arXiv:2305.01070v1 [cs.DS] 1 May 2023

匹配的鲁棒通信复杂度：EDCS 达到 5/6 的近似比

AmirAzarmehr Soheil Behnezhad东北大学 东北大学

摘要

我们研究最大匹配的鲁棒通信复杂度。一个含有 n 个顶点的任意图 G 的边，会以随机且均匀的方式独立分配给 Alice 和 Bob。Alice 必须向 Bob 发送一条消息，使 Bob 能够找出整个图 G 的一个（近似）最大匹配。我们具体研究这样一种协议所能达到的最佳近似比 ‹，在该协议中 Alice 只向 Bob 通信 $O(n)$ 位。

由于鲁棒通信模型与随机顺序流模型有关，人们对该模型的兴趣日益浓厚。Assadi 和 Behnezhad 的一项算法 [ICALP'21] 表明，当 $0 < \varepsilon_0 < 10^{-18}$ 为一个较小的常数时，可以实现 $(2/3 + \varepsilon_0 \sim .667)$ 的近似比，这仍是一般图中已知的最佳近似结果。对于二分图，Assadi 和 Behnezhad[Random'21] 将近似比提升至 0.716，但其协议的计算效率较低（即需要指数级时间）。

本文研究 Bernstein 在 [ICALP'20] 提出的一种随机顺序流处理算法所隐含的自然且高效的协议。该算法基于边度约束子图（EDCS）[Bernstein 和 Stein；ICALP'15]。Bernstein 的结果立即表明，该协议在鲁棒通信模型中可实现几乎（$2/3 \sim .666$）的近似比。我们提出了新的分析，证明该协议可实现远好于此的近乎（$5/6 \sim .833$）近似比。这显著提升了先前针对一般图和二分图的近似结果。我们还证明了对 Bernstein 协议的分析是紧的。

1 引言

给定一个 n顶点图 $G = (V, E)$，一个 匹配 是由 G 中的一组两两不共享顶点的边构成的，而最大匹配是边数最多的匹配。本文研究 Yao 的（单向）通信模型中的匹配问题 [13]。边集 E 由 Alice 和 Bob 两位参与者分割。Alice 必须向 Bob 发送一条消息，使 Bob 能够找到整个图 G的（近似）最大匹配。我们尤其关注 Alice 发送的消息大小与输出解近似比之间的权衡。除了这是一个自然的问题之外，该通信模型还与流处理算法密切相关，因此多年来一直受到广泛研究 [10, 12, 8, 1, 2]。

众所周知，要获得精确的 最大匹配，需要 $\Omega(n^2)$ 比特的通信 [9]。也就是说，Alice 将自己的全部输入发送给 Bob 的朴素协议是最优的。近似解的情况则更有意思。显然，任何近似算法都需要 $\Omega(n)$ 个通信字数，因为可以将整个匹配交给 Alice。因此，许多先前研究 [10, 11, 12, 1, 2], 探讨了一个自然的问题：在通信复杂度近乎最优、为 $O(n) = O(n\text{poly log})$ 的协议下，能够实现的最佳近似比是多少。

不难看出，如果 Alice 将其输入图的一个最大匹配发送给 Bob，那么 Bob 就能找到一个 1/2 近似匹配。然而，Bernstein 和 Stein [6] 提出了一种更精巧的方法，它基于强大的边度约束子图（*EDCS*） [6] ，能够达到（几乎）2/3 的近似比（参见 Assadi 和 Bernstein 的论文 [3]）。事实证明，在对边进行对抗性划分的情况下，这一近似比是最优的。在他们的开创性论文中，Goel、Kapralov 和 Khanna [10] 证明，要获得优于 2/3 的近似比，需要$_{n^{1+1/(\log 对数 n)} \gg n \text{poly log } n}$ 通信。

上述通信模型是双重最坏情况模型，因为输入图和边的划分均由对手选择。本文研究所谓的鲁棒通信模型——`遵循 Chakrabarti、Cormode 和 McGregor [8]的做法——在该模型中，图 G 仍由对手选择，但其边现在随机地分配给 Alice 和 Bob（即每条边都独立且等概率地分给 Alice 或 Bob ）。该模型超越了上述双重最坏情况的情形，并有助于揭示一个问题的难度究竟源于输入图本身，还是源于其边的一种病态划分。研究鲁棒通信模型的另一个动机，是它与随机顺序数据流之间的联系。具体而言，几乎所有已知的随机顺序数据流下界都是在这一鲁棒通信模型中证明的。

尽管现有的对抗式划分协议已经意味着在鲁棒通信模型中可以达到（近似）2/3 近似率，但 Assadi 和 Behnezhad[1] 提出的随机顺序流处理算法给出了更好的界限。他们的算法以 Bernstein [5],提出的基于 EDCS 的算法为基础，再加入若干条较短的增广路径，从而对某个固定常数 $0 < \varepsilon_0 < 10^{-18}$实现 $(2/3+\varepsilon_0)$ 近似率。这仍是一般图中已知的最佳近似率。对于二分图，Assadi 和 Behnezhad [2] 采用了完全不同的方法，达到了更高的 0.716 近似率，尽管他们的协议需要双指数时间。

本文对 Bernstein 的基于 EDCS 的协议进行了新的分析， [5] 表明即使不进行任何增强，该协议也已实现远优于 2/3 的近似比。

定理 1. *Bernstein* 协议 *[5]* 以高概率实现 $(1-\varepsilon)5/6 \sim .833$ 近似，在鲁棒通信模型中使用 $O(n \cdot \log n \cdot \text{poly}(1/\varepsilon))$ 通信字数。

定理 1将一般图和二分图的最先进的近似比分别从 .667 [1]和 .716 [2]提升至 。833。我们注意到，Bernstein 协议的运行时间相对于输入规模是线性的；因此，定理 1除了提高近似比之外，还将 [2] 算法的运行时间从双指数级降至线性级。除了这些定量上的改进外，我们认为，定理 1更重要的定性意义在于，EDCS 在文献中曾被用于各种模型下仅获得 2/3或略高于 2/3 的近似比，而如今可以用于在鲁棒通信模型中获得显著更优的近似比。

我们的分析可以应用于更一般的多方单向鲁棒通信模型。在该模型中，输入不是在两位玩家 Alice 和 Bob 之间划分，而是随机划分给k 位参与方（模型的正式定义见第 3 节）。这个通信模型尤其值得关注，因为其中任何针对任意 k的下界，也意味着随机顺序数据流的下界。我们给出以下结果，它推广了定理 1：

定理 2。对于任意 $k \geq 2$ 和任意 $\varepsilon > 0$，*Bernstein* 协议 *[5]* 在 k方单向鲁棒通信模型中，使用长度为$(1-\varepsilon)(\frac{2}{3}+\frac{1}{3k})$的消息，实现最大匹配的 $O(\overline{n}\cdot \log n \cdot \mathrm{poly}(1/\varepsilon))$近似。

我们注意到，在随机顺序数据流设置中，目前已知的最大匹配最佳近似比为（$2/3+10^{-18}$），由 Assadi 和 Behnezhad 提出 [1]。定理 2意味着，要么存在更好的最大匹配随机顺序流处理算法（很可能如此），要么若要通过多方通信模型证明紧下界，就必须考虑至少 $k \geq 10^{18}/3$ 位参与方！

最后，我们证明我们对定理 1 和 2 的保证对于 Bernstein 协议是紧的。也就是说，我们证明：

定理 3。对于任意 $k \geq 2$，存在一个无限图族 G，使得 *Bernstein* 协议在 k 方单向鲁棒通信模型中的期望近似比至多为 $(\frac{2}{3}+\frac{1}{3k})$。

2 技术概述

Bernstein 协议构造了两个大小为 $O(n\log\ n)$ 的子图 H和 U，并将二者都传递给 Bob。子图 H 完全由 Alice 构造，她仅通过公开其输入图中 ε 的比例来完成这一过程。该构造保证，对于某个充分大的常数$\beta \geq 1$，每条边 $(u,v) \in H$都满足 $\deg_H(u) + \deg_H(v) \leq \beta$。也就是说， H的边度上界为 β。这已经意味着， H 至多包含 $O(n\beta) = O(n)$ 条边。子图 U 则简单地定义为图 G （提供给 Alice 或 Bob）中所有剩余边 (u,v) 的集合，这些边满足 $\deg_H(u) + \deg_H(v) \leq \beta - 1$。换言之，所有剩余的“未充分填充”边（其边度小于 β ）都会被加入 U。尽管目前尚不清楚能否以保证 $|U| = O(n\log\ n)$ 的方式构造 H，但 Bernstein [5] 证明了这确实是可能的。最后，Bob 返回其收到的所有边所构成的最大匹配。

可以证明，子图 $H \cup U$ 包含一个边度受限子图 (EDCS)，该子图 G 已知包含基图 G for$\beta \geq 1/\varepsilon$ [6, 4] 的 $(2/3 - O(\varepsilon))$- 近似最大匹配。这已意味着，在我们的模型中可达到（近似） 2/3 的近似比。这一保证

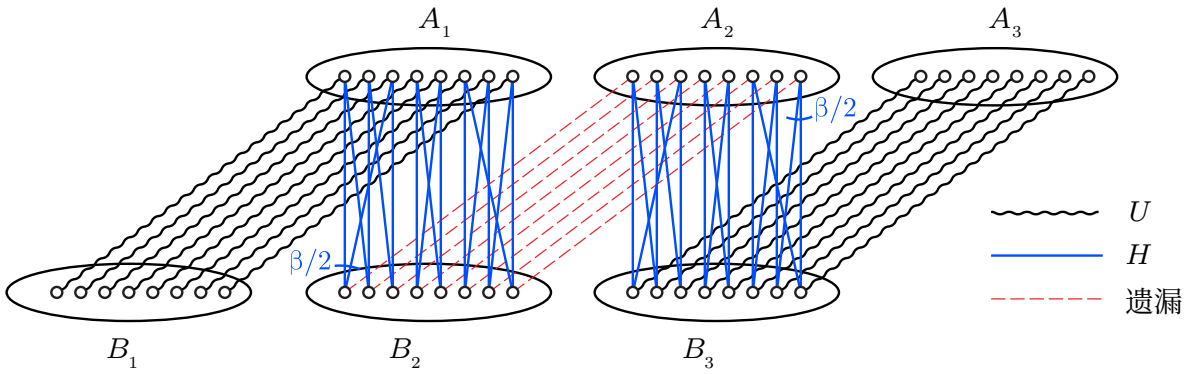


图 1：一个例子，其中 Bernstein 协议中的子图 $H \cup U$ 不包含优于 2/3 近似的匹配。

事实上，对于 $H \cup U$ 中包含的最大匹配而言，这一界是紧的，如 图 1 所示。在 图 1 的示例中，未匹配的（红色虚线）边在 H 中的边度为 β，因此不属于 U。尽管 图 1 中的示例图具有完美匹配，但 $H \cup U$ 中的任何匹配最多只能匹配 2/3 的顶点。

关键在于，尽管 $H \cup U$ 可能只包含图的一个 2/3 近似匹配，但 Bob 还可以访问集合 E_B，其中包含最初在随机划分中分配给他的边。因此，我们不应关注 $H \cup U$，而应关注 $H \cup U \cup E_B$ 中所包含的最大匹配的大小。现在让我们重新审视图 1中的例子。正如我们所讨论的，集合 H 仅利用少量 ε 边构造而成。此外，在给定 H的条件下，无论边如何在 Alice 和 Bob 之间划分，子图 U 都完全确定。这意味着，即使在给定 和 H 的条件下，U每条虚线边被分配给 Bob 的概率仍约为 1/2。因此，在图 1的例子中，我们得到一个接近 5/6 的近似：将 U 中的黑色 2/3 近似匹配与一半的虚线边结合起来，就能得到一个接近$\frac{2}{3} + \frac{1}{2} \cdot \frac{1}{3} = \frac{5}{6}$的近似。我们指出，这个例子已经表明，我们的 $5/6$近似保证对于 Bernstein 协议而言是紧的，见定理 1 （正式证明见定理 3）。

该示例的优良性质图 1 在于子图 $H \cup U$ 包含一个 2/3 近似匹配 M （U中的黑色匹配），从图中移除其顶点后，仍会在 G 中留下一个 1/3 近似匹配（红色虚线边）。如果我们能证明这一点对每个图都成立，就能立即得到 Bernstein 协议的（几乎）5/6 近似分析。然而，不幸的是，这一性质并非对所有图都成立。在第6节中，我们给出 H、U 的示例，使得对于每个匹配 M 在 $H \cup U$中，都有

$$|M| + \frac{1}{2}\mu(G - V(M)) \leq 0.75\mu(G),$$

其中，$\mu(G - V(M))$ 表示从图 G 中移除 M 的顶点后，剩余的最大匹配的大小。这意味着，这一思路不足以保证 Bernstein 协议达到（近似）5/6 的近似比。

在我们的分析中，我们不是先在 $H \cup U$ 中确定一个 $2/3$-近似匹配，再利用 E_B 中的边对其进行增广，而是先通过固定一个任意的最大匹配 M^*，并取其中分配给 Bob 的一半边，来确定一个更小的 $1/2$-近似匹配。这个较小的 1/2 近似匹配的优势在于，它可以得到更好的增广。

具体而言，我们证明，这个 1/2 近似匹配在期望意义下，可以利用 $H \cup U$ 中的边，通过一个大小为（几乎） $\mu(G)/3$ 的匹配进行增广，从而使最终匹配的大小达到（几乎）$\frac{1}{2}\mu(G)+\frac{1}{3}\mu(G) = \frac{5}{6}\mu(G)$。在可用顶点中能够找到大小为 $\bar{\mu}(G)/3$ 的匹配，这一论证是我们分析的核心，并通过分数匹配加以形式化。

3 预备知识

我们首先正式定义最大匹配的鲁棒通信模型。

定义 3.1。 在 k 方单向鲁棒通信模型中，每条边都会被独立且均匀地分配给其中一位参与方。第 i 位参与方获得分配给自己的边以及来自第 $(i-1)$ 位参与方的消息 m_i 后，决定向第 $(i+1)$ 位参与方发送什么消息。第k位也是最后一位参与方，负责报告一个匹配。该模型中协议的通信复杂度定义为参与方之间传递的消息中所含单词数的最大值，即 $\max_i |m_i|$，其中 $|m_i|$ 表示 m_i中的单词数。

In 对于情形 $k = 2$，我们将第一位参与方称为 ***Alice***，第二位参与方称为 ***Bob*** .

我们用 $\mu(G)$ 表示图 G 中最大匹配的大小。对于一条边 e，我们将其边度定义为其两个端点的度数之和。

匹配理论背景

命题 3.2 (众所周知的结论)。 设 G 是任意一个图，且 x 是 G 上的一个分数匹配，使得对于任意满足 $|S|$ 小于 $\frac{1}{\varepsilon}$ 的顶点集 $S \subseteq V$，我们有

$$\sum_{e \in G[S]} x_e \leq \left\lfloor \frac{|S|}{2} \right\rfloor.$$

于是，有 $\mu(G) \geq (1-\varepsilon) \sum_e x_e$。

证明概要。 设 z 为另一个分数匹配，其中 $z_e = (1-\varepsilon)x_e$。如果 x 满足花不等式，即对于所有大小至多为 $\frac{1}{\varepsilon}$ 的 S，均有 $\sum_{e \in G[S]} x_e \leq \lfloor \frac{|S|}{2} \rfloor$，那么 z 对所有 S 都满足该不等式。为此，设 S 是一个大小至少为 $\frac{1}{\varepsilon}$ 的奇数顶点集，满足 $\sum_{e \in G[S]} x_e \geq \lfloor \frac{|S|}{2} \rfloor$。那么有：

$$\sum_{e \in G[S]} z_e = (1-\varepsilon) \sum_{e \in G[S]} x_e \leq \sum_{e \in G[S]} x_e - \frac{1}{2} \leq \frac{|S|}{2} - \frac{1}{2} \leq \left\lfloor \frac{|S|}{2} \right\rfloor.$$

因此，存在一个大小至少为 $\sum_e z_e = (1-\varepsilon)\sum_e x_e$ 的整数匹配。 □

以下定义由 Bernstein 提出 [5]。同一论文中的该命题在我们的分析中发挥着关键作用。

定义 3.3。 如果对于所有边 $(u,v) \in E_H$ 均有$d_H(u) + d_H(v) \leq \beta$，则图 H 具有有界边度 β。

定义 3.4。 给定一个图 G 和一个子图 $H \subseteq G$，边 $(u,v) \in E_G \setminus E_H$称为$(H,\beta,\lambda)$-欠满的，当且仅当 $d_H(u) + d_H(v) < (1-\lambda)\beta$。

命题 *3.5*（引理 *3.1*，出自 [5]）。固定任意 $\varepsilon \in [0, \frac{1}{2}]$，令 λ, β 为满足 $\lambda \leq \frac{\varepsilon}{384}$，$\beta \geq 50\lambda^{-2} log\left(\frac{1}{\lambda}\right)$ 的参数。考虑任意图 G 及其任意具有有界边度 β 的子图 H。令 U 包含 $G \setminus H$ 中所有 $(H, \beta, 3\lambda)$-不足边。则 $\mu(H \cup U) \geq \left(\frac{2}{3} - \varepsilon\right) \mu(G)$ 。

集中不等式

我们的证明中使用以下集中不等式。

命题 3.6 (切尔诺夫界)。设 X_1、...、X_n 是取值于 $[0,1]$ 的独立随机变量。设 $X = \sum X_i$ 且 $\mu = \mathbb{E}[X]$。那么，对于任意 $0 < \delta \leq 1$ 和 $0 < a \leq \mu$，有

$$\Pr(X \geq (1+\delta)\mu) \leq \exp\left(-\frac{\delta^2 \mu}{3}\right) \qquad and \qquad \Pr(X \geq \mu + a) \leq \exp\left(-\frac{a^2}{3\mu}\right).$$

定义 3.7 ([7])。函数 $f\colon \{0,1\}^n \to \mathbb{N}$ 是自界函数，若存在函数 f_1、...、$f_n : \{0,1\}^{n-1} \to \mathbb{N}$，使得对于所有满足以下条件的 $x \in \{0,1\}^n$

$$0 \leq f(x) - f_i(x^{(i)}) \leq 1 \qquad \forall i \in [n],$$

and

$$\sum_{i=1}^{n} \left(f(x) - f_i(x^{(i)})\right) \leq f(x).$$

其中，$x^{(i)}$ 是通过从 x 中去掉第 i 个分量得到的。

命题 3.8 ([7])。取一个自界函数 $f\colon \{0,1\}^n \to \mathbb{N}$，以及独立的 $0-1$ 个变量 X_1、...、X_n。定义 $Z = f(X_1, \ldots, X_n)$。那么，有

$$\Pr(Z \leq \mathbb{E}\, Z - t) \leq \exp\left(\frac{-t^2}{2\,\mathbb{E}\, Z}\right).$$

4 Bernstein 协议的新分析

本节旨在证明定理 1 和 2。我们将分析两方模型中的 Bernstein 协议（协议 1）。为使这一分析适用于多方模型，我们假设每条边都以概率 $p \leq \frac{1}{2}$ 独立地分配给 Bob。

我们先描述两方模型中的协议。多方协议与之相当类似，我们将在定理 2 的证明中介绍。令 E_A 为分配给 Alice 的边集，E_B 为分配给 Bob 的边集。此外，固定一个常数 $\varepsilon \in [0, \frac{1}{2}]$，并令 $\lambda = \frac{\varepsilon}{384}$，且 $\beta = 50\lambda^{-4}$。该协议的形式化描述见协议 1。

断言 4.1 表明 Alice 可以执行步骤 2，并且协议 1的通信复杂度为 $O(n \cdot \log n \cdot \mathrm{poly}(1/\varepsilon))$。注意，考虑到将边分配给 Alice 和 Bob 时的随机性，E_s 可以视为从整个边集中抽取的均匀样本，其中每条边被选中的概率为 ε。

断言 4.1 ([5]中的引理 4.1)。*Alice* 只需查看 E_s 的边，就能构造一个子图 $H \subseteq E_s$，使其度有界于 β，并且以高概率 $E_r = E(G) \setminus E_s$ 至多含有 $O(n \cdot \log n \cdot \mathrm{poly}(1/\varepsilon))$条 *{style id='32'}(H, β, λ)*-欠满边。

协议 1： 双参与单向鲁棒通信模型中通过 EDCS 实现的 Bernstein 协议。

Alice:

1. 对 E_A 进行子采样 E_s，其中每条边都以概率 $\frac{\varepsilon}{1-p}$ 独立地被选入。
2. 从 H 中取一个有界边度 β 的 E_s 子图，使得 H、β、λ-欠满边在 $E_r = E(G)$ 中的数量 $\setminus E_s$ 以高概率为 $O(n \cdot \log n \cdot \mathrm{poly}(1/\varepsilon))$。（关于 断言 4.1 的存在性，请参见 H。）
3. 找出 $E_A \setminus E_s$ 的 (H, β, λ)-不足边，并将其称为 U_A。
4. 向 Bob 传达 $H \cup U_A$。

Bob:

1. 返回 $E_B \cup H \cup U_A$ 中的最大匹配。

在分析中，我们首先构造一个分数匹配 x，其期望大小为 $\left(\frac{2}{3} - O(\varepsilon)\right)\mu(G$，其支撑集包含于 $H \cup U$。接着，我们证明可以由 y 得到一个分数匹配 x，使其支撑集包含于 $E_B \cup H \cup U_A$，且期望大小至少为 $\left(\frac{2}{3} + \frac{p}{3} - O(\varepsilon)\right)\mu(G$。最后，我们利用 y 的结构证明，它的存在意味着 $E_B \cup H \cup U_A$ 具有一个整数匹配，其大小几乎与 y 的大小相同。在 第 5 节中，我们证明以高概率也能达到该近似比。

首先，我们说明如何在给定 $H \cup U$ 的情况下构造分数匹配 x；回忆一下，U 是 E_r 中 $(H$、β、$\lambda)$-不足边的集合。固定 E_r 中的一个最大匹配 M^*。令 M_{in} 为 M^* 中出现在 $H \cup U$ 里的边，即 $M_{\text{in}} = M^* \cap (H \cup U)$，并令 $M_{\text{out}} = M^* \setminus M_{\text{in}}$。

- 从 $H_1 = H$ 和 $U_1 = U$ 开始。
- 对于 $i = 1, \ldots\ldots, \lambda\beta$：
- 令 M_i 为 $H_i \cup U_i$ 中的一个最大匹配。
- 令 $H_{i+1} = H_i \setminus (M_i \setminus M_{\text{中}})$，$U_{i+1} = U_i \setminus (M_i \setminus M_{\text{中}})$。
- 对于每条边 e，令 $x_e = \frac{|\{i : e \in M_i\}|}{\lambda\beta}$。

可以将这个过程理解为：从 $H \cup U$ 开始，每次取一个最大匹配 M_i，并从图中移除 $M_i \setminus M_{\text{in}}$。然后，令 x_e 等于我们取过的匹配中包含 e 的匹配所占的比例。注意，匹配 M_1、$\ldots M_{\lambda\beta}$ 只能在 M_{in} 中相交。我们将使用 命题 3.5 来证明 x 的期望大小至少为 $\left(\frac{2}{3} - O\left(\varepsilon\right)\right)\mu\left(G\right)$。

引理 4.2。 成立的是 $\mathbb{E}\left[\sum_e x_e\right] \geq \left(\frac{2}{3} - \frac{5}{3}\varepsilon\right)\mu(G)$。

证明。 我们应用 命题 3.5 于 $G_i = (H \cup E_r) \setminus \left(\bigcup_{j<i} M_j \setminus M_{\text{in}}\right)$，$H_i$和 U_i。从 H中移除一个匹配后，对于任意一条边，其在 H 中的度最多减少 2。此外，U 包含所有

属于 H、β、λ-不足边的 E_r。因此，U_i 包含 $G_i \setminus H_i$ 中所有 H_i-度小于 $(1-\lambda)\beta - 2(i-1) \geq (1-3\lambda)\beta$ 的边。因此，命题 3.5 蕴含

$$|M_i| \geq \left(\frac{2}{3} - \varepsilon\right) \mu(G_i).$$

此外，注意到 G_i 总是包含 M^*，因此有 $\mu(G_i) = \mu(E_r)$，并且我们得到：

$$\sum_e x_e \geq \frac{1}{\lambda\beta} \sum_i |M_i| \geq \left(\frac{2}{3} - \varepsilon\right) \mu(E_r)$$

考虑到 $\mathbb{E}[\mu(E_R)] \geq (1-\varepsilon)\mu(G)$ 这一事实，我们得到：

$$\mathbb{E}\left[\sum_e x_e\right] \geq \left(\frac{2}{3} - \varepsilon\right)(1-\varepsilon)\mu(G) \geq \left(\frac{2}{3} - \frac{5}{3}\varepsilon\right)\mu(G). \quad \square$$

为了描述如何得到 y，我们以 E_s 为条件，从而固定 H、U 和 x。y 的支撑集包含于 $E_B \cup H \cup U_A$ 中，也就是 Bob 最终能够访问的边。我们证明，y 的期望大小至少约为 $p \cdot \mu(E_r) + (1-p)\sum_e x_e$，其中随机性来自剩余边 E_r 如何在 Alice 和 Bob 之间分配。去掉关于 E_s 的条件后，根据引理 4.2，该值的期望大于 $\left(\frac{2}{3} + \frac{p}{3} - O(\varepsilon)\right)\mu(G)$。

抽取 E_B 后，取一个匹配 M'，其中 M_{in} 的每条边都以概率 p 独立地被包含，且 $M_{\text{out}} \cap E_B$ 的每条边都以概率 $1-\varepsilon$ 独立地被包含。注意，在以 E_s 为条件时，M_{out} 的每条边被分配给 Bob 的概率为 $\frac{p}{1-\varepsilon}$。因此，M_{out} 的每条边最终进入 M' 的概率为 $\frac{p}{1-\varepsilon}(1-\varepsilon) = p$，也就是说，$M'$ 独立地以概率 p 包含 M^* 的每条边。

对于任意边 $e \notin M^*$，定义 p_e 为 e 不与 M' 中任何边相邻的概率。注意，p_e 恰好等于 $(1-p)$ 的 M' 中与 e 相邻的边数次方。我们如下定义匹配 $\hat{y}$：

$$\hat{y}_e = \begin{cases} 1 & \text{if } e \in M', \\ x_e & \text{if } e \in M^* \setminus M', \\ 0 & \text{if } e \notin M^* \text{ and } e \text{ is adjacent to an edge of } M', \\ (1-p) \cdot \dfrac{x_e}{p_e} & \text{otherwise.} \end{cases}$$

然后，我们将 $\hat{y}$ 缩小 $1+\varepsilon$ 倍，并将部分边置零，从而得到一个分数匹配。形式化地，我们令：

$$y_{(u,v)} = \begin{cases} 0 & \text{if } \hat{y}_u/(1+\varepsilon) > 1 \text{ or } \hat{y}_v/(1+\varepsilon) > 1, \\ \dfrac{\hat{y}_{(u,v)}}{1+\varepsilon} & \text{otherwise.} \end{cases}$$

引理 4.3。在 E_s 的条件下，有

$$\mathbb{E}\left[\sum_e y_e\right] \geq (1-3\varepsilon)p \cdot \mu(E_r) + (1-3\varepsilon)(1-p)\sum_e x_e - 2\varepsilon\mu(G).$$

证明。本证明中的所有论证均以 E_s 为条件。

断言 4.4。 对于每个顶点 u，均有 $\mathbb{E}[\hat{y}_u] = p \cdot \chi_{M^*}(u) + (1-p)x_u$，其中 $\chi_{M^*}(u)$ 若 u 被 M^* 覆盖则等于1，否则等于零。

证明。首先，考虑一个被 M^* 覆盖的顶点 u，比如由边 $e^* \in M^*$ 覆盖。当 e^* 出现在 M' 中时，我们有 $\hat{y}_{e^*} = 1$，而对于与 u 相邻的所有其他边 e，$\hat{y}_e$ 的值等于零。因此，我们将有 $\hat{y}_u = 1$，即 $\mathbb{E}[\hat{y}_u \mid e^* \in M^*] = 1$。

现在，我们以 $e^* \notin M'$ 为条件。在这种情况下，我们将得到 $\hat{y}_{e^*} = x_{e^*}$。此外，对于任意其他一条与 u 相邻的边 e，e 不与 M' 中任何边相邻的概率等于 $\frac{p_e}{1-p}$。因此，以 $\frac{p_e}{1-p}$ 的概率，有 $\hat{y}_e = (1-p)\cdot\frac{x_e}{p_e}$ 成立，而在其他情况下，我们将得到 $\hat{y}_e = 0$。因此，我们可以写出：

$$\mathbb{E}\left[\hat{y}_u \mid e^* \notin M'\right] = x_{e^*} + \sum_{\substack{e \ni u \\ e \neq e^*}} \frac{p_e}{1-p} \cdot \left((1-p)\cdot\frac{x_e}{p_e}\right) = \sum_{e \ni u} x_e = x_u.$$

因此，对于被 M^* 覆盖的顶点 u，有

$$\mathbb{E}[\hat{y}_u] = p \cdot \mathbb{E}\left[\hat{y}_u \mid e^* \in M'\right] + (1-p)\cdot\mathbb{E}\left[\hat{y}_u \mid e^* \notin M'\right] = p + (1-p)x_u.$$

顶点 u 未被 M^* 覆盖的情形可类似推得。对于每条与 u 相邻的边 e，以概率 p_e 有 $\hat{y}_e = (1-p)\cdot\frac{x_e}{p_e}$，否则有 $\hat{y}_e = 0$。因此

$$\mathbb{E}[\hat{y}_u] = \sum_{e \ni u} p_e \cdot \left((1-p)\cdot\frac{x_e}{p_e}\right) = (1-p)x_u. \quad \square$$

以下断言有助于我们证明，将 $\hat{y}$ 缩小并将部分边置零时，不会损失太多。

断言 4.5。 对于每个顶点 u，若 u 未被 M^* 或 $x_u \leq \frac{1}{2}$ 覆盖，则有 $\hat{y}_u \leq 1$；否则有 概率 $(\hat{y}_u > 1+\varepsilon) \leq \varepsilon$。

证明。考虑一个未被 M^* 覆盖的顶点 u。对于每条与 u 相邻的边 e，有 $p_e \geq 1-p$，因为 e 至多有一条相邻边属于 M^*。因此我们有：

$$\hat{y}_u \leq \sum_{e \ni u} (1-p)\cdot\frac{x_e}{p_e} \leq \sum_{e \ni u} x_e \leq 1.$$

现在考虑一个被 M^* 覆盖的顶点 u，假设它由边 $e^* \in M^*$ 覆盖。若 $e^* \in M'$，则有 $\hat{y}_u = 1$。因此我们假设 $e^* \notin M'$，相应地有 $\hat{y}_e = x_e$。对于任意其他边 e，有 $p_e \geq (1-p)^2$。因此我们有：

$$\hat{y}_u \leq x_{e^*} + \sum_{\substack{e \ni u \\ e \neq e^*}} (1-p)\cdot\frac{x_e}{p_e} \leq x_{e^*} + \sum_{\substack{e \ni u \\ e \neq e^*}} \frac{x_e}{1-p} \leq 2\sum_{e \ni u} x_e = 2x_u.$$

因此，如果 $x_u \leq \frac{1}{2}$ 成立，那么可推出 $\hat{y}_u \leq 1$。

对于其他情况，我们使用 切尔诺夫界 来证明，以高概率 $\hat{y}_u$ 不会远大于 1。如前所述，如果 e^* 出现在 M'中，则有 $\hat{y}_u = 1$。因此，我们以 $e^* \notin M'$为条件。我们将 $X = y_u - y_{e^*}$ 表示为取值于 $[0, 4/\lambda\beta]$的一组独立随机变量之和。注意，由于 M_{in} 之外的边最多出现在一个 M_i中，对于$e \notin M^*$我们有 $x_e \leq \frac{1}{\lambda\beta}$。

取一条边 $e=(u,\ v)\neq e^*$。如果 v 在 M^* 中没有与 u 的另一个邻点匹配，那么 y_e 的值与和 u 相邻的其他边的值相互独立。它以概率 $\frac{p_e}{1-p}$ 等于 $(1-p)\cdot\frac{x_e}{p_e}\leq\frac{2}{\lambda\beta}$，否则为零。

如果 v 在 M^* 中与 u 的另一个邻点 v' 匹配。令 $e'=(u,\ v')$。那么 $y_e+y'_e$ 的值与邻接于 u 的其他边的值相互独立。它以概率 $(1-p)$ 等于 $(1-p)\cdot\frac{x_e+x'_e}{(1-p)^2}\leq\frac{4}{\lambda\beta}$，否则为零。

因此，通过将匹配在一起的边两两配对，我们可以将 X 表示为 $[0,4/\lambda\beta]$ 中独立随机变量之和。根据 X 在 断言 4.4 中的计算，其期望等于$x_u-x_{e^*}\leq 1$。根据 切尔诺夫界 可得：

$$\begin{aligned}\Pr\left(\hat{y}_u>1+\varepsilon\right)&\leq\Pr\left(X>\mathbb{E}\,X+\varepsilon\right)\\&\leq\Pr\left(X\cdot\frac{\lambda\beta}{4}>\mu\frac{\lambda\beta}{4}+\varepsilon\frac{\lambda\beta}{4}\right)\\&\leq\exp\left(-\frac{\varepsilon^2\lambda^2\beta^2/16}{3\mu\lambda\beta/4}\right)\end{aligned}$$

（根据 切尔诺夫界，注意到 $X\lambda\beta/4$ 是 $[0,1].$ 中独立随机变量之和）

$$\begin{aligned}&=\exp\left(-\frac{\varepsilon^2\lambda\beta}{12}\right)\\&<\varepsilon,\end{aligned}$$

（由于 $\lambda=\frac{\varepsilon}{384}$ 和 $\beta=50\lambda^{-4}$）

完成证明。 □

我们分析 $\mathbb{E}\left[y_u\right]$。考虑分两步生成 y_u。首先，对于每条边 (u,v)，如果 $\hat{}y_v\leq 1+\varepsilon$，我们令$y_{(u,v)}$ 等于 $\frac{1}{1+\varepsilon}\hat{}y_{(u,v)}$；否则令其为零。然后，如果 $\hat{}y_u>1+\varepsilon$，我们将所有与 u 相邻的边上的 y 置零。

在第一步中，将 $\hat{}y_u$ 缩小会使我们损失一个因子（$1+\varepsilon$）。此外，根据 断言 4.5，当我们将边（$u,\ v$）置零时，因为 $y_v>1+\varepsilon$，每条边都会损失一个 ε-比例，因此 $\mathbb{E}\left[\hat{}y_u\right]$ 也会损失相应比例。在第二步中，同样根据 断言 4.5，如果 $x_u\leq\frac{1}{2}$ 或 u 未被 M^* 覆盖，则我们不会损失任何东西。否则，我们以至多 ε 的概率将所有边置零。我们有 $\hat{}y_u\leq 2x_u\leq 2$，因此会损失一个加性因子 2ε。总而言之，对于任意顶点 u，我们有：

$$\mathbb{E}\left[y_u\right]\geq(1-\varepsilon)\frac{\mathbb{E}\left[\hat{y}_u\right]}{1+\varepsilon}-2\varepsilon\geq(1-3\varepsilon)\left(p\cdot\chi_{M^*}(u)+(1-p)x_u\right)-2\varepsilon,$$

对于任意满足 $x_u\leq\frac{1}{2}$ 的顶点 u，我们得到：

$$\mathbb{E}\left[y_u\right]\geq(1-3\varepsilon)\left(p\cdot\chi_{M^*}(u)+(1-p)x_u\right).$$

注意到，由于 x 的各分量之和至多为 $\mu(G)$，因此至多有 $2\mu(G)$ 个顶点满足 $x_u\geq\frac{1}{2}$。因此，将最后两个方程对 u 求和，我们得到：

$$\begin{aligned}\sum_u\mathbb{E}\left[y_u\right]&=\sum_{u\in V(M^*)}\mathbb{E}\left[y_u\right]+\sum_{u\notin V(M^*)}\mathbb{E}\left[y_u\right]\\&\geq\left((1-3\varepsilon)\sum_{u\in V(M^*)}p+(1-p)x_u\right)+\left((1-3\varepsilon)\sum_{u\notin V(M^*)}(1-p)x_u\right)-2\varepsilon\cdot 2\mu(G)\end{aligned}$$

$$= (1-3\varepsilon)2p \cdot |M^*| + (1-3\varepsilon)(1-p)\sum_u x_u - 4\varepsilon\mu(G).$$

回忆一下，$|M^*| = \mu(E_r)$。最后，将等式两边除以 2，我们得到：

$$\sum_e \mathbb{E}[y_e] \geq (1-3\varepsilon)p \cdot \mu(E_r) + (1-3\varepsilon)(1-p)\sum_e x_e - 2\varepsilon\mu(G). \qquad \square$$

现在我们去掉对 E_s 的条件限制。

引理 4.6。有 $\mathbb{E}\left[\sum_e y_e\right] \geq \left(\frac{2}{3} + \frac{p}{3} - 6\varepsilon\right)\mu(G)$。

证明。我们有：

$$\begin{aligned}
\mathbb{E}\left[\sum_e y_e\right] &= \mathbb{E}\left[\mathbb{E}\left[\sum_e y_e \;\middle|\; E_s\right]\right] \\
&\geq \mathbb{E}\left[(1-3\varepsilon)p \cdot \mu(E_r) + (1-3\varepsilon)(1-p)\sum_e x_e - 2\varepsilon\mu(G)\right] \\
&\geq (1-3\varepsilon)p \cdot (1-\varepsilon)\mu(G) + (1-3\varepsilon)(1-p)\left(\frac{2}{3} - \frac{5}{3}\varepsilon\right)\mu(G) - 2\varepsilon\mu(G) \\
&\qquad\qquad \text{(by } \mathbb{E}[\mu(E_r)] = (1-\varepsilon)\mu(G) \text{ and Lemma 4.2)} \\
&\geq \left(\frac{2}{3} + \frac{p}{3} - \left(\frac{7}{6}p + \frac{29}{6}\right)\varepsilon\right)\mu(G) \\
&\geq \left(\frac{2}{3} + \frac{p}{3} - 6\varepsilon\right)\mu(G).
\end{aligned}$$

$\square$

我们证明 $E_B \cup H \cup U_A$ 具有一个整数匹配，其大小几乎与 y 的大小相同。

引理 4.7。在 $E_B \cup H \cup U_A$ 中存在一个大小为 $(1-3\varepsilon)\sum_e y_e$ 的匹配。

证明。注意，对于每条边 e，除去 M^* 中构成匹配的边之外，都有 $y_e \leq \frac{4}{\lambda\beta} \leq \varepsilon^3$。因此，对于任意满足 $|S|$ 小于 $\frac{1}{\varepsilon}$ 的顶点集 $S \subseteq V$，我们有：

$$\sum_{e \in G[S]} x_e = \sum_{e \in G[S] \cap M^*} x_e + \sum_{e \in G[S] \setminus M^*} x_e \leq |G[S] \cap M^*| + \frac{1}{\varepsilon^2}\varepsilon^3 \leq \left\lfloor \frac{|S|}{2} \right\rfloor + \varepsilon.$$

因此，我们可以将命题 3.2 应用于 $(1-2\varepsilon)\,y$，得到 $\mu\left(E_B \cup H \cup U_A\right) \geq (1-3\varepsilon)\sum_e y_e$。

$\square$

定理 *1* 的证明。根据断言 4.1，Bernstein 协议（协议 1）仅使用$O(n \cdot$对数$\, n \cdot$多项式$(1/\varepsilon))$个通信字数即可实现。

根据引理 4.6，存在一个期望大小为 y的分数匹配 $\left(\frac{2}{3} + \frac{p}{3} - 6\varepsilon\right)\mu(G)$。结合引理 4.7，我们可以得出协议 1 达到了 $\left(\frac{2}{3} + \frac{p}{3} - 9\varepsilon\right)$近似比。关于如何证明以高概率也能达到该近似比，请参阅第 5 节。最后，令 $p = \frac{1}{2}$ 并对 ε 重新缩放，即可证明该定理。

$\square$

定理 *2* 的证明。我们需要调整协议 1 以适用于 k方模型。第一方将以概率 $\varepsilon/(1-1/k)$ 独立采样其每条边，以获得 E_s。随后它将

构造具有有界边度的子图 $H \subseteq E_s$，并将其与 (H, β, λ)-不足边一起发送给下一方。除最后一方外，每一方都会将分配给自己的 (H, β, λ)-不足边，以及收到的消息中的边，传递给下一方。最后，最后一方将报告由其可访问的所有边组成的图中的最大匹配。

因此，设置 $p = \frac{1}{k}$ 后，前 $k-1$ 个参与方将在我们的分析中充当 Alice，而最后一个参与方充当 Bob。因此，通过与 定理 1 的证明 类似的论证，协议 1 实现了 $(\frac{2}{3} + \frac{1}{3k} - 9\varepsilon)$ 近似比，并且对 ε 进行重新缩放即可证明该定理。 □

5 从期望到高概率

在本节中，我们将证明，只需稍作修改，Bernstein 协议（协议 1）就能以高概率实现 $\frac{5}{6}$-近似。为此，当边的数量过少时，Alice 应将所有边发送给 Bob。

断言 5.1. 不失一般性，我们可以假设 $\mu(G) = \Omega(\log n)$.

证明。 可以使用收费论证来证明， G中的边数小于 $2n\mu(G)$。固定 G中的一个最大匹配 M 。对于任意边 e，向 M 中一条与 e相邻的边收取一个单位。这样的边必定存在，因为 M 是一个最大匹配。这样，我们对 G中的每条边各收费一次，而 M 中的每条边通过其每个端点最多被收费 n 次。

要了解为什么该断言成立，请注意，如果 $\mu(G)$ 太小，即 $\mu(G) = O(\log n)$，图中的边数将为 $O(n \log n)$，Alice 可以将所有边发送给 Bob。 □

引理 5.2。 假设 $\mu(G) = \Omega(\log n)$，无论 协议 *1* 在期望意义下达到何种近似比，都将以高概率达到。

证明。 我们以采样边集 E_s 为条件，从而固定 H 和 U。Bob 可以访问 $H \cup U$ 的边，因为这些边要么分配给 Bob，要么由 Alice 通知 Bob。令 $e_1, \ldots, e_k$ 表示其他边，即 $E_r \setminus U$ 的边。这些边中的每一条都以概率 $\frac{1/2}{1-\varepsilon}$ 独立地分配给 Bob。

我们定义一个自界函数 $f\colon \{0,1\}^k \to \mathbb{Z}$。对于 $x \in \{0,1\}^k$， $f(x)$ 的值等于 $H \cup U \cup E_x$ 的最大匹配，其中 $E_x = \{e_i \mid x_i = 1\}$。等价地， $f(x)$ 是将边 $\{e_i \mid x_i = 1\}$ 分配给 Bob 时的输出匹配大小，即 $E_B \cup H \cup U_A$ 等于 $E_x \cup H \cup U$。此外，令 $f_i(x^{(i)}) = f(x_1, \ldots x_{i-1}, 0, x_{i+1}, \ldots x_k)$。

任取一个 $x \in \{0,1\}^k$。注意， $f_i(x^{(i)})$ 等于 $\mu(E_x \setminus e_i)$，并且从图中移除一条边，最多会使最大匹配减少 1。因此，有：

$$0 \le f(x) - f_i(x^{(i)}) \le 1, \qquad \forall i : 1 \le i \le k.$$

取 $H \cup U \cup E_x$ 中的最大匹配 M，并令 I 表示 $E_x \cap M$ 中边的索引，即 $I = \{i \mid x_i = 1$ 和 $e_i \in M\}$。对于任意 $i \notin I$，边集 $E_x \setminus e_i$ 包含 M。因此， $f_i(x^{(i)})$ 等于 $f(x)$，我们有：

$$\sum_{i=1}^{k} \left(f(x) - f_i(x^{(i)}) \right) \le |I| = f(x).$$

因此， f 是一个自界函数。

现在我们可以应用 命题 3.8。令 X_i 为指示变量，表示 e_i 被分配给 Bob，即 X_i 当 $e_i \in E_B$ 时等于 1。令 $Z = f(X_1, \ldots, X_k)$，并令 $\mu = \mathbb{E}[Z] = r\mu(G)$。也就是说， Z是输出匹配的大小，而 r 是 协议 1 在期望意义下达到的近似比。根据 命题 3.8，我们有：

$$\Pr\left(Z \le r\mu(G) - \sqrt{2\mu(G)\log n}\right) \le \exp\left(-\frac{2\mu(G)\log n}{2\mu(G)}\right) = \frac{1}{n}.$$

因此，以高概率，协议 1 输出大小为 $(1 - o(1))r\mu(G)$ 的匹配。注意，根据假设，偏差为 $o(1)$，且 $\mu(G) = \Omega(\log n)$。 □

6 Bernstein 协议的一些实例

本节中，我们首先证明 定理 3 ，说明我们在定理 1和 2 中对 Bernstein 协议的分析是紧的。然后，我们将形式化阐述在第 2 节中所作的一条注释。

定理 3的证明。 考虑一个二分图 $G(L, R)$，满足 $|L| = |R|$。其中 L 由三个大小相等的顶点组 A_1、 A_2 和 A_3组成，类似地， R 由 B_1、 B_2 和 B_3组成。导出子图 $M_1 = G[A_1, B_1]$、 $M_2 = G[A_2, B_2]$和 $M_3 = G[A_3, B_3]$ 是完美匹配。导出子图 $K_1 = G[A_1, B_2]$ 和 $K_2 = G[A_2, B_3]$ 是完全二分图，且图中没有其他边（见图 1）。注意， G 有一个完美匹配，即最大匹配的大小等于 $|L| = |R|$。令 $\beta \le \frac{|V(G)|}{12k}$，其为$O(|V(G)|)$。 ——

令 E_L 为分配给最后一个参与方的边集。为了给多方协议的近似比设定上界，我们为 $E_L \cup H \cup U$构造一个顶点覆盖。众所周知，最小顶点覆盖的大小大于最大匹配的大小。以高概率，第一个参与方可以使 H 完全包含在 $K_1 \cup K_2$中，从而 U 将等于$M_1 \cup M_3$，且 M_2 中不会出现任何边。

令 $X = V(M_2 \cap E_L) \cap A_2$，即 X 从分配给最后一个参与方的 M_2 中的每条边各包含一个端点。我们断言， $A_1 \cup B_3 \cup X$ 是 $E_L \cup H \cup U$的一个顶点覆盖。这是因为 A_1覆盖了 M_1 和 K_1的边， B_3 覆盖了 M_3 和 K_2的边，而 X 覆盖了所有剩余的边，也就是 $E_L \cap M_2$。

在如上所述的 H 条件下， M_2 的每条边被分配给 Bob 的概率为 $\frac{1/k}{1-\varepsilon}$。因此，顶点覆盖的期望大小等于

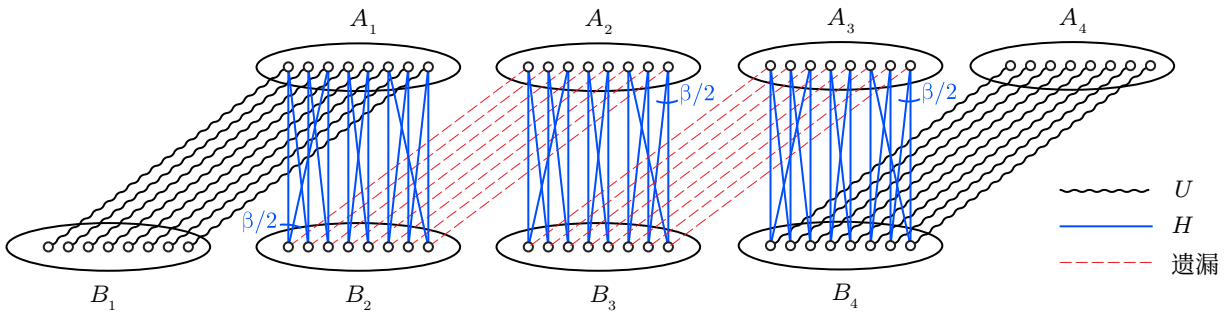
$$|A_1| + |B_3| + \frac{1/k}{1-\varepsilon}|A_2| = \left(\frac{1}{3} + \frac{1}{3} + \frac{1/k}{1-\varepsilon}\cdot\frac{1}{3}\right)\mu(G) \le (1+2\varepsilon)\left(\frac{2}{3} + \frac{1}{3k}\right)\mu(G).$$

令 ε 任意小，即可证明该定理。 □

如第 2 节所述，定理 3 中讨论的图（见图 1）具有一个良好性质，即存在一个较大的匹配 M，使得 $G - V(M)$ 也有一个较大的匹配。由于 Bernstein 协议的输出期望大小至少为 $|M| + \frac{1}{2}\mu(G - V(M))$，在此情形下它等于 $\frac{5}{6}\mu(G)$，这一性质似乎可能有助于分析该协议。然而，下面的断言表明，这样的 M 并不总是存在。

命题 6.1. 存在一个图 G，其顶点数可以任意大，使得对于 $\beta \leq |V(G)|/4$，存在一组 H和 U，使得 $H \cup U$中的每个匹配 M都满足 $|M| + \frac{1}{2}\mu(G - V(M)) \leq 0.75\mu(G)$。其中， H是一个具有有界边度 β 的子图，而 U 是 $G \setminus H$ 中欠满边的集合，也就是说，是 $G \setminus H$ 中 H-度小于 $\beta - 1$ 的边。

证明。设 $G(L, R)$ 是一个二分图，满足 $|L| = |R|$。其中 L 由四组大小相同的顶点 A_1、 A_2、 A_3 和 A_4 组成，类似地， R 由 B_1、 B_2、 B_3 和 B_4 组成。导出子图 $M_1 = G[A_1, B_1]$、 $M_2 = G[A_2, B_2]$、 $M_3 = G[A_3, B_3]$ 和 $M_4 = G[A_4, B_4]$ 是完美匹配。导出子图 $K_1 = G[A_1, B_2]$、 $K_2 = G[A_2, B_3]$ 和 $K_3 = G[A_3, B_4]$ 是完全二分图，图中没有其他边。注意， G 有一个完美匹配（见图 2）。



Fig图 2：一个例子，其中每个匹配 M都满足$|M| + \frac{1}{2}\mu(G - V(M)) \leq 0.75\mu(G$).

设 H 是 $\beta/2$-正则子图 $K_1 \cup K_2 \cup K_3$。相应的 U 等于 $M_1 \cup M_4$，且 $M_2 \cup M_3$ 中没有任何边出现在 $H \cup U$ 中。我们证明 $\max_M |M| + \frac{1}{2}\mu(G - V(M)) \leq 0.75\mu(G)$，其中 M 遍历 $H \cup U$ 中的所有匹配。若一个匹配使 $|M| + \frac{1}{2}\mu(G - V(M))$ 达到可能的最大值，则称其为最优匹配。

首先，我们证明存在一个包含 $M_1 \cup M_4$ 中所有边的最优匹配。为此，取一个最优匹配 M。取 A_1 中任意一个顶点 u ，并设其在 M_1 中的相邻边为 e。如果 u 被某条边 $e' \in M$ 覆盖，那么从 M 中移除 e' 并加入 e ，不会使 $|M| + \frac{1}{2}\mu(G - V(M))$ 减小。因为这不会改变 $|M|$，并且只会使 $\mu(G - V(M))$ 增大。当 u 未被 M 覆盖时，将该 e 加入 M 会使 $|M|$ 增加一，而 $\mu(G - V(M))$ 至多减少一。对于 B_4 中的顶点，同样的论证也成立。因此，通过反复加入这样的边，我们可以得到一个包含 $M_1 \cup M_4$ 的最优匹配。

现在我们可以将注意力限制在 $A_2 \cup A_3 \cup B_2 \cup B_3$ 上。我们断言，无论 M（即 $M \setminus (M_1 \cup M_2) = M \cap K_2$）的其余部分是什么， $|M| + \frac{1}{2}\mu(G - V(M))$ 的值都相同。因为如果 $|M \cap K_2|$ 等于 k，则有 $\mu(G - V(M)) = \frac{1}{2}\mu(G) - 2k$。因此， $|M| + \frac{1}{2}\mu(G - V(M))$ 的最优值等于

$$|M_1| + |M_2| + k + \frac{1}{2}\left(\frac{1}{2}\mu(G) - 2k\right) = \frac{3}{4}\mu(G).$$

要理解为何 $\mu(G - V(M)) = \frac{1}{2}\mu(G) - 2k$，注意到由于 $M_1 \cup M_4 \subseteq M$， $B_2 \cup A_3$ 中的任意顶点都是 $G - V(M)$ 中的单元素。因此， $G - V(M)$ 中的最大匹配由 $M_2 \cup M_3$ 中不与 M 中的任何边相邻的边组成，其大小为 $\frac{1}{2}\mu(G) - 2k$。 □

致谢

第二作者感谢 David Wajc 就如何通过 EDCS 超越 2/3 近似所进行的启发性讨论。

参考文献

[1] Sepehr Assadi 和 Soheil Behnezhad。随机顺序流式匹配中超越三分之二近似。在第*48*届自动机、语言与编程国际研讨会（*ICALP 2021*），*2021*年*7*月*12*日至*16*日，苏格兰格拉斯哥（线上会议），第19:1–19:13页，2021年。[2] Sepehr Assadi 和 Soheil Behnezhad。二分图匹配的鲁棒通信复杂度。在近似、随机化与组合优化：算法与技术，*APPROX/RANDOM 2021*，*2021*年*8*月*16*日至*18*日，华盛顿大学，美国华盛顿州西雅图（线上会议），第48:1–48:17页，2021年。[3] Sepehr Assadi 和 Aaron Bernstein。迈向匹配问题稀疏化的统一理论。在第二届算法简洁性研讨会（*SOSA 2019*），*2019*年*1*月*8*日至*9*日，美国加利福尼亚州圣迭戈，第11:1–11:20页，2019年。[4] Soheil Behnezhad。通过 Gallai-Edmonds 分解改进 EDCS 的分析。*CoRR*，abs/2110.05746，2021年。[5] Aaron Bernstein。改进随机顺序数据流中匹配的界。在第*47*届自动机、语言与编程国际研讨会（*ICALP 2020*），*2020*年*7*月*8*日至*11*日，德国萨尔布吕¨肯（线上会议），载于《LIPIcs》第168卷，第12:1–12:13页。Schloss Dagstuhl - Leibniz 信息学中心，2020年。[6] Aaron Bernstein 和 Cliff Stein。二分图中的全动态匹配。载于《自动机、语言与编程——第*42*届国际研讨会（*ICALP 2015*）论文集》，日本京都，*2015*年*7*月*6*日至*10*日，第一部分，载于《计算机科学讲义》第9134卷，第167–179页。施普林格，2015年。[7] Stephane Boucheron、Gabor Lugosi 和 Pascal Massart。关于自界函数的集中性。《电子概率杂志》，14(none):1884–1899，2009年。[8] Amit Chakrabarti、Graham Cormode 和 Andrew McGregor。通信与流计算的鲁棒下界。载于第*40*届*ACM*计算理论研讨会年会论文集，加拿大不列颠哥伦比亚省维多利亚，*2008*年*5*月*17*日至*20*日，第641–650页，2008年。[9] Joan Feigenbaum、Sampath Kannan、Andrew McGregor、Siddharth Suri 和 Jian Zhang。半流式模型中的图问题。理论计算机科学，348(2-3):207–216，2005年。[10] Ashish Goel、Michael Kapralov 和 Sanjeev Khanna。最大二分图匹配的通信与流式复杂度。载于第二十三届*ACM-SIAM*离散算法研讨会论文集（*SODA 2012*），日本京都，*2012*年*1*月*17*日至*19*日，第468–485页，2012年。

[11] Michael Kapralov。流式模型中匹配的更优界。载于第二十四届*ACM-SIAM*离散算法研讨会论文集（*SODA 2013*），美国路易斯安那州新奥尔良，*2013*年*1*月*6*日至*8*日，第1679–1697页。SIAM，2013年。[12] Michael Kapralov。边到达模型中最大匹配近似的空间下界。载于*2021*年*ACM-SIAM*离散算法研讨会（*SODA 2021*），线上会议，*2021*年*1*月*10*日至*13*日，第1874–1893页。SIAM，2021年。[13] 姚期智。与分布式计算相关的一些复杂性问题（初步报告）。载于第*11*届*ACM*计算理论研讨会年会论文集，*1979*年*4*月*30*日至*5*月*2*日，美国佐治亚州亚特兰大，第209–213页。ACM，1979年。