Massar 和 Ronald de Wolf。非定域性与通信复杂度。《现代物理评论》，82:665–698，2010年3月。[BCW98] Harry Buhrman、Richard Cleve 和 Avi Wigderson。量子与经典的通信和计算。载于第三十届 ACM 计算理论研讨会论文集，第 63–68 页。ACM，1998年。[BDKW14] Amos Beimel、Sebastian Ben Daniel、Eyal Kushilevitz 和 Enav Weinreb。通信复杂度中的选择、一致与消除。《计算复杂性》，23(1):1–42，2014年。[Bel64] J. S. Bell。论爱因斯坦–波多尔斯基–罗森悖论。《物理学》，1:195，1964年。[BK18] Mark Braverman 和 Gillat Kol。将交互式通信压缩至外部信息。载于第50届 ACM SIGACT 计算理论研讨会论文集，STOC，第 964–977 页，2018年。[BMY15] Balthazar Bauer、Shay Moran 和 Amir Yehudayoff。将协议内部压缩至熵。载于《近似、随机化与组合优化：算法与技术》，APPROX/RANDOM 2015，第 481–496 页，2015年。[Bol01] B´ela Bollob´as。《随机图》。剑桥高等数学研究丛书。剑桥大学出版社，第二版，2001年。[BR14] Mark Braverman 和 Anup Rao。信息等于摊销通信。《IEEE信息理论汇刊》，60(10):6058–6069，2014年。[Bra15] Mark Braverman。交互式信息复杂度。《SIAM计算杂志》，44(6):1698–1739，2015年。

[BRWY13] Mark Braverman、Anup Rao、Omri Weinstein 和 Amir Yehudayoff。通信复杂度中的直积定理。载于第54届 IEEE 计算机科学基础年会（FOCS 2013），2013年10月26日至29日，美国加利福尼亚州伯克利，第746–755页，2013年。[BYJKS04] Ziv Bar-Yossef、T. S. Jayram、Ravi Kumar 和 D. Sivakumar。数据流与通信复杂度的信息统计方法。《计算机系统科学杂志》，68(4):702–732，2004年。[CR12] Amit Chakrabarti 和 Oded Regev。间隔汉明距离通信复杂度的最优下界。《SIAM计算杂志》，41(5):1299–1317，2012年。[ER60] P´al Erd˝os 和 Alfr´ed R´enyi。关于随机图的演化。载于匈牙利科学院数学研究所出版物，第17–61页，1960年。[FJK+16] Lila Fontes、Rahul Jain、Iordanis Kerenidis、Sophie Laplante、Mathieu Lauri`ere 和 J´er´emie Roland。相对差异不能区分信息复杂度与通信复杂度。《ACM计算理论汇刊》，9(1):4:1–4:15，2016年。[FKNN95] Tomas Feder、Eyal Kushilevitz、Moni Naor 和 Noam Nisan。摊销通信复杂度。《SIAM计算杂志》，24(4):736–750，1995年8月。[FRPU94] Uriel Feige、Prabhakar Raghavan、David Peleg 和 Eli Upfal。利用噪声信息进行计算。《SIAM计算杂志》，23(5):1001–1018，1994年10月。[GKR16] Anat Ganor、Gillat Kol 和 Ran Raz。布尔函数的信息与通信复杂度之间的指数分离。《计算复杂性》，63(5):46:1–46:31，2016年。[GKR21] Anat Ganor、Gillat Kol 和 Ran Raz。通信复杂度与外部信息之间的指数分离。《SIAM计算杂志》，50(3)，2021年。[HPZZ21] Dawei Huang、Seth Pettie、Yixiang Zhang 和 Zhijun Zhang。集合交集与多重相等性测试的通信复杂度。《SIAM计算杂志》，50(2):674–717，2021年。[HW07] Johan H˚astad 和 Avi Wigderson。集合不相交性的随机化通信复杂度。计算理论，3(11):211–219，2007年。[IW03] Piotr Indyk 和 David P. Woodruff。不同元素问题的紧下界。载于第44届计算机科学基础研讨会论文集（FOCS 2003），第283–288页，2003年。[JK10] Rahul Jain 和 Hartmut Klauck。经典通信复杂度与查询复杂度的划分界。载于第25届 IEEE 计算复杂性年会论文集（CCC 2010），第247–258页，2010年。[KN97] Eyal Kushilevitz 和 Noam Nisan。通信复杂度。剑桥大学出版社，美国纽约州纽约市，1997年。[Kol16] Gillat Kol。乘积分布下的交互式压缩。载于第48届 ACM SIGACT 理论计算年会（STOC）论文集，第987–998页。ACM，2016年。[Koz15] Alexander Kozachinskiy。关于间隔汉明距离通信复杂度的一些界。CoRR，abs/1511.08854，2015年。

[KS92] Bala Kalyanasundaram 和 Georg Schnitger。集合交集的概率通信复杂度。SIAM 离散数学杂志，5(4):545–557，1992年。[MS83] Florence Jessie MacWilliams 和 Neil James Alexander Sloane。纠错码理论，北荷兰数学丛书第16卷。爱思唯尔，1983年。[Nis93] Noam Nisan。阈值门的通信复杂度。载于《组合学：Paul Erd˝os八十华诞纪念文集》，波尧伊学会数学研究丛书第1卷，第301–315页。J´anos Bolyai Mathematical Society，1993年。[Orl90] Alon Orlitsky。最坏情形交互通信 I：两条消息几乎最优。IEEE 信息论汇刊，36(5):1111–1126，1990年。[Orl91] Alon Orlitsky。最坏情形交互通信 II：两条消息并非最优。IEEE 信息论汇刊，37(4):995–1005，1991年。[PV16] Carlos Palazuelos 和 Thomas Vidick。非局域博弈与算子空间理论综述。数学物理学杂志，57(1):015220，2016年。[Raz92] Alexander A. Razborov。集合不相交性的分布复杂度。理论计算机科学，106(2):385–390，1992年。[RS18] Anup Rao 和 Makrand Sinha。信息与通信分离的简化证明。计算理论，14(1):1–29，2018年。[RY20] Anup Rao 和 Amir Yehudayoff。通信复杂度。剑桥大学出版社，2020年。[She12] Alexander A. Sherstov。间隙汉明距离的通信复杂度。计算理论，8(1):197–208，2012年。[She18] Alexander Sherstov。乘积分布下的交互式通信压缩。《SIAM计算杂志》，47(2):367–419，2018年。[Vid12] Thomas Vidick。向量在大集合上的重叠度集中不等式及其在间隙汉明距离问题通信复杂度中的应用。芝加哥理论计算机科学杂志，2012，2012年。[Vio15] Emanuele Viola。加法的通信复杂度。组合学，35(6):703–747，2015年12月。[Wat18] Thomas Watson。小优势下的通信复杂度。载于第33届计算复杂性会议（CCC）论文集，第9:1–9:17页，2018年。

A 大输出函数的模型

通信复杂度的一种标准定义要求：通信协议结束时，可以根据通信记录和公共随机性确定计算结果（矩形界采用的正是这种模型）。很容易找到这样的例子：按照这种定义，所需交换的通信量远超直觉上的合理值。例如，

例 A.1。考虑函数 $f\colon \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^n, f(x,y) = x$，并假设我们希望在承诺 $x = y$ 下计算它。

如果协议的结果必须能够从通信内容和公共随机性中看出，那么计算 f 需要 n 比特的通信，尽管双方参与者从一开始就知道$f(x,y)$。

本节将正式定义输出模型，并证明分离结果。最有趣的模型或许是最弱的几种：二选一（定义 A.99定义 A.）、分割（定义 A.1313定义 A.）和异或模型（定义 44.11定义 .）。

A.1 开放模型

我们从模型的形式化定义开始；该模型揭示了关于计算结果的最多信息。我们称之为开放模型。

这是这样一种模型，其中划分界[JK10], JK以文献中出现的形式给出下界。

定义 A.2 (开放计算)。若存在一个映射 $\mathcal{O}\colon \mathcal{T}_\pi \times \mathcal{R}^{\mathsf{pub}} \to \mathcal{Z}$，使得对所有 $(x,y) \in \mathcal{X} \times \mathcal{Y}$，则称协议 Π 在错误为 ϵ 时能够开放地计算 f：

$$\Pr_{r,r_A,r_B}[\mathcal{O}(t_\pi, r) = f(x,y)] \geq 1 - \epsilon.$$

A.2 局部模型

在前一个模型中，协议是揭示性的，也就是说，计算结果不能是只有参与者知晓的秘密。在局部模型中，我们只要求双方参与者在协议结束时都能输出函数的值（对于关系，则输出相同的有效结果）。

定义 A.3（局部计算）。如果存在两个映射 $\mathcal{O}_A$ 和 $\mathcal{O}_B$，满足 $\mathcal{O}_A$：$\mathcal{T}_\pi \times \mathcal{R}^{\mathsf{pub}} \times \mathcal{R}_A \times \mathcal{X} \to \mathcal{Z}$，并且类似地$\mathcal{O}_B$：$\mathcal{T}_\pi \times \mathcal{R}^{\mathsf{pub}} \times \mathcal{R}_B \times \mathcal{Y} \to \mathcal{Z}$，使得对于所有 $(x,y) \in \mathcal{X} \times \mathcal{Y}$，协议 Π 可以以 ϵ 错误局部计算 f，则称该协议具有此性质：

$$\Pr_{r,r_A,r_B}[\mathcal{O}_A(t_\pi, r, r_A, x) = \mathcal{O}_B(t_\pi, r, r_B, y) = f(x,y)] \geq 1 - \epsilon.$$

Bauer 等人 [BMY15] BMY指出，对于全函数和关系，确定性开放模型与局部模型的通信复杂度相同。例 A.11 例 A.展示了在承诺条件下计算函数时，确定性复杂度之间存在差异。

对于随机通信，以下全函数使局部模型与开放模型之间产生了分离，见定理 A.55 定理 A.：

定义 A.4（带输出的相等性问题）。$\mathbf{EQ}_n^{\mathsf{out}}$：$\{0,1\}^n \times \{0,1\}^n \to \{0,1\}^n \cup \{\top\}$ 定义为

$$\mathbf{EQ}_n^{\mathsf{out}}(x,y) = \begin{cases} x & if\ x = y \\ \top & otherwise \end{cases}$$

0	⊤	⊤	⊤	⊤	⊤	⊤	⊤
⊤	1	⊤	⊤	⊤	⊤	⊤	⊤
⊤	⊤	2	⊤	⊤	⊤	⊤	⊤
⊤	⊤	⊤	3	⊤	⊤	⊤	⊤
⊤	⊤	⊤	⊤	4	⊤	⊤	⊤
⊤	⊤	⊤	⊤	⊤	5	⊤	⊤
⊤	⊤	⊤	⊤	⊤	⊤	6	⊤
⊤	⊤	⊤	⊤	⊤	⊤	⊤	7

图 5：$\mathbf{EQ}_3^{\mathsf{out}}$的通信矩阵

定理 A.5。 $\forall f\colon \mathcal{X}\times\mathcal{Y}\to\mathcal{Z}$ 满足 $k=\lceil$对数$|\mathcal{Z}|\rceil$ 以及 $\epsilon>0$,

$$R_\epsilon^{\mathsf{loc}}(f)\le R_\epsilon^{\mathsf{open}}(f)\le R_\epsilon^{\mathsf{loc}}(f)+k,\qquad and$$
$$R_{1/4}^{\mathsf{loc}}(\mathbf{EQ}_n^{\mathsf{out}})\le 4,\qquad R_{1/4}^{\mathsf{open}}(\mathbf{EQ}_n^{\mathsf{out}})\in\Omega(n).$$

我们给出该定理的完整证明，但由于对于两个模型 $\mathcal{M}_1$ 和 $\mathcal{M}_2$，形如 $R_\epsilon^{\mathcal{M}_1}(f)\le R_\epsilon^{\mathcal{M}_2}(f)$ 或 $R_\epsilon^{\mathcal{M}_1}(f)\le R_\epsilon^{\mathcal{M}_2}(f)+k$ 的所有结果都可以用本质上相同的证明来证明，因此在后续类似定理的证明中，我们将省略这些证明，只证明分离结果。

定理 *A.5* 的证明*5*定理 *A.*.**证明** $R_\epsilon^{\mathsf{loc}}(f)\le R_\epsilon^{\mathsf{open}}(f)$**：** 对于函数 f，一个开放协议也是一个用于计算 f 的局部协议，因为参与者可以将开放协议的映射 $\mathcal{O}_\mathsf{A}$ 和 $\mathcal{O}_\mathsf{B}$ 用作映射 $\mathcal{O}$ （忽略双方参与者的随机性和输入）。

证明 $R_\epsilon^{\mathsf{open}}(f)\le R_\epsilon^{\mathsf{loc}}(f)+k$**：** 设 Π 是一个用于计算 f 的局部协议，其错误率至多为 ϵ。考虑由先运行协议 Π，然后 Alice 通过通信信道发送 $\mathcal{O}_\mathsf{A}(t_\pi,\ r,\ r_\mathsf{A},\ x)$（即她在结束 Π 时本会输出的内容）以局部计算 f 的协议 Π'。这只需要额外 k 比特的通信。现在 Π' 是一个开放协议，因为外部观察者可以将通信记录的最后 k 比特作为可能的 $f(x,\ y)$。

$\mathbf{EQ}^{\mathsf{out}}$ 的下界和上界均可直接由本手稿中前面给出的命题和定理推出。

局部模型上界： 参与者采用 **EQ** 的标准协议（命题 3*3.5*5命题 .）。如果两个字符串不同，他们输出 $\top$；否则 Alice 输出 x，Bob 输出 y。

开放模型下界： 考虑开放协议 Π 的映射 $\mathcal{O}$，并注意，对于所有 x，$\Pr_r[\mathcal{O}(\Pi(x,\ x,\ r)\ r)=x]\ge 3/4$。设参与者拥有一个公共随机源 $\mathcal{R}^{\mathsf{pub}}$，其对应于 $\{0,1\}^k$ 上的均匀随机分布。那么上述陈述意味着 $|\mathcal{O}^{-1}(x)|\ge\frac{3}{4}\cdot 2^k$。由于 $\cup_x\mathcal{O}^{-1}(x)\subseteq\mathcal{T}_\pi\times\{0,1\}^k$，我们有 $\frac{3}{4}\cdot 2^k\cdot 2^n\le 2^{\mathrm{CC}(\Pi)}\cdot 2^k$，因此 $\mathrm{CC}(\Pi)\ge n+\log\binom{3}{4}\in\Omega(n)$。当公共随机源不是 $\{0,1\}^k$ 上的均匀分布时，这一结论同样成立，因为任何非均匀随机源都可以通过均匀随机源以任意精度模拟。

□

在附录 C 中附录 C 我们将其推广，证明对于一个问题，任何开放协议都需要 $\Omega(k)$ 通信。这个结果源于一个称为弱划分界的下界 [FJK+16]FJK。

A.3 单边模型

本节中，我们考察通信复杂度模型，在这些模型中，我们要求协议结束时，一位参与者能够输出函数的值（对于关系，则输出一个有效结果）。单向问题通常就是在这一模型中表述的。

定义 A.6 (单边计算)。 如果存在映射 $\mathcal{O}_{\mathsf{A}}\colon \mathcal{T}_\pi \times \mathcal{R}^{\mathsf{pub}} \times \mathcal{R}_{\mathsf{A}} \times \mathcal{X} \to \mathcal{Z}$，使得对所有$(x, y) \in \mathcal{X} \times \mathcal{Y}$，则称协议 Π 以 ϵ 错误计算 f，由Alice执行：

$$\Pr_{r,r_{\mathsf{A}},r_{\mathsf{B}}}[\mathcal{O}_{\mathsf{A}}(t_\pi, r, r_{\mathsf{A}}, x) = f(x, y)] \geq 1 - \epsilon.$$

Bob 计算的定义与此类似。
如果协议单向计算 f，即由 *Alice* 或 *Bob* 计算 f，则称该协议为单向 计算。

我们对单边模型的定义相当于两个模型的最小值，每个模型分配给一个参与者。

定义 A.7 (单边恒等问题)。 $\mathbf{id}_n^{\mathsf{A}}\colon \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^n$ 定义为

$$\mathbf{id}_n^{\mathsf{A}}(x, y) = x$$

id_n^{B} 的定义类似，只是 *Alice* 和 *Bob* 的角色相反。

0	0	0	0	0	0	0	0
1	1	1	1	1	1	1	1
2	2	2	2	2	2	2	2
3	3	3	3	3	3	3	3
4	4	4	4	4	4	4	4
5	5	5	5	5	5	5	5
6	6	6	6	6	6	6	6
7	7	7	7	7	7	7	7

0	1	2	3	4	5	6	7
0	1	2	3	4	5	6	7
0	1	2	3	4	5	6	7
0	1	2	3	4	5	6	7
0	1	2	3	4	5	6	7
0	1	2	3	4	5	6	7
0	1	2	3	4	5	6	7
0	1	2	3	4	5	6	7

图 6：$\mathbf{id}_3^{\mathsf{A}}$ 和$\mathbf{id}_3^{\mathsf{B}}$的通信矩阵

定理 A.8。 $\forall f\colon \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$ 其中 $k = \lceil \log|\mathcal{Z}| \rceil$，$\lambda \in [0,1]$ 以及 $\epsilon > 0$

$$R_\epsilon^{\mathsf{uni}}(f) \leq R_\epsilon^{\mathsf{loc}}(f) \leq R_\epsilon^{\mathsf{open}}(f) \leq R_\epsilon^{\mathsf{uni}}(f) + k,$$
$$D^{\mathsf{loc}}(f) \leq D^{\mathsf{A}}(f) + D^{\mathsf{B}}(f), \qquad R_\epsilon^{\mathsf{loc}}(f) \leq R_{\lambda\epsilon}^{\mathsf{A}}(f) + R_{(1-\lambda)\epsilon}^{\mathsf{B}}(f), \quad \textit{and}$$
$$D^{\mathsf{uni}}(\mathbf{id}_n^{\mathsf{A}}) = D^{\mathsf{A}}(\mathbf{id}_n^{\mathsf{A}}) = D^{\mathsf{B}}(\mathbf{id}_n^{\mathsf{B}}) = 0, \qquad R_{1/4}^{\mathsf{loc}}(\mathbf{id}_n^{\mathsf{A}}) = R_{1/4}^{\mathsf{loc}}(\mathbf{id}_n^{\mathsf{B}}) \in \Omega(n).$$

第一行对关系也成立，但第二行不成立：考虑关系 f 作为反例：
$\{0,1\}^n \times \{0,1\}^n \to 2^{\{0,1\}^n}$，$f(x, y) = \{x, y\}$。在两个单边模型中（$D^{\mathsf{A}}(f) = D^{\mathsf{B}}(f) = 0$），这个问题都不需要任何通信，但在局部模型中，由于参与者需要就单一输出达成一致，因此无论在确定性情形还是随机化情形下，通信量都为 $\Omega(n)$（$D^{\mathsf{loc}}(f) \geq R_\epsilon^{\mathsf{loc}}(f) \in \Omega(n)$）。

定理 *A.8* 的证明*8 定理 A.*。 我们省略前两行的证明，因为它们仅基于使用相同的协议但采用不同的适当映射，或者通过通信信道发送较低层模型的输出。

我们证明一个稍强的分离结果： $R_{1/4}^{\mathsf{B}}$**(标识**$_n^{\mathsf{A}}$**)** $\in \Omega(n)$。

A Alice 模型上界： Alice 输出她的 x，这不需要任何通信 .

Bob 模型下界： 考虑 $D^{\mathsf{B}}_{1/4}(\mathrm{id}^{\mathsf{A}}_n, \mu)$，其中 μ 为均匀分布。Bob 必须输出 2^n 个等概率答案中的一个。通信量为 C 时，Bob 只能给出 2^C 个不同的答案，因此 Bob 出错的概率为 $\geq 1 - 2^{C-n}$。由于要求 Bob 的错误概率小于 $1/4$，我们有：$C \geq n + \log\binom{3}{4}$，因此 $R^{\mathsf{B}}_{1/4}(\mathrm{id}^{\mathsf{A}}_n) \in \Omega(n)$。

□

A.4 二选一模型

在单边模型中，协议结束时输出结果的参与者是固定的。具体而言，这不取决于输入。在二选一模型中，我们放宽了这一条件：在二选一模型中正确计算一个函数，对应于这样一种执行：协议结束时：

- 一名参与者输出一个特殊符号 $\top \notin \mathcal{Z}$（表示沉默）
- 另一名参与者输出 $f(x, y)$。

直观地说，我们不仅要求其中一名参与者输出正确答案，还要求她知道自己的答案很可能是正确的，同时另一名参与者也知道对方有一个可以输出的好答案。如果我们只要求一名参与者给出正确答案，那么所有布尔函数都可以在该模型中通过零通信来求解。相比之下，我们的模型不会使布尔函数的通信复杂度变得平凡。

定义 A.9 (二选一计算)。如果存在两个映射 $\mathcal{O}_{\mathsf{A}}$ 和 $\mathcal{O}_{\mathsf{B}}$，满足 $\mathcal{O}_{\mathsf{A}}: \mathcal{T}_\pi \times \mathcal{R}^{\mathsf{pub}} \times \mathcal{R}_{\mathsf{A}} \times \mathcal{X} \to \mathcal{Z} \cup \{\top\}$，且类似地 $\mathcal{O}_{\mathsf{B}}: \mathcal{T}_\pi \times \mathcal{R}^{\mathsf{pub}} \times \mathcal{R}_{\mathsf{B}} \times \mathcal{Y} \to \mathcal{Z} \cup \{\top\}$，使得对于所有 $(x, y) \in \mathcal{X} \times \mathcal{Y}$，则称协议 Π 以 ϵ 错误进行二选一计算 f：

$$\Pr_{r, r_{\mathsf{A}}, r_{\mathsf{B}}}\left[(\mathcal{O}_{\mathsf{A}}(t_\pi, r, r_{\mathsf{A}}, x), \mathcal{O}_{\mathsf{B}}(t_\pi, r, r_{\mathsf{B}}, y)) \in \{(f(x,y), \top), (\top, f(x,y))\}\right] \geq 1 - \epsilon.$$

下一个命题表明，任何二选一协议都可以转换为另一个错误不高于原协议、且仅多使用一比特通信成本的二选一协议，使得协议结束时总是恰有一个参与者输出一个值 $\mathcal{Z}$，另一个保持沉默（输出 $\top$）。

命题 A.10。考虑函数 $f: \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$，以及一个用于计算 f 的二选一协议 Π，其错误为 $\epsilon > 0$，通信成本为 C。那么存在一个通信成本为 $(C+1)$ 的二选一协议 Π'，以相同的错误计算 f，且其映射满足协议结束时总是只有一个参与者发言：

$$\forall x, y, r_{\mathsf{A}}, r_{\mathsf{B}}, r, t_{\pi'} = \Pi'(x, y, r_{\mathsf{A}}, r_{\mathsf{B}}, r):$$
$$\left(\mathcal{O}'_{\mathsf{A}}(t_{\pi'}, r, r_{\mathsf{A}}, x), \mathcal{O}'_{\mathsf{B}}(t_{\pi'}, r, r_{\mathsf{B}}, y)\right) \in (\mathcal{Z} \times \{\top\}) \cup (\{\top\} \times \mathcal{Z}).$$

命题 *A.10* 的证明命题 *A.10*。设 Π 是针对 f 和 $\mathcal{O}_{\mathsf{A}}$的$\mathcal{O}_{\mathsf{B}}$ 二选一协议，Π' 为其相关映射。我们将协议 Π 定义为一种协议：它首先按照 t_π的方式运行（获取一份通信记录）；当我们在针对 Π的协议中到达一个叶节点时，Alice 按照以下规则向 Bob 发送一个比特的通信内容：

- 如果 $\mathcal{O}_{\mathsf{A}}(t_\pi, r, r_{\mathsf{A}} x) = \top$，Alice 向 Bob 发送 0。
- 否则，Alice 向 Bob 发送 1。

令 c_{A} 为这个控制位，由 Alice 在新协议 Π'的最后一轮发送。接着，Alice 保持映射 $\mathcal{O}_{\mathsf{A}}$ 不变，而 Bob 的新映射 $\mathcal{O}'_{\mathsf{B}}$ 满足：

$$\mathcal{O}'_{\mathsf{B}}(t_{\pi'}, r, r_{\mathsf{B}}, y) = \begin{cases} \top & \text{if } c_a = 1, \\ \mathcal{O}_{\mathsf{B}}(t_\pi, r, r_{\mathsf{B}}, y) & \text{if } c_a = 0 \text{ and } \mathcal{O}_{\mathsf{B}}(t_\pi, r, r_{\mathsf{B}}, y) \neq \top, \\ z & \text{picked u.a.r. in } \mathcal{Z}, \text{ otherwise.} \end{cases}$$

直观地说，Alice 告诉 Bob 是否发言，而他会照做。只有在两人原本都打算发言或都打算保持沉默的情况下，这种做法才会改变参与者的输出，因此这一过程不会增加错误。

□

定义 A.11 (条件恒等问题)。 函数 $\mathbf{CondId}_n$： $\{0,1\}^n \times \{0,1\}^n \to \{0,1\}^n$ 定义如下

$$\mathbf{CondId}_n(x, y) = \begin{cases} x & \text{if } x_0 = y_0, \\ y & \text{otherwise,} \end{cases}$$

其中 x_0 是 x的第一个比特， y也类似。

0	0	0	0	4	5	6	7
1	1	1	1	4	5	6	7
2	2	2	2	4	5	6	7
3	3	3	3	4	5	6	7
0	1	2	3	4	4	4	4
0	1	2	3	5	5	5	5
0	1	2	3	6	6	6	6
0	1	2	3	7	7	7	7

图 7：$\mathbf{CondId}_3$ 的通信矩阵

定理 A.12。 $\forall f$： $\mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$ 满足 $k = \lceil$对数$|\mathcal{Z}| \rceil$ 和 $\epsilon > 0$

$$R_\epsilon^{\mathsf{1of2}}(f) \leq R_\epsilon^{\mathsf{uni}}(f) \leq R_\epsilon^{\mathsf{loc}}(f) \leq R_\epsilon^{\mathsf{open}}(f) \leq R_\epsilon^{\mathsf{1of2}}(f) + k + 1, \quad \textit{and}$$
$$D^{\mathsf{1of2}}(\mathbf{CondId}_n) \in O(1), \qquad R_\epsilon^{\mathsf{uni}}(\mathbf{CondId}_n) \in \Omega(n).$$

定理 *A.12* 的证明*12* 定理 *A.*。 同样，我们关注分离结果。

二选一模型上界： Alice 和 Bob 互相发送 x_0 和 y_0。如果 $x_0 = y_0$，则 Alice 输出 x，否则 Bob 输出 y。这只需要 2 比特通信。

单边模型下界： 考虑 $D^{\mathsf{B}}_{1/4}(\mathbf{CondId}_n, \mu)$，其中 μ 是 (x, y) 上的均匀分布，且 $x_0 = y_0$。收到任意给定的 x 后，Bob 必须输出 2^{n-1} 个等概率答案之一。在通信量为 C 时，Bob 只能给出 2^C 个不同答案，因此 Bob 出错的概率为 $\geq 1 - 2^{C-n+1}$。由于 Bob 的错误率应低于 1_4，我们有： $C \geq n - 1 + \log\binom{3}{4}$，因此 $R^{\mathsf{B}}_{1/4}(\mathbf{CondId}_n) \in \Omega(n)$。由对称性，我们也有 $R^{\mathsf{A}}_{1/4}(\mathbf{CondId}_n) \in \Omega(n)$，因此 $R^{\mathsf{uni}}_{1/4}(\mathbf{CondId}_n) \in \Omega(n)$。

□

A.5 分割模型

在下一个模型中，我们允许两个参与者分摊答案。在一人输出模型中，其中一位参与者必须输出完整结果，而另一位则完全保持沉默。相比之下，在分割模型中，我们允许双方各自输出部分结果。我们只要求每个给定的比特恰好由一位参与者输出（另一位参与者对该比特保持沉默）。在有效的分裂计算中， $f(x,\ y)$的第一比特可以由Alice输出，而第二比特由Bob输出。

定义 A.13 (分裂计算)。 如果存在两个映射 $\mathcal{O}_\mathsf{A}$ 和 $\mathcal{O}_\mathsf{B}$，满足 $\mathcal{O}_\mathsf{A}$： $\mathcal{T}_\pi \times \mathcal{R}^{\mathsf{pub}} \times \mathcal{R}_\mathsf{A} \times \mathcal{X} \to \{0,1,*\}$，且类似地$\mathcal{O}_\mathsf{B}$： $\mathcal{T}_\pi \times \mathcal{R}^{\mathsf{pub}} \times \mathcal{R}_\mathsf{B} \times \mathcal{Y} \to \{0,1,*\}$，使得对于所有$(x,y) \in \mathcal{X} \times \mathcal{Y}$，则称协议 Π 以 ϵ 错误分裂计算 f：

$$\Pr_{r,r_\mathsf{A},r_\mathsf{B}}[\mathcal{O}_\mathsf{A}(t_\pi, r, r_\mathsf{A}, x) \bowtie \mathcal{O}_\mathsf{B}(t_\pi, r, r_\mathsf{B}, y) = f(x,y)] \geq 1 - \epsilon.$$

$$where\ (a \bowtie b)_i \begin{cases} a_i & if\ b_i = *, \\ b_i & if\ a_i = *, \\ * & otherwise. \end{cases}$$

我们称 $weave$ 二元算子 $\bowtie$ 为： $\{0,1,*\}^k \times \{0,1,*\}^k \to \{0,1,*\}^k$ 定义 A.13 末尾所述的13，定义 A.它会重新组合分配给各参与者的部分。

为了将此模型与二选一模型区分开来，我们引入一个问题，其中关于输出的信息自然地分割给两位参与者。我们以一种方式定义该问题，使其在分割模型中易于计算；而在二选一模型中，由于其中一位参与者必须汇总关于输出的完整信息，因此需要大量通信。

定义 A.14 (分裂恒等问题). $\mathbf{SplitId}_n$： $\{0,1\}^n \times \{0,1\}^n \to \{0,1\}^n$ 定义为

$$\mathbf{SplitId}_n(x,y)_i = \begin{cases} x_i & if\ i = 0 \mod 2, \\ y_i & otherwise. \end{cases}$$

0	0	2	2	0	0	2	2
1	1	3	3	1	1	3	3
0	0	2	2	0	0	2	2
1	1	3	3	1	1	3	3
4	4	6	6	4	4	6	6
5	5	7	7	5	5	7	7
4	4	6	6	4	4	6	6
5	5	7	7	5	5	7	7

图 8：**分裂恒等问题**3 的通信矩阵

定理 A.15. $\forall f$： $\mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$ 其中 $k = \lceil \log|\mathcal{Z}| \rceil$ 且 $\epsilon > 0$

$$R_\epsilon^{\mathsf{spl}}(f) \leq R_\epsilon^{\mathsf{1of2}}(f) \leq R_\epsilon^{\mathsf{spl}}(f) + \lfloor k/2 \rfloor + 1, \quad and$$
$$D^{\mathsf{spl}}(\mathbf{SplitId}_n) \in O(1), \qquad R_\epsilon^{\mathsf{1of2}}(\mathbf{SplitId}_n) \in \Omega(n).$$

定理 *A.15* 的证明15定理 *A.*。 这里有一个细微之处：参与者在分割协议结束时，可能会犯下符号过多或过少的错误 $*$ 。我们的证明$R_\epsilon^{\mathsf{1of2}}(f) \leq R_\epsilon^{\mathsf{spl}}(f) + \lfloor k/2 \rfloor + 1$不能依赖这一假设：例如，我们不能说“符号较少的参与者先发言”，因为这可能导致协议产生歧义。

证明 $R_\epsilon^{\mathsf{1of2}}(f) \leq R_\epsilon^{\mathsf{spl}}(f) + \lfloor k/2 \rfloor + 1$**：** 设 Π 是一个最优分割协议。在 Π结束时，Alice 计算她在分割协议中会输出多少个 $*$ 符号。如果该数量大于 $\lfloor k/2 \rfloor$，她就发送 1 比特；否则发送 0。如果她发送了 0，接着她会发送 $\lfloor k/2 \rfloor$比特，其中前若干比特按顺序对应于她在分割协议中按顺序会输出的非$*$符号。如果她发送了 1，则由 Bob 发送他在分割协议中会输出的前 $\lfloor k/2 \rfloor$ 个非$*$ 比特。在这两种情况下，如果可发送的比特不足，参与者会按需补上 0，以达到 $\lfloor k/2 \rfloor$ 比特。

如果发送其分割输出中的非$*$ 符号的是 Alice，那么 Bob 会在一取二协议的最后一步输出结果之前，用 Alice 发送的比特替换其分割输出中的 $*$符号。如果发送其非$*$ 比特的是 Bob，情况则对称。如果用于替换 $*$ 符号的比特过多或不足，就丢弃多余比特，或直接补上 0。

该协议没有歧义（它不依赖 Alice 和 Bob 合计恰好没有 k颗星这一条件），并且只要原始协议在分割模型中正确，它在二选一模型中也正确。

分　　离结果再次界定了不会产生过多错误的矩形大小　　s.

分割模型上界： Alice 将 x 的奇数位置替换为 $*$，Bob 将 y 的偶数位置替换为 $*$。随后，他们各自输出替换后的字符串，从而在分割模型中计算$\mathbf{SplitId}_n(x,\ y)$。这无需通信。

二选一模型下界： 考虑 $D^{1\text{of}2}_{1/4}(\mathbf{SplitId}_n,\ \mu)$，其中 μ 是定义在$(x,\ y)$上的均匀分布，满足奇数 i 对应 $x_i = 0$，偶数 i对应 $y_i = 0$；再考虑这个缩减后（但仍为全定义）的问题的$\widetilde{\ }$通信矩阵 $M_{\mathbf{SplitId}_n}$。这将输入数量减少到 2^n。令 Π 为通信量为 $C = D^{1\text{of}2}_{1/4}(\mathbf{SplitId}_n,\ \mu)$的最优确定性二选一协议。

$\sim\Pi$用条纹矩形划分通信矩阵 $M_{\mathbf{SplitId}_n}$：在任意给定的矩形中，二选一协议的输出只能取决于行或列之一，不能同时取决于二者。但对于我们的问题，通信矩阵中的每个单元格都有不同的输出，因此，任何宽和高都至少为 2 的矩形，至少有一半单元格会产生错误。

宽或高至多为 1 的矩形最多包含 $2^{n/2}$ 个元素，因此，错误率低于其所含元素一半的矩形至多覆盖$2^{C+n/2}$ 个元素。因此，至少有 $2^n - 2^{C+n/2}$ 个输入被错误率至少为 $1/2$ 的矩形覆盖，所以Π的错误率至少为$2^{-n} \cdot \frac{1}{2}\left(2^n - 2^{C+n/2}\right)$。该错误率必须低于$\frac{1}{4}$，因此：

$$\frac{1}{4} \geq 2^{-n} \cdot \frac{1}{2}\left(2^n - 2^{C+n/2}\right) \Rightarrow C \geq n/2 - 1$$

这就完成了我们的证明，即 $R^{1\text{of}2}_{1/4}(\textbf{分割标识}_n) \geq D^{1\text{of}2}_{1/4}(\textbf{分割标识}_n,\ \mu) \in \Omega(n)$。

□

A.5.1 异或模型

在我们最后讨论的模型中，协议结束时两个参与者都会输出一个 k 比特串。当两个字符串逐位异或的结果等于 $f(x,\ y)$ 时，该计算便正确计算出 $f(x,\ y)$ 的值。

定义 4.1 (异或计算)。考虑一个输出集合为 $\mathcal{Z} = \{0,1\}^k$ 的函数 f。如果存在两个映射 $\mathcal{O}_\mathsf{A}$ 和 $\mathcal{O}_\mathsf{B}$，满足$\mathcal{O}_\mathsf{A}$：$\mathcal{T}_\pi \times \mathcal{R}^{\mathsf{pub}} \times \mathcal{R}_\mathsf{A} \times \mathcal{X} \to \{0,1\}^k$，且同样满足 $\mathcal{O}_\mathsf{B}$：$\mathcal{T}_\pi \times \mathcal{R}^{\mathsf{pub}} \times \mathcal{R}_\mathsf{B} \times \mathcal{Y} \to \{0,1\}^k$，使得对于所有 $(x,y) \in \mathcal{X} \times \mathcal{Y}$，则称协议 Π 以 ϵ 的 f 错误进行异或计算，

$$\Pr_{r,r_\mathsf{A},r_\mathsf{B}}[\mathcal{O}_\mathsf{A}(t_\pi, r, r_\mathsf{A}, x) \oplus \mathcal{O}_\mathsf{B}(t_\pi, r, r_\mathsf{B}, y) = f(x,y)] \geq 1 - \epsilon.$$

Th异或模型与二选一模型的区别在于以下函数：

定义 A.16。**异或**$_n$：$\{0,1\}^n \times \{0,1\}^n \to \{0,1\}^n$ 定义为 **异或**$_n(x,\ y) = (x_i \oplus y_i)^{i\in[n]}$。

0	1	2	3	4	5	6	7
1	0	3	2	5	4	7	6
2	3	0	1	6	7	4	5
3	2	1	0	7	6	5	4
4	5	6	7	0	1	2	3
5	4	7	6	1	0	3	2
6	7	4	5	2	3	0	1
7	6	5	4	3	2	1	0

图 9：**异或**$_3$ 的通信矩阵

定理 A.17。 $\forall f:\ \mathcal{X}\times\mathcal{Y}\to\mathcal{Z}$ 其中 $k=\lceil\log|\mathcal{Z}|\rceil$ 以及 $\epsilon>0$，

$$R_\epsilon^{\mathsf{xor}}(f)\le R_\epsilon^{\mathsf{spl}}(f)\le R_\epsilon^{\mathsf{1of2}}(f)\le R_\epsilon^{\mathsf{uni}}(f)\le R_\epsilon^{\mathsf{xor}}(f)+k,\quad \textit{and}$$
$$D^{\mathsf{xor}}(\mathbf{XOR}_n)=0,\qquad R_\epsilon^{\mathsf{spl}}(\mathbf{XOR}_n)\in\Omega(n).$$

定理 *A.17* 的证明*17*定理 *A.*。 **异或模型的上界：** Alice 和 Bob 只需各自输出自己的输入，无需通信。

分割模型下界： 考虑 $D_{1/4}^{\mathsf{spl}}$**(异或**$_n$**,** μ**)**，其中 μ 是均匀分布。设 Π 是一个最优的确定性一取二通信协议$C=D_{1/4}^{\mathsf{spl}}$**(异或**$_n$**,** μ**)**。

Π 将通信矩阵 $M_{\mathbf{XOR}_n}$ 划分为 2^C 个矩形。首先假设，在每个矩形中，输出的每个比特都由一个固定的参与者输出。稍后我们会看到，即使不作此假设，我们的论证仍然成立。

在每个 2^C 矩形中，至少有一位参与者输出的结果比特数少于 $n/2$ 。考虑一个 Bob 最多输出结果一半比特的矩形。那么，在该矩形的一行中，至多有 $2^{n/2}$ 种不同的输出。但对于**异或**$_n$ 问题，给定一行中的所有单元格都有不同的输出。我们将说明，这限制了错误很少的矩形的大小。

设一个矩形至少包含 $2^{3n/2+1}$ 个元素。由于每行或每列至多包含 2^n 个元素，这样的矩形至少包含 $2^{n/2+1}$ 行和列。因此，在分割模型中，输出至多一半输出位的参与者，在某一行或某一列中至多输出 $2^{n/2}$ 个不同的字符串；而该行或列包含超过 $2^{n/2+1}$ 个不同的值，所以该矩形中至少一半的元素存在错误。

如果参与者并非总是以相同方式分割输出位，则考虑最大的行集合，使得 Alice 输出给定的那一组输出位；再考虑最大的列集合，使得 Bob 输出给定的那一组输出位。如果 Alice 在这些行上输出的输出位集合与 Bob 在这些列上输出的输出位集合并非互补，则该矩形中至少一半的元素存在错误。如果这些集合恰好将输出位划分开来，我们就沿用之前的论证：假设 Bob 在我们定义的子矩形中输出至多一半的比特。那么在该子矩形的任何一行中，正确的单元格至多有 2^n 个；而子矩形之外的行也大多存在错误，因此该矩形中至少一半的元素存在错误。

至多有 $2^{C+3n/2+1}$ 个元素位于错误严格小于一半的矩形中，因此该协议产生的错误至少为 $\frac{1}{2}\cdot 2^{-2n}\left(2^{2n}-2^{C+3n/2+1}\right)$。错误必须小于 $1\,4$，因此：
–

$$C\ge n/2-2$$

由此完成了我们的证明： $R_{1/4}^{\mathsf{spl}}$（**异或**$_n$） $\ge D_{1/4}^{\mathsf{spl}}$（**异或**$_n$， μ） $\in\Omega(n)$。

□

A.6 模型之间的关系

下一个命题总结了定理 A.5、定理 A.A.8、A.A.12、A.A.15 A.和 A.17A.。

命题 A.18. $\forall f:\ \mathcal{X}\times\mathcal{Y}\to\mathcal{Z}$ 其中 $k=\lceil$对数$|\mathcal{Z}|\rceil$ 且 $\epsilon>0$ 我们有：

$$
\begin{aligned}
R_\epsilon^{\mathsf{open}}(f)\geq R_\epsilon^{\mathsf{loc}}(f)&\geq \max\Big(R_\epsilon^{\mathsf{A}}(f),R_\epsilon^{\mathsf{B}}(f)\Big) && (1)\\
&\geq \min\Big(R_\epsilon^{\mathsf{A}}(f),R_\epsilon^{\mathsf{B}}(f)\Big)=R_\epsilon^{\mathsf{uni}}(f)\\
&\geq R_\epsilon^{\mathsf{1of2}}(f)\geq R_\epsilon^{\mathsf{spl}}(f)\geq R_\epsilon^{\mathsf{xor}}(f)\\
R_{2\epsilon}^{\mathsf{loc}}(f)&\leq R_\epsilon^{\mathsf{A}}(f)+R_\epsilon^{\mathsf{B}}(f) && (2)\\
R_\epsilon^{\mathsf{open}}(f)&\leq R_\epsilon^{\mathsf{uni}}(f)+k && (3)\\
R_\epsilon^{\mathsf{open}}(f)&\leq R_\epsilon^{\mathsf{1of2}}(f)+k+1 && (4)\\
R_\epsilon^{\mathsf{1of2}}(f)&\leq R_\epsilon^{\mathsf{spl}}(f)+\lceil k/2\rceil+1. && (5)\\
R_\epsilon^{\mathsf{uni}}(f)&\leq R_\epsilon^{\mathsf{xor}}(f)+k. && (6)
\end{aligned}
$$

对于确定性通信和仅使用私有随机性的通信，同样的陈述也成立。除子命题 *2* 之外的所有陈述 *2* 子命题也适用于关系和非确定性通信。

命题 A.18 命题 A.表明，这些模型构成一个自然的层级，并且可以按通信密集程度从高到低排列。我们还在正文的图 1 中总结了这一层级1，图 。该图还展示了本节及附录中的其他分离问题。

B 我们的结果总结

本节总结本文的结果。表 1 表总结了我们研究过的、揭示不同输出模型之间存在差距的问题。表 2 表总结了 **GapMAJ∘异或** 在各种模型中的界。表 3 表总结了错误缩减界和去随机化结果。

	open	局部	单向	1 取 2	XOR
$\mathbf{EQ}_n^{\mathsf{out}}$	$R_{1/3}^{\mathcal{M}}\in\Theta(n)$	$R_{1/3}^{\mathcal{M}}\in\Theta(1)$			
t **–INT**$_n$	$R_{1/3}^{\mathcal{M}}\in\Theta(t\cdot\log(n))$	$R_{1/3}^{\mathcal{M}}\in\Theta(t)$			
$\mathbf{id}_n^{\mathsf{A}}$	$R_{1/3}^{\mathcal{M}}\in\Theta(n)$		$D^{\mathcal{M}}=0$		
条件标识符$_n$	$R_{1/3}^{\mathcal{M}}\in\Theta(n)$			$D^{\mathcal{M}}=2$	
$\mathbf{MAX}_n$	$R_{1/3}^{\mathcal{M}}\in\Theta(n)$			$R_{1/3}^{\mathcal{M}}\in\Theta(\log(n)$	
$t-\mathbf{FtFD}_n$	$R_{1/3}^{\mathcal{M}}\in\Theta(\log(n))$			$R_{1/3}^{\mathcal{M}}\in\Theta(\log(t)+\log\log(n))$	
$\mathbf{XOR}_n$	$R_{1/3}^{\mathcal{M}}\in\Theta(n)$				$D^{\mathcal{M}}=0$
$\mathbf{GapMAJ}_{N,\ k,\ 1/3}$**∘异或**	$R_{1/3}^{\mathcal{M}}\in\Theta(k)$				$R_{1/3}^{\mathcal{M}}=0$
$\mathbf{GapMAJ}_{N,\ k,\ 2/5}$**∘异或**	$R_{1/3}^{\mathcal{M}}\in\Theta(k)$				$R_{1/3}^{\mathcal{M}}\in O(1)$

表 1：所有模型中我们分离问题的通信复杂度汇总。问题的定义和证明见附录A附录A 和 HH。在此表中，n 表示输入长度，k 表示输出长度，$\mathcal{M}$ 表示输出模型，$\mathcal{M}\in\{$开放，loc、A、B、uni、1of2、xor$\}$，而 t 表示一个实例的汉明重量。

关于间隙多数问题的上界总结见表 2Table。我们猜想存在一个与我们所述确定性 $O(\epsilon Nk)$上界相匹配的下界。研究

上界		
$\epsilon' \geq \epsilon$	$R^{\mathsf{xor}}_{\epsilon'}$	0
	$R^{\mathsf{xor},\ \mathsf{priv}}_{\epsilon'}$	$\log(N)$
	$R^{\mathsf{open}}_{\epsilon'}$	$2k$
	$R^{\mathsf{open},\ \mathsf{priv}}_{\epsilon'}$	$2k + \log(N)$
$0 < \epsilon' < \epsilon$	$R^{\mathsf{xor}}_{\epsilon'}$	$O\big(\text{最小}\big(C_{\epsilon,\epsilon'}, N + \text{对数}\big(\frac{1}{\epsilon'}\big)\big)\big)$
$\epsilon' = 0$	D^{uni}	$(2\epsilon N + 1)k$

表 2：GapMAJ∘**异或**的上界，证明见附录 G附录 G。在此表中，N，k，ϵ 是间隙多数问题的参数，而 ϵ' 是误差参数。

错误缩减		
模型	上界	（条件）
open 局部 单向	$R_{\epsilon'}(f) \leq C_{\epsilon,\epsilon'} \cdot R_{\epsilon}(f)$	
二选一	$R_{\epsilon'}(f) \leq C_{\epsilon,\ \epsilon'}(R_{\epsilon}(f)\ + 1)\ + C'_{\epsilon\epsilon'}$	
分割	$R_{\epsilon'}(f) \leq C_{\epsilon,\epsilon'} R_{\epsilon}(f) + O\big(C_{\epsilon,\epsilon'}\big)$	
XOR	$R_{\epsilon'}(f) \leq C_{\epsilon,\ \epsilon'} R_{\epsilon}(f) + O\big(C_{\epsilon,\ \epsilon'}\big)$	
	$R_{\epsilon'}(f)\ \leq 50\ln\big(\frac{12}{\epsilon'}\big) R_{\epsilon}(f) + C_{\epsilon,\ \epsilon'} R_{\epsilon}(g) + O\big(C_{\epsilon,\ \epsilon'} + \log(k)\big)$	$(f = g^{\otimes k})$
去随机化		
模型	上界	（条件）
open 局部	$D(f)\ \in O\Big(2^R\Big(R + \log\big(\frac{1}{\frac{1}{2}-\epsilon}\big)\Big)\Big)$	
单向	$D(f)\ \in O\Big(2^R\Big(R + \log\big(\frac{1}{\frac{1}{2}-\epsilon}\big)\Big)\Big)$	
二选一	$D(f)\ \in O\Big(2^R\Big(R + \text{对数}\big(\frac{1}{\frac{1}{2}-\epsilon}\big)\Big)\Big)$	
	$D(f)\ \in O\Big(2^R\Big(R + \text{对数}\big(\frac{1}{\frac{1}{2}-\epsilon}\big)\Big) + \text{对数}(k)\Big)$	
分割	$D(f)\ \in O\Big(2^R\Big(R + \log\big(\frac{1}{\frac{1}{2}-\epsilon}\big)\Big) + k\Big)$	
	$D(f)\ \in O\Big(2^R\Big(R + \text{对数}\big(\frac{1}{\frac{1}{2}-\epsilon}\big)\Big) + 2^R\big(\frac{1}{2}-\epsilon\big)^{-2}k\Big)$	
XOR	$D(f)\ \in O\Big(2^R\Big(R + \log\big(\frac{1}{\frac{1}{2}-\epsilon}\big)\Big) + 2^R\big(\frac{1}{2}-\epsilon\big)^{-2}k\Big)$	

表 3：错误缩减和去随机化方案总结。在上述所有陈述中，f 是一个输出长度为 k的函数，ϵ 是初始误差参数，ϵ' 是目标误差参数，$R = R^{\mathcal{M}}_{\epsilon}(f)$，$C_{\epsilon,\ \epsilon'} \in O\Big(\epsilon\big(\frac{1}{2}-\epsilon\big)^{-2}\log\big(\frac{1}{\epsilon'}\big)\Big)$ 和$C'_{\epsilon,\ \epsilon'} \in O\Big(\log\big(\frac{1}{\epsilon'}\big) + \log\Big(\frac{1}{\frac{1}{2}-\epsilon}\Big)\Big)$。

我们在本文中看到，这个问题的通信复杂度具有理论意义，因为通信复杂度中的基本结果，即错误缩减和去随机化，都与异或模型中的 **GapMAJ∘异或** 问题有关。通过定理 6**GapMAJ∘异或**，改进 6.44Theorem. 的确定性上界将带来更好的去随机化结果。类似地，改进随机化上界也可能通过引理 55.55Lemma .. 反过来，鉴于我们已经为 **GapMAJ∘异或** 的私有硬币 XOR 通信复杂度证明了一个 $\log(N)$ 上界，若能证明其确定性通信复杂度具有 $\Omega(Nk)$ 下界，就表明我们在异或模型中的去随机化定理（定理 66.44Theorem.）接近紧确。

C 弱划分界

弱划分界可用于证明开放模型的下界。我们利用它说明，该模型对“非平凡”或“典型”输出的数量非常敏感；这些输出出现得足够频繁，其含义将在定义 C.3 中精确定义3Definition C.。

定义 C.1 (弱划分界[FJK+16]FJK)。 我们定义（使用记号 $\beta = \sum_{x,\ y} \beta_{x,\ y}$）

$$\begin{aligned} \mathrm{wprt}_\epsilon^\mu(f) = \max_{\alpha \geq 0,\, \beta_{xy} \geq 0} \quad & (1-\epsilon)\alpha - \beta \\ \textit{subject to :} \quad & \alpha\mu(R \cap f^{-1}(z)) - \beta(R) \leq 1 \qquad \forall R, z, \qquad (7) \\ & \alpha\mu_{xy} - \beta_{xy} \geq 0 \qquad \forall (x,y). \qquad (8) \end{aligned}$$

f 的非分布弱划分界为 $\mathrm{wprt}_\epsilon(f) = \max_\mu \mathrm{wprt}_\epsilon^\mu(f)$。

请注意，此处的定义与 Fontes 等人给出的定义略有不同 [FJK+16]FJK。对于布尔函数，这两种表述是等价的；而该论文所考虑的正是这种情形。

命题 C.2 ([JK1010，JKFJK+16]FJK)。 设 $0 < \epsilon < 1/2$ 并设 $f\colon \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$ 是一个函数。那么，

$$\log(\mathrm{wprt}_\epsilon(f)) \leq \log(\mathrm{prt}_\epsilon(f)) \leq R_\epsilon^{\mathsf{open}}(f).$$

右侧来自 *[JK10] JK*，左侧来自 *[FJK+16]FJK*。

接下来我们引入相对于分布 μ 的 ϵ-输出最小集这一概念。稍微滥用一下记号，在无需指明我们隐含指代的是哪个 f 时，将 $\mu(f^{-1}(z))$ 记作 $\mu(z)$。

定义 C.3。 令 $\mathcal{Z}$ 为函数 $f\colon \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$ 的输出集合。
进一步假设 $\mathcal{Z} = \{z_1, z_2, \ldots, z_n\}$ 关于 μ 已排序，即：

$$i \leq j \Rightarrow \mu(z_i) \geq \mu(z_j).$$

然后将 $\xi_\epsilon^\mu(f)$ 定义为：

$$\xi_\epsilon^\mu(f) = \min\left\{ k \;\middle|\; \sum_{i=1}^{k} \mu(z_i) \geq 1 - \epsilon \right\}.$$

定理 C.4。 设 $0 < \epsilon < 1/2$，令 $f\colon \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$ 为一个函数，令 μ 为 $\mathcal{X} \times \mathcal{Y}$ 上的一个分布。那么，

$$\xi_\epsilon^\mu(f) - 1 \leq \mathrm{wprt}_\epsilon^\mu(f).$$

定理 C.4 的证明4定理 C.。 按 μ 对输出集合排序（即 $z_1 \le z_2 \le \ldots \le z_n$），并设 $z_{\min} = z_{\xi_\epsilon^\mu(f)}$。考虑以下变量赋值：

$$\alpha = \frac{1}{\mu(z_{\min})}, \qquad \beta_{xy} = \max\left(0, \mu_{xy} \cdot \left(\alpha - \frac{1}{\mu(f(x,y))}\right)\right).$$

于是，wprt_ϵ^μ 的第一个约束得到满足。事实上，令 z 满足 $\mu(z) \le \mu(z_{\min})$（因此，对所有 $(x,y) \in f^{-1}(z)$，$\beta_{xy} = 0$ 都成立）。那么，对所有 R：

$$\begin{aligned}
&\alpha \cdot \mu(R \cap f^{-1}(z)) - \beta(R) \\
\le\ &\alpha \cdot \mu(R \cap f^{-1}(z)) - \beta(R \cap f^{-1}(z)) \\
=\ &\alpha \cdot \mu(R \cap f^{-1}(z)) \le \alpha\mu(z) = \frac{\mu(z)}{\mu(z_{\min})} \le 1.
\end{aligned}$$

当 z 满足 $\mu(z) > \mu(z_{\min})$ 时，对所有 R：

$$\begin{aligned}
&\alpha \cdot \mu(R \cap f^{-1}(z)) - \beta(R) \\
\le\ &\alpha \cdot \mu(R \cap f^{-1}(z)) - \beta(R \cap f^{-1}(z)) \\
=\ &\alpha \cdot \mu(R \cap f^{-1}(z)) - \left(\alpha - \frac{1}{\mu(z)}\right)\mu(R \cap f^{-1}(z)) \\
=\ &\frac{\mu(R \cap f^{-1}(z))}{\mu(z)} \le 1.
\end{aligned}$$

第二个约束也得到满足：

$$\forall x, y: \quad \alpha\mu_{xy} - \beta_{xy} = \alpha\mu_{xy} - \max\left(0, \mu_{xy}\left(\alpha - \frac{1}{\mu(z^{xy})}\right)\right) \ge 0.$$

该可行解的值为：

$$\begin{aligned}
(1-\epsilon)\alpha - \beta &= (1-\epsilon)\frac{1}{\mu(z_{\min})} - \sum_{z:z=z_i, i<\xi_\epsilon^\mu(f)} \beta(z_i) \\
&= \left(1 - \epsilon - \sum_{z:z=z_i, i<\xi_\epsilon^\mu} \mu(z_i)\right)\frac{1}{\mu(z_{\min})} + \xi_\epsilon^\mu(f) - 1 \\
&\ge \xi_\epsilon^\mu(f) - 1.
\end{aligned}$$

□

D 错误缩减

D.1 随机图引理的证明

第 5 节中陈述的随机图引理5.11 Section .以及用于求解 **GapMAJ∘异或**，是对 Erd˝os 和 R´enyi [ER60]ER 所证明结果的一个简单变体。他们证明的结果基于一种随机图模型：从所有可能的边中随机选取固定数量的边；而我们关注的模型是以固定概率选取每条边 p，且各边的选取相互独立。这两种模型的渐近行为已知基本相似。对随机图理论感兴趣的读者可参阅 [Bol01]Bol。

引理 5 的证明5.77引理 .. 我们注意到，与 [ER60] ER中的情况一样，如果不存在顶点数超过 $(1-\alpha)n$ 的连通分量，那么我们可以将顶点划分为两个不连通集合，其大小分别为 n_0和 n_1，使得 $\frac{\alpha}{2}n \le n_0 \le n_1 \le \left(1 - \frac{\alpha}{2}\right)n$。

给定一个将顶点划分为大小分别为 n_0 和 n_1 的集合的划分，这两个集合不连通的概率为 $(1-p(n))^{n_0 n_1}$。由于 $p(n)=\frac{c}{n}$，且可能的划分少于 2^n 种，因此不存在顶点数超过 $(1-\alpha)n$ 的连通分量的概率有如下上界：

$$2^n\left(1-\frac{c}{n}\right)^{n_0 n_1} \le 2^n e^{-c\frac{n_0 n_1}{n}} \le 2^n e^{-c\frac{\alpha}{2}\left(1-\frac{\alpha}{2}\right)n} = e^{\left(\ln(2)-\frac{\alpha}{2}\left(1-\frac{\alpha}{2}\right)c\right)n}$$

□

D.2 异或模型中的错误缩减

D.2.1 二选一模型中的错误缩减

二选一模型本身已具有非平凡性。如果我们重复该协议，在某些运行中 Alice 会输出结果，而在另一些运行中 Bob 会输出结果，并且双方的多数输出都有可能是错误的。在该模型中，一种简单的错误缩减方法是将二选一协议转换为单边模型（命题 A.18[18 命题 A.]），然后应用定理 5[5.22，定理。]以获得$R^{1of2}_{\epsilon'}(f) \le C_{\epsilon,\ \epsilon'} \cdot (R^{1of2}_{\epsilon}(f)\ + k)$。

我们证明，可以消除对输出长度 k 的额外依赖。我们表明，参与者可以将多数结果的候选项数量缩减至最多四个。通过哈希，可以以高概率筛选出获胜结果。

定理 D.1。设 $0<\epsilon'<\epsilon<\frac{1}{2}$, $C_{\epsilon,\epsilon'} = \frac{2\epsilon}{(1-\epsilon)\left(\frac{1}{2}-\epsilon\right)^2}\ln\left(\frac{4}{\epsilon'}\right)$ 且 $C'_{\epsilon,\epsilon'} \le 18+4\log\left(\frac{1}{\epsilon'}\right)+4\log\left(C_{\epsilon,\epsilon'}\right)$。对于所有函数 $f\colon \mathcal{X}\times\mathcal{Y}\to\mathcal{Z}$,

$$R^{1of2}_{\epsilon'}(f) \le C_{\epsilon,\epsilon'}(R^{1of2}_{\epsilon}(f)+1)+C'_{\epsilon,\epsilon'}.$$

定理 *D.1* 的证明[1定理 D.]. 固定一个错误率至多为 f 的二选一协议 ϵ，并应用命题 A.10[10 命题 A.]，从而现在有一个二选一通信协议和若干映射，使得在任何一次执行结束时恰有一位参与者发言。根据霍夫丁不等式（引理 5[5.11引理 .]），如果参与者进行 $T=\lceil 8\epsilon(1-\epsilon)\left(\frac{1}{2}-\epsilon\right)^{-2}\ln\left(\frac{4}{\epsilon'}\right)\rceil$ 次执行，那么以至少 $1-\frac{\epsilon'}{2}$ 的概率，在该参与者的至少 $\frac{1}{2}+\frac{1}{2}\left(\frac{1}{2}-\epsilon\right)>\frac{1}{2}$ 次执行中，其中一位参与者输出$f(x,y)$（另一位保持沉默）。

参与者想要确定正确的输出。我们认为，他们只需付出极少的额外通信和错误代价就能做到这一点。注意，如果某个值在上述执行中被输出的次数严格超过一半，那么它必定被两位参与者中的一位输出了严格超过 $T/4$ 次。因此，每位参与者只需考虑在自己这边出现次数严格超过 $T/4$ 次的输出。我们将 Alice 和 Bob 两边分别识别为 $f(x,y)$ 的候选项的输出记作 $(z^{\mathsf{A}}_i)_{i\in[n_{\mathsf{A}}]}$ 和 $(z^{\mathsf{B}}_j)_{j\in[n_{\mathsf{B}}]}$，其中 n_{A} 和 n_{B} 分别是两边候选项的数量。由于共有 T 次执行，每个候选项被输出的次数都严格超过 $T/4$，且某个值被输出的次数严格超过 $T/2$，因此候选项最多有 3 个，并且 $n_{\mathsf{A}}\le 2$ 和 $n_{\mathsf{B}}\le 2$。

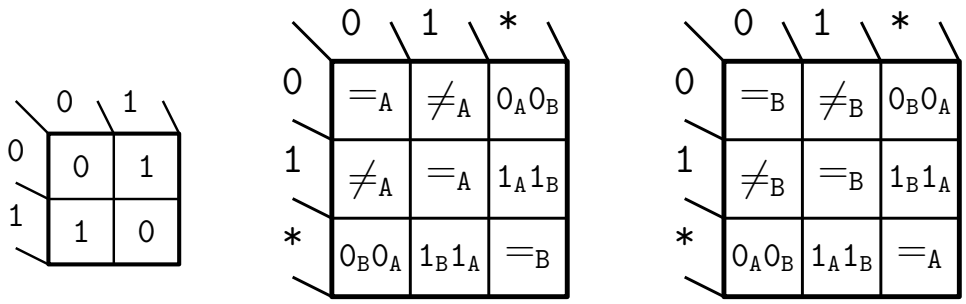
参与者利用公共随机性选取一个随机哈希函数 $h:\ \mathcal{Z}\to[m]$其中 m 的选取应确保：以高概率，参与者选出的候选项之间不会发生碰撞（z^{A}_i）$_{i\in[n_{\mathsf{A}}]}$ 和（z^{B}_j）$_{j\in[n_{\mathsf{B}}]}$。由于给定碰撞发生的概率为$\frac{1}{m}$，且候选项对共有 $n_{\mathsf{A}}+n_{\mathsf{B}}\le\binom{4}{2}$ 个，取 $m=\lceil\frac{12}{\epsilon'}\rceil\ge\frac{12}{\epsilon'}$即可保证此类碰撞发生的概率仅为 $\le\frac{\epsilon'}{2}$。随后，参与者交换各自候选项的哈希值$h_1,\dots,h_{n_{\mathsf{A}}+n_{\mathsf{B}}}$（对应于 $h(z^{\mathsf{A}}_i)$ 和 $h(z^{\mathsf{B}}_j)$，分别属于 $i\in[n_{\mathsf{A}}]$ 和 $j\in[n_{\mathsf{B}}]$），通信量为 $4\lceil\log(m)\rceil$ 比特。对于每个 k，Alice 计算 $\alpha_k=|\{i:\ h(z^{\mathsf{A}}_i)=h_k\}|$，Bob 计算 $|\{j:\ h(z^{\mathsf{B}}_j)=h_k\}|$。Alice 将她的计数（$\alpha_k$）$_{k=1,\dots,n_{\mathsf{A}}+n_{\mathsf{B}}}$ 发送给 Bob（通信量为 $\le 4\lceil\log(T)\rceil$）。Bob 回复 $k\in[n_{\mathsf{A}}+n_{\mathsf{B}}]$，其中 h_k 是通过 h哈希到该值的输出最多的哈希值。如果该哈希值是 Alice 的某个候选项 z^{A}_i的哈希值，她就输出该候选项；否则，由 Bob 输出他对应的候选项。将偏差（Hoeffding）导致的错误与碰撞导致的错误相加后，该协议的错误至多为 ϵ'。

□

D.2.2 分割模型中的错误缩减

值得注意的是，分割模型中的错误缩减可以通过与异或模型方案非常相似的方法实现。请注意，仅将星号替换为零并套用异或方案并不足够，因为输出也应当进行分割（而异或方案中的输出不一定具有这种形式）。更准确地说，应用定理55.33 定理 。将表明$R_{\epsilon'}^{\mathsf{xor}}(f) \le C_{\epsilon,\ \epsilon'} \cdot R_{\epsilon}^{\mathsf{spl}}(f) + O\big(C_{\epsilon,\ \epsilon'}\big)$。

我们在异或模型中缩减错误时所用的关键观察是：当 GapMAJ 矩阵的两行 i和 j 在逐位异或后得到相同字符串时，也就是说， $X_i \oplus Y_i = X_j \oplus Y_j$，我们观察到 $X_i \oplus X_j = Y_i \oplus Y_j$。这使我们能够通过对两个可局部计算的字符串进行一次相等性测试，来检验两行逐位异或后是否得到相同字符串。我们将这一局部操作称为“兼容性构件”，即一个函数 g ，参与者在行对上局部应用该函数，使得检验 $X_i \oplus Y_i = X_j \oplus Y_j$的问题可归约为检验$g(X_i, X_j)$ 与 $g(Y_i, Y_j)$ 之间的相等性。在异或模型中，兼容性构件 g就是逐位异或。逐位兼容性构件如图 10a 所示。10a图 a。



(a) 异或小工具　(b) Alice 的分割小工具 g_{A}　(c) Bob 的分割小工具 g_{B}

图 10：异或模型和分割模型的兼容性构件矩阵。

事实证明，我们可以在分割模型中进行类似的操作，只需稍作调整。双方参与者不再是在检验相等性之前对行对应用相同的小工具，而是分别应用不同的小工具。他们逐比特应用于行对的函数是变换 g_{A} 和 g_{B} ，如图 1010b图 b 和 1010cc 所示。这些函数经过选择，满足以下性质。

命题 D.2. 对于所有 X_i， X_j， Y_i，以及 $Y_j \in \{0,1,*\}^k$，还有 g_{A}， g_{B} 如图 *1010b图 b*和 *1010cc*，

$$X_i \bowtie Y_i = X_j \bowtie Y_j \Leftrightarrow g_{\mathsf{A}}(X_i, X_j) = g_{\mathsf{B}}(Y_i, Y_j)$$

这些函数刻画了一对行何时输出相同结果：如果 Alice 在 X_i 和 X_j的某个位置输出两个星号，那么 Bob 就需要在他的字符串 Y_i 和 Y_j的同一位置输出两个 0 或两个 1。类似地，如果 Alice 在索引 X_i 的行中输出一个星号，而在行 X_j中输出一个 0，那么 Bob 需要在同一索引处，在 Y_i 中输出一个 0，并在 Y_j中输出一个星号，从而使两行产生相同的结果。

命题 D.22命题 D.意味着，分割模型中的错误缩减归结为求解 **GapMAJ**与 weave 小工具（⋈）的组合，这与异或模型中的错误缩减归结为求解**GapMAJ**∘**异或**的情形相同。由此得到以下类似结果：定理 D.33定理 D.。

定理 D.3。 令 $0 < \epsilon' < \epsilon < \frac{1}{2}$, $C_{\epsilon,\ \epsilon'} = 8\epsilon\big(\frac{1}{2} - \epsilon\big)^{-2}\ln\big(\frac{4}{\epsilon'}\big)$。对于所有 $f\colon \ \mathcal{X} \times \mathcal{Y} \to \{0,1\}^k$,

$$R_{\epsilon'}^{\mathsf{spl}}(f) \le C_{\epsilon,\epsilon'} \cdot R_{\epsilon}^{\mathsf{spl}}(f) + O\big(C_{\epsilon,\epsilon'}\big).$$

E 直和问题的错误缩减

本节给出定理 55.88 定理 .给出的完整证明，该定理为形如 $f = g^{\otimes k}$. 的函数给出了错误缩减方案

定理 5 的证明5.88定理 .. 考虑一个用于 $f = g^{\otimes k}$、错误率至多为 ϵ 的 XOR 协议，以及一个用于 g、错误率至多为 ϵ 的协议。要实现错误率 ϵ' 的协议按如下方式进行。

步骤 1: [将范围限制为至多两个候选项。] 参与者运行 f，总共进行 $T_{\epsilon'} = 50\ln\left(\frac{12}{\epsilon'}\right)$ 次迭代。令 (a_i, b_i) 表示 Alice 和 Bob 在第 i^{th} 次迭代中本会输出的结果。与定理 51 证明中的步骤 相同5.66，定理 。以高概率，$|\{i : a_i \oplus b_i = f(x,y)\}| \geq \frac{2}{5}T_{\epsilon'}$。

> 与定理 52 证明中的步骤 相同5.66，定理 。随后，参与者求解随机的 **EQ** 实例，以找出计算结果相同的较大迭代子集。以高概率（$\geq 1 - \frac{\epsilon'}{6}$），他们最多计算 $O(T_{\epsilon'})$ 个 **EQ**实例，错误率为 $\frac{\epsilon'}{6}$。
> 以高概率（$\geq 1 - 3 \cdot \frac{\epsilon'}{6}$），参与者应该已经找到了一个或两个集合，每个集合至少包含 $\frac{11}{30}T_{\epsilon'}$ 次迭代，且集合中的所有迭代都计算出相同的值。如果只找到一个这样的大集合，参与者就输出 a_i 和 b_i，其中 i 是该大集合中任意一次迭代的索引。否则，令 i_1 和 i_2 分别为两个大集合中各自代表一次迭代的索引。

步骤 2: [查找一个关键索引 l。] 参与者将按照 i_1^{th} 次迭代或 i_2^{th} 次迭代的结果输出。为在两者之间作出选择，他们查找 $a_{i_1} \oplus a_{i_2}$ 和 $b_{i_1} \oplus b_{i_2}$ 之间的首个差异。这会得到一个索引 $l \in [k]$，在该索引处，两个可能的输出不同。我们称之为关键索引。

步骤 3。 [在关键索引 l 上求解 GHD。] 我们对 f C_ϵ 的第 l^{th} 比特异或计算 ϵ' 次。这会得到一个规模为 C_ϵ 的间隙汉明距离实例，ϵ' 其解以高概率确定正确输出的第 l^{th} 比特。参与者确定哪次迭代（i_1 或 i_2）在第 l^{th} 比特上正确，并按照该次迭代的结果输出。

总之，我们得到计算 f 且错误率为 ϵ' 时的以下上界。

$$R^{\mathsf{xor}}_{\epsilon'}(f) \leq \left(50\ln\left(\frac{12}{\epsilon'}\right)\right) \cdot R^{\mathsf{xor}}_{\epsilon}(f) + R^{\mathsf{loc}}_{\epsilon'/6}\left(\mathbf{EQ}_k^{\otimes O(T_{\epsilon'})}\right) + R^{\mathsf{loc}}_{\epsilon'/6}(\mathbf{FtFD}_k)$$
$$+ C_{\epsilon,\epsilon'} \cdot R^{\mathsf{xor}}_{\epsilon}(g) + R^{\mathsf{loc}}_{\epsilon'/6}\left(\mathbf{GHD}^{C_{\epsilon,\epsilon'}}_{(1/4+\epsilon/2)C_{\epsilon,\epsilon'},(3/4-\epsilon/2)C_{\epsilon,\epsilon'}}\right).$$

最后，我们通过应用查找第一个差异的已知上界[FRPU94]FRPU(命题 33.22命题 .)，来解决多个相等性实例 [FKNN9595，FKNN第 6] (命题 33.55命题 .)；而间隙汉明距离可以通过交换完整输入来解决，这在本质上是最优的[CR1212，CRVid1212，VidShe12]She。 □

F 消除随机性

F.1 通信记录分布估计

本节证明引理 66.33 引理 .以及 F.11F.。

引理 6 的证明6.33 引理 .。设 Π 是一个通信协议，并且 $\gamma = \delta|\mathcal{T}_\pi|^{-1}$。给定 (x, y)，参与者考察 Π的协议树：

这棵树的每个节点都表示协议的一次部分执行，因此我们用单词 $w \in \{0,1\}^*$ 来标记树的每个节点，该单词表示 Alice 和 Bob 为到达此节点而进行的通信。特别地，叶节点由完整通信记录标记，即单词$w \in \mathcal{T}_\pi$。我们将用记号 $w_{<i}$ 表示 w 中长度为 $(i-1)$ 的前缀。每个内部节点都属于 Alice 或 Bob，这一属性决定了在协议执行的这一特定时刻由谁发送下一条消息。它有 $|\mathcal{T}_\pi|$ 个叶节点。

对于每个内部节点 w，我们可以为其指定一个概率分布 p_w，该分布对应于接下来发送哪条消息（0 或 1）。如果该节点属于 Alice，则此分布完全由 x 决定

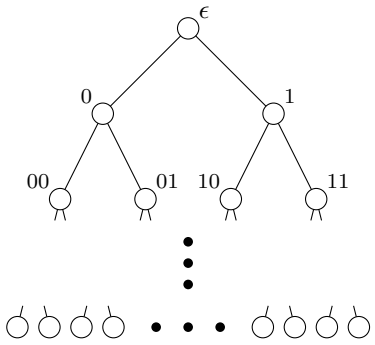


图 11：表示协议 Π 在给定（x，y）上的可能执行过程的树。

，否则由 y 决定。其随机性来自参与者的私有随机性，其支撑集是下一条消息的集合（此处为 0 或 1 ）。在协议树中的节点 w 处，Alice 下一条消息发送 1 的概率记为 $p_w(1 \mid x)$。

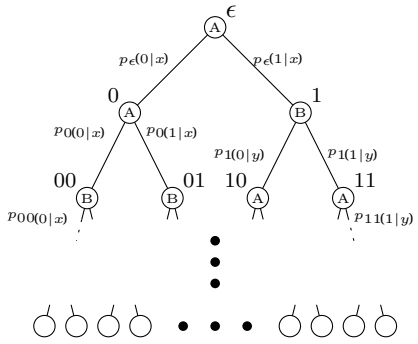


图 12：与图 11 相同的树11 图。其中节点根据其所有者和概率分布进行标记。注意 $p_w(0 \mid x) + p_w(1 \mid x) = 1$。

对于输入（x，y）上标签为 w的叶节点，协议的一次执行最终到达 w 的概率为：

$$p(w \mid x, y) = \underbrace{\left(\prod_{\substack{1 \le i \le |w| \\ w_{<i} \in \text{Alice}}} p_{w_{<i}}(w_i \mid x) \right)}_{\alpha(w|x)} \times \underbrace{\left(\prod_{\substack{1 \le i \le |w| \\ w_{<i} \in \text{Bob}}} p_{w_{<i}}(w_i \mid y) \right)}_{\beta(w|y)}.$$

对于每个 $w \in \mathcal{T}_\pi$，Alice 完全了解 $\alpha(w \mid x)$，Bob 完全了解 $\beta(w \mid y)$。现在我们描述 Bob 模型和开放模型中的实际协议。

步骤 1. 对于每个 $w \in \mathcal{T}_\pi$，Alice 发送满足以下条件的最小非负整数 $d_w < \left\lceil \frac{1}{\gamma} \right\rceil$：

$$\gamma \cdot d_w \le \alpha(w \mid x) \le \gamma \cdot (d_w + 1).$$

这需要 $|\mathcal{T}_\pi| \cdot \left\lceil \log \frac{1}{\gamma} \right\rceil$ 的通信。

现在，对于所有满足以下条件的 w，Bob 都知道 Alice 的 $\alpha(w \mid x)$ 的一个近似值 $\alpha'(w \mid x)$：

$$\alpha'(w \mid x) \le \alpha(w \mid x) \le \alpha'(w \mid x) + \gamma.$$

由于对于所有 w，$p'(w \mid x, y) := \alpha'(w \mid x)\beta(w \mid y)$（Bob 已知）均满足 $\beta(w \mid y) \in [0, 1]$，因此：

$$\forall w \in \mathcal{T}_\pi : p'(w \mid x, y) \le p(w \mid x, y) \le p'(w \mid x, y) + \gamma.$$

也就是说，Bob 对 $p(.\mid x,\ y)$ 的真实概率有一个估计，该估计从不高估真实值，并且逐点与真实值相差不超过 γ。

步骤 2. Bob 不能直接输出 $p'(.\mid x,\ y)$，因为它可能不是一个概率分布。然而， $p(.\mid x,\ y)$ 是一个概率分布，因此：

$$1-\gamma|\mathcal{T}_\pi| \le \sum_w p'(w\mid x,y) \le 1.$$

定义 $C := 1-\sum_w p'(w\mid x,\ y)$ 和 $p''(w\mid x,\ y) = p'(w\mid x,\ y) + \frac{C}{|\mathcal{T}_\pi|}$ 对所有 w 都成立。由于$0 \le C \le \gamma|\mathcal{T}_\pi|$， $p''(.\mid x,\ y)$ 是一个分布，同时也是$p(.\mid x,\ y)$ 的逐点 γ-近似值。我们的 $\gamma = \delta|\mathcal{T}_\pi|^{-1}$ 选择使得 $p''(.\mid x,\ y)$ 与 $p(.\mid x,\ y)$ 的统计距离为 δ-近，因此 Bob 可以输出它。因此

$$D^{\mathsf{B}}(\mathbf{TDE}_{\Pi,\delta}) \le |\mathcal{T}_\pi| \cdot \left\lceil \log \frac{|\mathcal{T}_\pi|}{\delta} \right\rceil .$$

这就完成了引理6的证明6.33Lemma.。

□

我们将在局部模型和开放模型中证明一个类似的结论，并将其用于证明其他去随机化结果。

引理 F.1. 设 Π 是一个私有硬币通信协议， $\mathcal{T}_\pi$ 是其可能的转录集合。对于任意$0<\delta<\frac{1}{2}$、$D^{\mathsf{loc}}(\mathbf{TDE}_{\Pi,\delta}) \le D^{\mathsf{open}}(\mathbf{TDE}_{\Pi,\delta}) \le 2|\mathcal{T}_\pi| \cdot \left\lceil \log \frac{2|\mathcal{T}_\pi|}{\delta} \right\rceil$。 ——

请注意，在局部模型中，我们要求两个参与者都输出叶节点上分布的相同近似值。在引理6.33中，引理 。只需一个参与者发送估计值，另一个参与者使用其精确值，但在这里，第二个参与者也必须发回结果。因此，**TDE** 协议在局部模型中的通信复杂度略高于单边模型。

在局部模型中，运行**TDE**后，两个参与者对叶节点上的分布有相同的估计。每个参与者还知道自己在每个叶节点上的输出分布，因此双方的多数答案都很明确。在开放模型中，外部观察者公开计算**TDE**后的情况也类似。因此，例如，沿用[KN9797的KN引理 3。8],引理 F.11 引理 F.可推出定理 F.33定理 F.。

引理 *F.1* 的证明*1*引理 *F.*。令 $\gamma = \frac{\delta}{2}|\mathcal{T}_\pi|^{-1}$。我们按照引理 6 的证明6.33 引理 .进行，但将第二步替换为以下步骤。

步骤 2'。Bob 不在第一步之后直接输出，而是将 $p'(.\mid x,\ y)$ 的一个近似值发回给 Alice。更准确地说，对于所有 w，他发送 d'_w， $0 \le d'_w < \left\lceil \frac{1}{\gamma} \right\rceil$， 使得：

$$\gamma \cdot d'_w \le p'(w\mid x,y) \le \gamma\cdot(d'_w+1).$$

这同样需要通信 $|\mathcal{T}_\pi| \cdot \left\lceil \log(\frac{1}{\gamma}) \right\rceil$。因此，外部观察者知道 $p''(w\mid x,y) := \gamma \cdot d'_w$ 对所有 w 均成立，它满足：

$$\forall w \in \mathcal{T}_\pi : p''(w\mid x,y) \le p(w\mid x,y) \le p''(w\mid x,y) + 2\cdot\gamma.$$

我们定义 $C := 1-\sum_w p''(w\mid x,y)$ 和 $p'''(w\mid x,y) = p''(w\mid x,y) + \frac{C}{|\mathcal{T}_\pi|}$ 对所有 w 均成立。这个$p'''(.\mid x,y)$ 是一个分布，也是 $p(.\mid x,y)$ 的逐点近似值，并且可以由外部观察者计算。根据我们对 $\gamma = \frac{\delta}{2}|\mathcal{T}_\pi|^{-1}$ 的选择，我们得到了所需的输出，因此

$$D^{\mathsf{open}}(\mathbf{TDE}_{\Pi,\delta}) \le 2|\mathcal{T}_\pi| \cdot \left\lceil \log \frac{2|\mathcal{T}_\pi|}{\delta} \right\rceil,$$

至此，引理 F.1 的证明得以完成1引理 F..

□

F.2 随机性消除的证明细节（第6节6第）

以下引理将有助于证明我们的结果：

引理 F.2。 设 U 和 V 是分别取值于其定义域 $\mathcal{U}$ 和 $\mathcal{V}$ 上的随机变量。对于所有$u \in \mathcal{U}$，令我们考虑在事件 $[U = u]$ 条件下的随机变量 $V_{U=u}$ V。假设存在两个常数 δ_U 和 δ_V，以及两个分别取值于与 U 和 V 相同定义域上的随机变量 U' 和 V'，使得：

$$\Delta(U, U') \leq \delta_U \qquad \forall u \in \mathcal{U} : d_\infty(V_{U=u}, V'_{U'=u}) \leq \delta_V.$$

则：

$$d_\infty(V, V') \leq \delta_U + \delta_V.$$

引理 *F.2* 的证明2引理 *F.*。 我们来证明 $\forall v \in \mathcal{V}$， $|$概率$[V = v] -$ 概率$[V' = v]| \leq \delta_U + \delta_V$。固定任意一个 $v \in \mathcal{V}$，则概率 $\Pr[V = v]$ 和 $\Pr[V' = v]$ 可以写成：

- 概率$[V = v] = \sum_{u \in \mathcal{U}}$概率$[U = u] \cdot$概率$[V = v \mid U = u]$,
- 概率$[V' = v] = \sum_{u \in \mathcal{U}}$概率$[U' = u] \cdot$概率$[V' = v \mid U' = u]$。

因此，利用上述两个假设，我们得到：

$$\begin{aligned}
&\Pr[V = v] - \Pr[V' = v] \\
&= \sum_{u \in \mathcal{U}} \big(\Pr[U = u] \cdot \Pr[V = v \mid U = u] - \Pr[U' = u] \cdot \Pr[V' = v \mid U' = u]\big) \\
&\leq \sum_{u \in \mathcal{U}} \big(\big(\Pr[U = u] - \Pr[U' = u]\big) \Pr[V = v \mid U = u] + \delta_V \Pr[U' = u]\big) \\
&\leq \sum_{u \in \mathcal{U} : \Pr[U=u] > \Pr[U'=u]} \big(\Pr[U = u] - \Pr[U' = u]\big) + \delta_V \\
&\leq \delta_U + \delta_V.
\end{aligned}$$

按照相同的证明方法，我们可以证明 $\Pr[V = v] - \Pr[V' = v] \geq -(\delta_U + \delta_V)$；将两者结合，便得到所需的结果：

$$\forall v \in \mathcal{V} : \big|\Pr[V = v] - \Pr[V' = v]\big| \leq \delta_U + \delta_V.$$

□

F.3 异或模型下的去随机化

开放模型和局部模型是对定理 66.11 定理的直接改编。

定理 F.3. 对于任意函数 f、错误 $\epsilon < \frac{1}{2}$ 和模型 $\mathcal{M} \in \{$开放、局部$\}$，若满足 $R^{\mathcal{M}} = {}_{R^{\mathcal{M}}, \epsilon}^{\text{priv}}(f)$：

$$D^{\mathcal{M}}(f) \leq 2 \cdot 2^{R^{\mathcal{M}}} \left(R^{\mathcal{M}} + \log\left(\frac{1}{\frac{1}{2} - \epsilon}\right) + 2\right)$$

F.3.1 二选一模型中的去随机化

有趣的是，在二选一模型中，$\epsilon = \frac{1}{3}$ 处存在一个用于去随机化的错误阈值。如果错误低于此阈值，只需解决相应的**TDE** 实例，之后其中一位参与者便知道多数结果。当错误接近 $1/2$时，多数结果可能有多个候选项，需要额外通信 $O(k)$。在这种情况下，我们通过使用 NBA问题的一种变体，将这一项降至 $O(\log(k))$。

定理 F.4. 对于任意函数 f和错误 $\epsilon < \frac{1}{2}$，若满足 $R = R^{\mathsf{priv}}_{\epsilon}(f)$：

$$D^{1\text{of}2}(f) \leq \begin{cases} 2^{R+1}\left(R + \log\left(\frac{4}{\frac{1}{3}-\epsilon}\right) + 1\right), & \text{if } \epsilon < \frac{1}{3}, \\ \left(2^{R+1}+2\right) \cdot \left(R + \log\left(\frac{8}{\frac{1}{2}-\epsilon}\right) + 1\right) + \log(k) + 4, & \text{for any } \epsilon < \frac{1}{2}. \end{cases}$$

定理 *F.4* 的证明4定理 *F.*。 取 Π为针对 f 的一个最优私有硬币一取二协议，其错误为 ϵ。令 σ 为一个精度参数，我们将在后面设定它。

当 $\epsilon < \frac{1}{3}$ 时，注意到其中一名参与者必须以大于$1\,3$的概率输出正确结果，而所有错误结果的输出概率均小于$1\,3$（带有额外的微小偏差）。因此，参与者只需针对1 引理 F.$\mathbf{TDE}_{\pi,\,\sigma}$运行引理 F.1 的局部协议，其中此处 $\sigma < \frac{1}{3} - \epsilon$，并让以大于$1\,3$ 的概率输出某个结果的参与者输出该结果。

现在我们转向更有趣的情形，即$1/3 \leq \epsilon < \frac{1}{2}$。令 $\delta = \frac{1}{2} - \epsilon$且 $\sigma < \frac{\delta}{3}$。参与者首先运行针对$\mathbf{TDE}_{\pi,\,\sigma}$的局部协议，从而获知该协议每个通信记录的概率的一个 σ 近似值。由引理 F.22，引理 F.由于每位参与者都确切知道她在每个叶节点的输出分布，对于所有 z，每位参与者都知道她在原始协议中输出 z 的概率，精度达到 σ 。

我们将 p^z_{A} 称为 Alice 输出 z 的概率，将 $\widetilde{p}^z_{\mathsf{A}}$ 称为她对该概率的近似值。对于 $z = f(x,\ y)$，我们有 $p^z_{\mathsf{A}} + p^z_{\mathsf{B}} \geq \frac{1}{2} + \delta$，因此 $\widetilde{p}^z_{\mathsf{A}} + \widetilde{p}^z_{\mathsf{B}} \geq \frac{1}{2} + \delta - \sigma$。

利用这一点，参与者将一些 z视为 $f(x,\ y)$ 的候选项 。Alice 将满足 $\widetilde{p}^z_{\mathsf{A}} \geq \frac{1}{4} + \frac{\delta-\sigma}{2}$ 的答案 n_{A} 视为 $(z_i^{\mathsf{A}})_{i\in[n_{\mathsf{A}}]}$。类似地，Bob 将满足 $\widetilde{p}^z_{\mathsf{B}} \geq \frac{1}{4} + \frac{\delta-\sigma}{2}$ 的答案 n_{B} 视为 $(z_j^{\mathsf{B}})_{j\in[n_{\mathsf{B}}]}$ 。

由于 $\sum_z \widetilde{p}^z_{\mathsf{A}} + \widetilde{p}^z_{\mathsf{B}} = 1$ （其中求和遍及所有 $z \in \mathcal{Z}$），我们有： $n_{\mathsf{A}} + n_{\mathsf{B}} \leq 3$。由于多数输出所代表的输出严格多于所有输出的一半，因此 $\max(n_{\mathsf{A}},\ n_{\mathsf{B}}) \leq 2$。

参与者使用 4 个比特相互发送值 $n_{\mathsf{A}},\ n_{\mathsf{B}}$ 。不失一般性，假设 $n_{\mathsf{A}} \geq n_{\mathsf{B}}$。则有以下四种情况：

1. $(n_{\mathsf{A}},\ n_{\mathsf{B}}) = (1,\ 0)$ 2. $(n_{\mathsf{A}},\ n_{\mathsf{B}}) = (2,\ 1)$ 3. $(n_{\mathsf{A}},\ n_{\mathsf{B}}) = (2,\ 0)$ 4. $(n_{\mathsf{A}},\ n_{\mathsf{B}}) = (1,\ 1)$。

前两种情况很简单：如果只有一个候选项（情况 11case），拥有该候选项的参与者就输出它。如果有三个候选项（情况 22case），拥有单个候选项的参与者知道它必须与另一方的某个候选项相同，并且是多数输出，因此输出该候选项。

对于剩下的两种情况，我们将使用 NBA问题协议的一个变体。对于情况 $(n_{\mathsf{A}},\ n_{\mathsf{B}}) = (2,\ 0)$，Alice（拥有两个候选项）向 Bob 发送两个候选项不同的某个比特的索引，例如 $i \in [\lceil \log(\mathcal{Z}) \rceil]$。Bob 回复 $\sum_{z:\ z_i=0} \widetilde{p}^z_{\mathsf{B}}$。因此 Alice 可以计算 $\sum_{z:\ z_i=0} \widetilde{p}^z_{\mathsf{B}} + \widetilde{p}^z_{\mathsf{A}}$。如果该数值大于 $1\,2$，则正确候选项是其第 i 个比特为 0 的候选项；否则，就是另一个候选项。

最后，我们考虑情形（$n_{\mathsf{A}},\ n_{\mathsf{B}}$） $= (1,\ 1)$。不失一般性，假设 Alice 的候选项 z_1^{A} 不正确，即 $z_1^{\mathsf{A}} \neq f(x,\ y) = z_1^{\mathsf{B}}$。然后我们注意到，Alice 输出 z_1^{A} 的概率以及 Bob 输出不同于 z_1^{B} 的概率都小于 $\epsilon = \frac{1}{2} - \delta$。为完成该协议，参与者将 $\widetilde{p}^{z_1^{\mathsf{A}}}{}_{\mathsf{A}}$ 和 $\widetilde{p}^{z_1^{\mathsf{B}}}{}_{\mathsf{B}}$ 交换到 σ 精度。于是：

- $p^{z_1^{\mathsf{B}}}_{\mathsf{B}} + p^{\top}_{\mathsf{B}} - p^{z_1^{\mathsf{A}}}_{\mathsf{A}} = p^{z_1^{\mathsf{B}}}_{\mathsf{B}} + \sum_{z\neq z_1^{\mathsf{A}}} p^z_{\mathsf{A}} \geq p^{z_1^{\mathsf{B}}}_{\mathsf{B}} + p^{z_1^{\mathsf{B}}}_{\mathsf{A}} \geq \frac{1}{2} + \delta,$
- $p^{z_1^{\mathsf{A}}}_{\mathsf{A}} + p^{\top}_{\mathsf{A}} - p^{z_1^{\mathsf{B}}}_{\mathsf{B}} = p^{z_1^{\mathsf{A}}}_{\mathsf{A}} + \sum_{z\neq z_1^{\mathsf{B}}} p^z_{\mathsf{B}} \leq 1 - p^{z_1^{\mathsf{B}}}_{\mathsf{A}} + p^{z_1^{\mathsf{B}}}_{\mathsf{B}} \leq \frac{1}{2} - \delta.$

每位参与者都能近似计算自己一侧输出概率之和，并能近似计算另一位参与者一侧候选输出的概率，因此他们可以近似计算上述各项之和。由于 $\sigma < \frac{\delta}{3}$，参与者可以确定自己是否得到了正确的输出。如果输出不正确，他们就让另一位参与者输出。

□

F.3.2 分割模型中的去随机化

分割模型中的去随机化可以通过与先前研究的模型中的去随机化类似的方法实现。

定理 F.5。令 $0 < \epsilon < 1/2$ 和 $f\colon \mathcal{X} \times \mathcal{Y} \to \mathcal{Z} = \{0,1\}^k$。令 $R = R^{\mathsf{spl},\ \epsilon\ \mathsf{priv}}(f)$，$M = 16 \cdot \left(\frac{1}{2} - \epsilon\right)^{-2} \cdot 2^R$。则：

$$D^{\mathsf{spl}}(f) \leq \begin{cases} 2^{R+1} \cdot \left(R + \log\left(\frac{4}{\frac{1}{3}-\epsilon}\right) + 1\right) + k, & \text{if } \epsilon < \frac{1}{3}, \\ 2^{R+1} \cdot \left(R + \log\left(\frac{8}{\frac{1}{2}-\epsilon}\right) + 1\right) + k \cdot \left(\frac{5-2\epsilon}{4} M + 1\right), & \forall \epsilon < \frac{1}{2}. \end{cases}$$

定理 *F.* 的证明 *5* $\epsilon < \frac{1}{3}$ *4,* $z \in \{0,1,*\}^k$ z 定理 *F.*。 **当……时的情形** 与定理 F. 的证明一样，定理 F.对于每个字符串，参与者通过求解一个 **TDE** 实例来估计其输出的概率。对于输出的每个索引，在随机化协议中，其中一位参与者必须以至少 $\frac{1-\epsilon}{2} > \frac{1}{3}$ 的概率输出正确的比特。Alice 向 Bob 发送 k 比特，指明在原始协议中，她以大于 $1\ 3$ 的概率输出相同的非-$*$ 符号的那些比特。在去随机化协议中，她输出这些比特，而 Bob 负责输出其余比特。对于他们输出的每个比特，他们都输出在原始随机化协议中出现频率最高的值。

情形为 $\epsilon < \frac{1}{2}$ 此情形与我们在定理 6 的证明中看到的类似6.44定理 。：我们构造一个 **GapMAJ** 与 weave（xx）小工具复合而成的实例，方式与我们构造**GapMAJ∘异或** 实例以对异或模型中的协议进行去随机化的方式相同。

□

与异或模型中一样，通过改进分割模型中 GapMAJ 的正确小工具化版本的确定性复杂度，可以改进这些界。

G GapMAJ∘异或界的证明

本节证明表 2 中关于表 2GapMAJ**GapMAJ∘异或**的通信复杂度的陈述。

回顾一下，**GapMAJ∘异或** 问题由四个参数确定：N 行数，k Alice 和 Bob 的行长度，ϵ不异或为隐藏的 k 位字符串 z 的行所占比例，以及 μ 行上的一个分布。一个 **GapMAJ∘异或** 实例可以这样理解：Alice 和 Bob 各自拥有一个 $N \times k$ 布尔矩阵，其中两矩阵按位异或中包含 z 的行的集合，其权重相对于 μ 高于 $1 - \epsilon$。下文中，ϵ' 是目标概率，即我们在求解 **GapMAJ∘异或** 实例时所能容忍的最大错误率。

当 $\epsilon \leq \epsilon'$ 时，在异或模型中，使用公共硬币选择一行并输出该行的平凡协议就足够了。因此我们得到以下结果。

命题 G.1。对于所有 N，k，ϵ，ϵ'，μ，满足 $0 \leq \epsilon \leq \epsilon'$，

$$R_{\epsilon'}^{\mathsf{xor,pub}}(\mathbf{GapMAJ}_{N,k,\epsilon,\mu} \circ \mathbf{XOR}) = 0.$$

由此命题可得以下上界。

推论 G.2。对于所有 N、k、ϵ、ϵ'、μ，满足 $0 \leq \epsilon \leq \epsilon'$

- $R_{\text{异或，私有}\,\epsilon'}$（$\mathbf{GapMAJ}_{N,\,k,\,\epsilon,\,\mu}$∘**异或**）$\leq \log(N)$，
- $R_{\text{开放，公开}\,\epsilon'}$（$\mathbf{GapMAJ}_{N,\,k,\,\epsilon,\,\mu}$∘**异或**）$\leq 2k$，
- $R_{\text{开放，私有}\,\epsilon'}$（$\mathbf{GapMAJ}_{N,\,k,\,\epsilon,\,\mu}$∘**异或**）$\leq 2k + \log(N)$，
- $D_{\text{统一}}$（$\mathbf{GapMAJ}_{N,\,k,\,\epsilon,\,\mu}$∘**异或**）$\leq$ $(2\epsilon N + 1)\,k$。

推论 *G.2* 的证明2推论 *G.*。 关键在于注意到，$\mathbf{GapMAJ}$∘**异或** 在 $\epsilon \leq \epsilon'$（命题 G.11命题 G.）时是平凡的。对于私有随机币，注意到参与者只使用 $\log(N)$ 个随机币，因此让 Alice（不失一般性）将她的随机币发送给 Bob 就足够了。对于开放协议，参与者可以交换他们的行。对于确定性协议，Alice（不失一般性）将她的 ）个最重的行（μ-wise）发送给 Bob，随后 Bob 便可以计算出现频率最高的 z，即他的行与 Alice 的行进行异或所得的对象。 □

当 $\epsilon > \epsilon'$时，我们引用第 5 节中的早期结果第 5 节 (定理 55.66定理 .)。

H 其他分离问题

在本附录中，我们给出图 1 和表 1 中分离问题的定义，并证明它们的复杂度取决于输出模型。这些问题都是常见问题的变体，额外对输入施加了一项约束：每位参与者的 n 比特输入中，最多有 t 比特为 1。记 $B_2(n,t) = \{x \in \{0,1\}^n$ 为以 0^n 为中心、半径为 t 的 $\{0,1\}^n$ 中的汉明球；记 $H(x) = -x\log(x) - (1-x)\log(1-x)$ 为期望值为 x 的伯努利随机变量的熵；并回顾其大小的以下界，记为 $V_2(n,t) = |B_2(n,t)|$:

引理 H.1（第 10 章，推论 9，见 [MS83]MS）。 设 $0 < t < n/2$。则：

$$\frac{1}{\sqrt{8t(1-t/n)}} 2^{n\cdot H(t/n)} \leq V_2(n,t) \leq 2^{n\cdot H(t/n)}.$$

在下文中，我们将考虑 $t \in o(n)$，并且只在以下情形中使用它：

$$\log(V_2(n,t)) \in \Omega(t \cdot \log(n)).$$

H.1 t-交集

由于不相交性是一个布尔问题，它无法区分我们的通信模型。然而，对于它的大输出变体——交集问题，情况并非如此；在该问题中，Alice 和 Bob 必须计算各自集合的实际交集。

我们回顾问题 $t - \mathbf{DISJ}_n$ 和 $t - \mathbf{INT}_n$ 的定义及其复杂度的已知结果，并证明 t –$\mathbf{INT}_n$ 能够区分局部模型与开放模型。

定义 H.2（t-不相交性问题）。 $t - \mathbf{DISJ}_n$： $B_2(n,\ t) \times B_2(n,\ t) \to \{0,1\}$ 定义如下：

$$t - \mathbf{DISJ}_n(X,Y) = \mathbf{1}_{X\cap Y=\emptyset}.$$

现在，我们定义这个问题的一种自然变体，其输出规模较大。

定义 H.3（t-交集问题）。 $t - \mathbf{INT}_n$： $B_2(n,\ t) \times B_2(n,\ t) \to B_2(n,\ t)$ 定义如下：

$$t - \mathbf{INT}_n(X,Y) = X \cap Y.$$

由于 $t-\mathbf{DISJ}_n$ 的输出是布尔值，其各种通信复杂度除一个比特外基本相同，因此在下面的陈述中无需指定通信模型：

定理 H.4. $R_\epsilon(t-\mathbf{DISJ}_n) = \Theta(t)$。

$\Omega(t)$ 下界直接来自 $\mathbf{DISJ}_n$ 在 [KS9292、KSRaz9292、RazBYJKS04], BYJKS 中的 $\Omega(n)$ 下界，而 $O(k)$ 上界则由 [HW07]HW证明。

定理 H.5. $R_\epsilon^{\mathsf{loc}}(t-\mathbf{INT}_n) = \Theta(t)$，以及 $R_\epsilon^{\mathsf{open}}(t-\mathbf{INT}_n) = \Theta(t\cdot\log(n))$。

该问题的 $O(t)$ 上界由 [BCK+14] 证明，而开放模型中的 $\Omega(t\cdot\log(n))$ 下界直接来自输出的规模（见附录 C附录 C 和引理 H.11Lemma H.），因为 $|B_2(n,\ t)| = V_2(n,\ t) \in \Omega(t\cdot\log(n))$（当 $t \in o(n)$ 时）。

H.2 t-查找第一个差异

正如交集可以看作不相交性问题的大输出变体，查找第一个差异也可以看作大于问题的大输出变体。

现在我们定义问题 $t-\mathbf{GT}_n$ 和 $t-\mathbf{INT}_n$，介绍关于它们复杂度的已知结果，并证明 $t-\mathbf{FtFD}_n$ 能够区分二选一模型和单边模型。

定义 H.6（t-大于问题）。 $t-\mathbf{GT}_n$： $B_2(n,t) \times B_2(n,t) \to \{0,1\}$ 定义如下：

$$t-\mathbf{GT}_n(x,y) = \mathbf{1}_{x>y}.$$

定义 H.7（t-寻找首个差异问题）。 $t-\mathbf{FtFD}_n$： $B_2(n,t) \times B_2(n,t) \to \{0,\dots,n\}$ 定义如下：

$$t-\mathbf{FtFD}_n(x,y) = \min(\{i : x_i \neq y_i\} \cup \{n\}).$$

定理 H.8。

$$R_\epsilon^{\mathsf{1of2}}(t-\mathbf{FtFD}_n) \in O\left(\log(t)+\log(\log(n))+\log\left(\frac{1}{\epsilon}\right)\right),$$
$$R_\epsilon^{\mathsf{uni}}(t-\mathbf{FtFD}_n) \in \Omega(\log(n)).$$

定理 *H.8* 的证明定理 *H.8*。

关于 $R_\epsilon^{\mathsf{1of2}}(t-\mathbf{FtFD}_n)$ 的上界。 作为直观说明，我们先给出在$t=1$情况下的一个协议。

在这种情况下，Alice 和 Bob 的 n 比特字符串 x 和 y 各自只包含一个 1。因此，考虑i^{A}，i^{B} 使得 $x_{i^{\mathsf{A}}} = 1$ 且 $y_{i^{\mathsf{B}}} = 1$。 i^{A}，$i^{\mathsf{B}} \in [n]$，因此它们可以写成两个$\lceil\log n\rceil$比特字符串。

参与者随后运行 Feige 等人的协议 [FRPU94]FRPU，以查找 i^{A} 和 i^{B}之间的首个差异。通过这样做，他们可以得知满足以下条件的最小 t $(i^{\mathsf{A}})_k \neq (i^{\mathsf{B}})_k$（或$\lceil\log(n)\rceil+1$，如果不存在），从而知道 $i^{\mathsf{A}} < i^{\mathsf{B}}$，$i^{\mathsf{A}} > i^{\mathsf{B}}$ 还是 $i^{\mathsf{A}} = i^{\mathsf{B}}$。编号最小的参与者因此知道 x和 y之间首个差异的索引，因为它是 $\min(i^{\mathsf{A}},\ i^{\mathsf{B}})$。

现在考虑 t 不受约束的情况。为了找到它们两个 n位、权重为 $\le t$的比特串之间的首个差异，两个参与者只需构造一个由其值为 1 的比特索引组成的 $\Omega(t\cdot\log(n))$-比特串（并进行适当填充），然后使用 Feige 等人 [FRPU94] FRPU协议，正如在 $t=1$ 情形中那样。更准确地说：

- 设 $w_x = |x| \leq t$ （分别为 $w_y = |y| \leq t$）是 x （分别为 y）的权重。现在，考虑索引 $i_{\mathsf{A}1}, \ldots, i_{\mathsf{A}t}$ 和 $i_{\mathsf{B}1}, \ldots, i_{\mathsf{B}t}$，它们位于 $\{0, \ldots, n-1\} \cup \{2^{\lceil \log(n+1) \rceil} - 1\}$ 中，满足：
  - $i_j^{\mathsf{A}} = 2^{\lceil \log(n+1) \rceil} - 1$ (an all-1 string) iff $j > w_x$
  - $x_{i_j^{\mathsf{A}}} = 1, \forall j <= t$
  - $i_j^{\mathsf{A}} < i_{j+1}^{\mathsf{A}}, \forall j < t$
    (and similarly for the i_j^{B}'s)

  每个 i_j^{A}都可以用 $\lceil \log(n+1) \rceil$ 比特表示，因此 Alice 会计算一个由所有 i_j^{A} 按顺序拼接而成的 $t\lceil \log n \rceil$ 位比特串s_x 。Bob 也会以类似方式计算s_y 。
  然后，两位参与者使用 Feige 等人的协议，找出 s_x 和 s_y之间的首个差异。我们将此差异的索引记为 i_{diff} 。
  然后，如果 $(s_x)_{i_{\mathsf{diff}}} = 0$，Alice 就知道首个差异的索引；否则，Bob 知道。确实，让我们考虑第一种情况：
  - Alice 在该索引处对应的值为 0，这意味着 s_x 的这一部分对应于原始 n 位字符串 x 中值为 1 的位置，因为我们会在末尾用 1 补齐。– 这个位置就是 Alice 有而 Bob 没有的最靠左的 1 所在的索引。事实上，在 i_{diff} 所属的位置之前，所有位置都是 Alice 和 Bob 共有的。因此，如果 Bob 在 i_{diff} 出现的位置也有一个 1，那么 Alice 和 Bob 在 i_{diff} 中找到差异这一事实就意味着 Bob 在更靠前的位置也有一个 1，这与 s_x 和 s_y 之间的首个差异恰好是 Alice 在该位置的值为 0 相矛盾。

  使用 Feige 等人的协议处理一个 $O(t \cdot \log(n))$ 比特的字符串，代价为 $O\left(\log \binom{t \cdot \log(n)}{\epsilon}\right)$，因此得到所述上界。

关于 $R_\epsilon^{\mathsf{uni}}(t$ –查找第一个差异$_n)$ 的下界。不妨设 Alice 是输出结果的参与者，并考虑这样的输入：Alice 总是收到全为 0 n 的比特串，而 Bob 收到一个随机的 n 比特串，其中只有一个比特为 1。在这类实例上求解查找第一个差异问题，可以让 Bob 以 R_ϵ^{A}**(查找第一个差异**$_n$**)** 的通信量，以高概率向 Alice 发送大小为 $\log(n)$ 比特的信息，因此得到 $\Omega(\log(n))$ 的下界。

□

请注意，我们的二选一去随机化定理（4定理 F.）表明，对于私有随机硬币通信复杂度，我们的上界是紧的；但在二选一模型或异或模型中，仍可能存在更高效的公共随机硬币协议。现在我们证明，Viola 的 $\Omega(\log(n))$ 公共随机硬币随机化下界 [Vio15]Vio对于 $\mathbf{GT}_n$ 意味着，即使可以使用公共硬币，该协议的上界也是紧的。

定理 H.9。

$$R_\epsilon(t - \mathbf{GT}_n) \in \Omega(\log(t) + \log\log(n))$$

并且作为推论，$R_\epsilon^{\mathsf{xor}}(t - \mathbf{FtFD}_n) \in \Omega(\log(t) + \log\log(n))$ 。

定理 *H.9* 的证明9定理 *H.*。 我们分别证明对 $\log(t)$ 和 $\log\log(n)$ 的依赖关系。

$R_\epsilon(t$ –$\mathbf{GT}_n) \in \Omega(\log(t))$。 我们指出，$\mathbf{GT}_t$ 可按与上一节中 t–$\mathbf{GT}_n$ 从 $\mathbf{DISJ}_t$归约到 t–$\mathbf{DISJ}_n$ 相同的方式进行归约，因此应用 Viola 下界 [Vio15] Vio可得：

$$R_\epsilon(t - \mathbf{GT}_n) \geq R_\epsilon(\mathbf{GT}_t) \in \Omega(\log(t))$$

$R_\epsilon(t-\mathbf{GT}_n) \in \Omega(\log\log(n))$。我们指出，$1-\mathbf{GT}_n$ 可归约到 $\mathbf{GT}^{\log(n)}$，因为比较两个二进制表示中恰有一个比特为 1 的数，相当于比较它们唯一的 1 所在位置的索引。因此，再次应用 Viola 下界[Vio15] Vio可得：

$$R_\epsilon(t-\mathbf{GT}_n) \geq R_\epsilon(1-\mathbf{GT}_n) \geq R_\epsilon(\mathbf{GT}_{\log(n)}) \in \Omega(\log\log(n))$$

□

H.3 最大值问题

定义 H.10（最大值问题）。$\mathbf{MAX}_n$：$\{0,1\}^n \times \{0,1\}^n \to \{0,1\}^n$ 定义为

$$\mathbf{MAX}_n(x,y) = \begin{cases} x, & if\ x \geq y, \\ y, & otherwise. \end{cases}$$

对于这个问题，我们有：

定理 H.11。

$$R_\epsilon^{\mathsf{1of2}}(\mathbf{MAX}_n) \in O(\log n), \qquad R_\epsilon^{\mathsf{uni}}(\mathbf{MAX}_n) \in \Omega(n).$$

当仅允许使用私有随机币时，该间隔保持不变（渐近意义下，忽略乘法常数和加法常数）。

定理 *H.11* 的证明*11定理 H.*。

R^{1of2} **上界**。参与者以高概率计算是否 $x \leq y$，使用$O(\log n)$通信；随后，如果 $x \leq y$，Alice 输出 x，否则 Bob 输出 y。

R^{uni} **下界**。只需证明 R^{A} 很大，因为根据对称性可知R^{B} 也同样很大。

证明相当简单：考虑 2^n 输入对 $\{(0,y) : y \in [0,2^n-1]\}$。对于这些输入，**最大值**$_n$ 问题只是一个单向通信问题：显然，他必须发送 $\Omega(n)$ 比特，Alice 才能以 $\geq 1-\epsilon$的概率正确猜出他的 y。

□