arXiv:2302.05966v1 [cs.DS] 12 Feb 2023

谱图论中的无限刘易斯权重

Amit Suliman* OmriWeinstein†希伯来大学 哥伦比亚大学
与希伯来大学

摘要

我们研究了根据图中各边的 ℓ_∞刘易斯权重对图重新加权所产生的谱含义。我们的主要动机是ER最小化问题（Saberi等人，SIAM 2008）：给定一个无向图G，目标是寻找正的归一化边权重$w \in \mathbb{R}^m_+$，使*e*ff*ective-resistances* ofG_w 的两两之和最小（基尔霍夫指数）。相比之下，ℓ_∞ 刘易斯权重使*maximum*有效电阻*of*边最小，但其近似成本低得多，尤其对于拉普拉斯矩阵而言。出于这一算法动机，我们研究了刘易斯权重所实现的ER近似比。

我们的第一个主要结果是，ℓ_∞刘易斯权重可为树上的ER最小化提供常数（≈ 3.12）近似。该证明引入了一种针对树上有效电阻（ℓ_2拥塞）的新技术——局部极化过程，这一技术在电网络分析中具有独立意义。对于一般图，我们证明了刘易斯权重所得近似比的上界 $\alpha(G)$，该上界始终为 $\leq \min\{\mathrm{diam}(G), \kappa(L_{w_\infty})\}$，其中 κ 是加权拉普拉斯矩阵的条件数。¯我们的所有近似算法均在*input-sparsity* 时间$O(m)$内运行，相较于Saberi等人用于精确ER最小化的$O(m^{3.5})$半定规划，这是重大改进。我们猜想，刘易斯权重可为*any*图提供$O(\log n)$近似，并通过实验表明，许多图上的近似比为$O(1)$。

最后，我们展示了 ℓ_∞-LW 重加权对图的 谱间隙 及其 谱稀疏性 的有利影响（Anari和Gharan，2015年）。在取得这些结果的过程中，我们证明了 Mohar经典界 关于 $\lambda_2(G)$ 的加权类比，并给出了矩阵杠杆分数的新刻画：它们等于包围椭球体积关于权重的梯度。

*本研究获得ERC起始资助项目CODY 101039914和ISF资助项目3011005535的支持。tima.suliman@gmail.com
†本研究获得ERC起始资助项目CODY 101039914和ISF资助项目3011005535的支持，omriwe@cs.huji.ac.il

目录

1 引言 21.1 主要结果.4

2 技术概述 5

3 预备知识 **7**

3.1 符号.7 3.2 杠杆分数与刘易斯权重.8 3.3 图的有效电阻.9 3.4 杠杆分数与有效电阻的新刻画.10

4 通过 ℓ_∞-刘易斯权重近似 ERMP **11**

4.1 ℓ_∞-LW 对树是 $O(1)$ 近似.12 4.1.1 定理 1.3 的证明.14 4.2 一般图上的 ERMP.18 4.2.1 算术-几何平均上界.18 4.2.2 对偶间隙上界.19 4.2.3 ERMP 的近似比上界.20

5 实验结果 **23**

6 谱学应用 局部变换 ℓ_∞-LW 重加权的应用（定理 1.6 的证明）24

6.1 界定代数连通度.24 6.2 最小化图的混合时间.25 6.3 最优谱稀疏树.27

7 计算刘易斯权重

287.1 以低精度计算 ℓ_∞-刘易斯权重 [Coh+20] .29

8 一般半正定矩阵的 A 与 D 设计之间的$\Omega(n/\log^2 n)$ 差距 **30**

8.1 HM 与 GM 比较.31 8.2 一般半正定矩阵的 A-最优设计与 D-最优设计之间的差距.31

A 附录 **36**

A.1 背景：凸优化分析.36 A.2 背景：实验最优设计.36 A.3 D-最优设计问题.37 A.4 命题 8.2 的证明.40 A.5 ERMP 的拉格朗日对偶问题.41 A.6 ERMP 的半正定规划表述.43 A.7 SEV 的最优性准则.43 A.8 上部 LT 的图示.43

1 引言

矩阵 $A \in \mathbb{R}^{m\times n}$ 的 ℓ_p-刘易斯权重推广了行的统计杠杆分数这一概念，它们以非正式的方式量化了行 $a_i^\top$ 在构成 A 的谱时的重要性。形式上，A 的 ℓ_p-刘易斯权重（ℓ_p-LW）是满足以下不动点方程的唯一向量 $w \in \mathbb{R}^m$

$$a_i^T \left(A^T \overline{W}^{1-2/p} A\right)^{-1} a_i = \overline{w}_i^{2/p} \quad , \quad \forall\, i \in [m] \tag{1.1}$$

其中 $\overline{\mathrm{W}} = \mathrm{diag}(\overline{\mathrm{w}}) \in \mathbb{R}^{m\times m}$。 [Lew78] 的这一定义可以看作对各行进行密度变换，要求每一行重新加权后的杠杆分数均为 1；也就是说，重新加权矩阵 $\mathrm{W}^{1/2-1/p}\mathrm{A}$ 的第 i 行最终应具有杠杆分数 w_i，因此得到“不动点”条件 (1.1)。近年来，p-范数刘易斯权重在优化和随机矩阵理论中有重要的算法应用，从 p-范数回归的降维（行采样 [CP14；Faz+21]），到线性规划最优障碍函数的设计 [LS14],，以及拉普拉斯矩阵的谱稀疏化（[SS11；KMP10]）。对于本文重点研究的 $\mathrm{p} = \infty$，ℓ_∞-LW 具有特别优美的几何解释——由 ℓ_∞-LW(A) 定义的椭球 $\mathcal{E} = \{\mathrm{x} \in \mathbb{R}^n \,|\mathrm{x}^T(\mathrm{A}^T\mathrm{W}\,\mathrm{A})^{-1}\mathrm{x} \le 1\}$，是包围点集 $\{\mathrm{a}_i\}_{i=1}^m$ 的最小体积椭球的对偶解，该椭球称为外约翰椭球：

$$\max_{M\succeq 0} \left\{\log\det M \;:\; a_i^T M a_i \le 1,\ \forall i \in [m]\right\}. \tag{1.2}$$

本文研究 ℓ_∞-刘易斯权重，将其作为设计和分析拉普拉斯矩阵的工具。由于杠杆分数（ℓ_2-刘易斯权重）在谱图算法 [ST03; SS11; KMP10], 中发挥着关键作用，因此自然会提出这样的问题：高阶刘易斯权重在图论和网络分析中扮演什么角色?

用图中各边的 ℓ_∞-刘易斯权重重新加权后，会带来哪些谱含义?

为了阐明这一动机以及 ℓ_∞-刘易斯权重在谱图论中的作用，将优化问题（1.2）解释为实验最优设计 [BV04], 的一个特例会很有帮助04], 其目标是找到固定线性测量 $\{a_i\}_{i=1}^m$的最优凸组合，从而为其最小二乘估计量诱导出最大置信椭球；这里的“最大”是相对于半正定矩阵上的某种偏序而言（例如洛文纳序，见附录 A.2）。对此类偏序，一个自然的选择是置信椭球的体积（行列式），这称为 *D*-最优设计：给定实验矩阵 $\{a_i \in\mathbb{R}^n\}_{i=1}^m$，（1.2）的原始问题为

$$\begin{array}{ll} \text{minimize} & \log\det\left(\sum_{j=1}^m g_j a_j a_j^T\right)^{-1} \\ \text{subject to} & \mathbf{1}^T g = 1\ ,\ g \ge 0. \end{array} \tag{D-optimal Design}$$

另一个自然的顺序 序是矩阵迹，通常称为*A*-最优设计 [BV04] :

$$\begin{array}{ll} \text{minimize} & \mathbf{Tr}\left(\sum_{j=1}^m g_j a_j a_j^T\right)^{-1} \\ \text{subject to} & \mathbf{1}^T g = 1\ ,\ g \ge 0, \end{array} \tag{A-optimal Design}$$

这是一个半定规划（SDP）。在这两个凸优化问题中，优化对象都是正的归一化权重向量 $g \in \mathbb{R}^m$，它们定义了点集 $\{a_i\}_{i=1}^m$的包围椭球。理解（D-最优设计）与（A-最优设计）之间的关系，是本文的核心主题。两种不同的目标函数具有自然的几何解释：（D-最优设计）旨在最小化椭球半轴长度的几何平均数，而（A-最优设计）旨在最小化其调和平均数。直观而言，这种调和平均数—几何平均数视角使这两个问题截然不同；事实上，我们证明，对于一般的 半正定矩阵，这两个目标之间存在 $\Omega(n/\log^2 n)$ 的差距：

定理 1.1（非正式表述）。 存在 $n \times n$ 个半正定实验矩阵 $A \geq 0$，使得 D-最优解（ℓ_∞-*LW(A)*）对于关于 A 的 A-最优设计问题，至多只能达到 $\Omega\left(n/\log^2 n\right)$-近似。

这个下界意味着，只有由于半正定矩阵的某些特殊结构，两个最小化问题之间才可能存在非平凡的关系。如前所述，本文关注的是图拉普拉斯矩阵。我们研究拉普拉斯矩阵的 A 对 D 最优设计，主要动机来自网络设计应用，因为在这种情况下，两个问题都对应于控制底层图的电流分布 和有效电阻。下面我们将解释这种对应关系。

电网络设计：最小化图上的有效电阻 给定一个无向图G，有效电阻 $R_{ij}(G)$是指在i 和j 两个端点之间施加单位电流源时，端点间出现的电势差（见定义 (3.3)）。直观地说，当两个节点之间有许多带高电导的边相连的路径时，i 与j 之间的有效电阻较小；当它们之间路径较少且电导较低时，有效电阻则较大。有效电阻在电网络分析、谱稀疏化 [SS11] 、拉普拉斯线性方程组求解器 [KMP10],、最大流算法 [Chr+11]、马尔可夫链的往返时间和覆盖时间 [Cha+96; Mat88],、连续时间平均网络、寻找稀疏树 [AG15], 以及生成随机生成树[KM09; MST15; Dur+17]等方面有许多应用。除此之外，图的有效电阻分布本身也是值得研究的自然图性质，正如 [Ale+17; GBS08] 所倡导的那样。

在一项开创性工作中，Ghosh、Boyd 和 Saberi [GBS08] 提出了有效电阻最小化问题（ERMP）：给定一个无权图G，寻找非负且归一化的边权重，使重加权图G_g中所有顶点对之间的有效电阻总和最小；该总和也称为基尔霍夫ff指数 $\mathcal{K}_g$（G） [LNT99]。该问题可以表示为以下半定规划:

$$\begin{aligned} \text{minimize} \quad & \mathcal{K}_g(G) := \sum_{i,j} R_{ij} \\ \text{subject to} \quad & g \geq 0,\ \mathbf{1}^T g = 1 \end{aligned} \tag{ERMP}$$

不难验证，ERMP 是一个关于拉普拉斯矩阵的 A-最优设计问题（见公式（ERMP））。ERMP 的这一表述具有如下解释——在G的节点上有实数x_1,.....、x_n，且它们的总和为零。G 中的每条边都对应一种可能的“测量”，其结果是相邻节点值之差，再加上噪声。给定用于估计$\boldsymbol{x}$的固定测量预算，问题是选择各条边测量所应占用的实验比例。最优比例正是 ERMP 权重。 [GBS08] 研究这一问题的动机，是改善底层网络的“电连通性”和混合时间。最小化电阻距离和电流，是路由网络（以及分子化学 [LNT99], ，该问题事实上就起源于此）中的重要基本操作；同时，它也可以被视为最大化重加权网络谱间隙和通信吞吐量的凸代理目标。因此，ERMP 可以被视为一种网络配置算法 [Hac+13；Mas+20]。

在 [GBS08], 中，作者设计了一种专用的内点法（IPM），可在 $O(m^{3.5})$ 时间内求解（ERMP）（更准确地说，需要求解 $\sim \sqrt{m}$ 个涉及拉普拉斯矩阵伪逆的通用线性方程组，而这些伪逆已不再是 SDD）。我们注意到，即使采用近期用于加速通用内点法的逆矩阵维护技术 [Jia+20；Hua+22], 最多也只能达到 $\widetilde{O}(mn^{2.5}+n^{0.5}m^{\omega})$，但这种加速（对于稀疏图而言）仅具有理论价值。无论哪种情况，上述运行时间对于大规模网络而言都相当可观，更不用说稠密网络了（$m \gg n$）。这促使我们转而采用近似算法。然而，对于 ERMP 半定规划而言，使用黑盒近似半定规划求解器（类似于 [AK16；Haz08] 的无投影算法）似乎天生过慢，这是由于其较大的宽度以及 MwU 的加性保证所致。[1]

[1]ERMP半定规划的宽度 为 $\geq \Omega(n^2)$，因此，基于乘法权重更新的近似半定规划求解器 [Haz08; AK16] 至少需要mn^2 时间才能实现非平凡的近似。此外，Hazan算法只能保证相对于目标矩阵的弗罗贝尼乌斯范数的加法 近似 $\epsilon\|C\|_F$，而在我们的情形下，该范数为 $\Omega(n^{1.5})$，见附录 A.6。

相比之下，ℓ_∞-刘易斯权重的近似计算成本低得多，尤其是对于拉普拉斯矩阵而言——尽管高精度计算约翰椭球（1.2）的速度不快于前述内点法（$\sim m^{3.5}\log(m/\epsilon)$[Nem99]），但对 ℓ_∞-LW 的低精度近似却便宜得多——[Coh+20；CP14]~提出了一种简单实用的近似算法，可达到任意小的精度 ϵ，其方法是进行 $O(1/\epsilon)$ 次重复的杠杆分数计算；对于拉普拉斯矩阵，这些计算可以在输入稀疏时间内完成 [ST03；KMP10]。幸运的是，对于 ERMP，低精度近似 ℓ_∞-LW 已经足够（见 §7 ）。我们指出，最近 [Faz+21] 给出了一种计算 ℓ_p-LW 的高精度算法，仅需进行 $\widetilde{O}(p^3\log(1/\epsilon))$ 次杠杆分数计算（在我们的情形下即为拉普拉斯线性方程组求解），因此令 $p = n^{o(1)}$，即可得到另一种 $O(m^{1+o(1)}\log(1/\epsilon))$ 时间的近似算法，对于大多数应用而言已经足够好 [Faz+21]。

与 ERMP 目标相比，可以证明，图拉普拉斯矩阵$L(G)$ 的 ℓ_∞-刘易斯权重可使最大有效电阻在边G上达到最小。事实上，我们证明了一个更强的结论：G 中各边的有效电阻，是加权拉普拉斯矩阵L_g 所诱导椭球的体积关于权重$g \in \mathbb{R}^m$ 的梯度。据我们所知，这为图的有效电阻（更一般地说，一般矩阵的统计杠杆分数，见 §3.4）提供了一种新的几何刻画：

引理 1.2. 有效电阻$(Gg) = -\nabla_g \log\det L_g^+$。

上述讨论以及 ℓ_∞-刘易斯权重近似的高效性，引出了以下自然问题：图的 ℓ_∞-刘易斯权重对基尔霍夫指数（最优 *ERMP* 权重）的近似效果如何？

本文的核心观点是，对边进行 ℓ_∞-刘易斯权重重新加权会对图的谱产生多种有利影响，并且能够为大规模（无向）网络提供一种高效的“预处理”操作，其作用远不止于基尔霍夫指数。我们将在下一小节总结主要发现。

1.1 主要结果

令 $\alpha_{A,\ D}(G)$ 表示 ℓ_∞-LW 针对 ERMP 目标（定义 (4.1)）所达到的近似比。我们的首个主要结果表明，对于树，$\alpha_{A,\ D}$ 是一个常数：

定理 1.3。 *ℓ_∞-LW 是树上 ERMP 问题的*3.12*-近似解* .

值得注意的是，存在一些树，使得$\alpha_{A,\ D} > 2.5$（见第 5 节），因此该界几乎是紧的。定理 1.3 的证明依赖一种专为树设计的新技术，它或许在电网络分析中也具有独立的研究价值——我们证明，总能对树进行局部修改，使近似比 $\alpha_{A,\ D}(T)$ 增大。更准确地说，我们的证明引入了一种基于树的 ℓ_2-拥塞的（有限）有效电阻“极化过程”，可反复应用该过程，从而得到最坏情况下的“极化”树族（蝴蝶结图，见图 1）。我们在第 2 节对证明作了高层次概述。

我们的第二个主要结果是关于一般图中 ℓ_∞-LW 近似比的上界。我们所得到的精确上界是谱参数 $\alpha_1(g)$ 的函数：$= \ ^2(n-1)^2\,\mathbf{Tr}L_g^+$，其中 G 的加权拉普拉斯矩阵是我们分析中的核心量。记 ~~$g_{\ell w}$~~为 ℓ_∞-LW 边权重，定义

$$\alpha_{min}(G) := \min\left\{\alpha_1(g_{\ell w})\ ,\ \left\|-\nabla_g\left(\log\alpha_1(g_{\ell w})\right)\right\|_\infty\right\}. \tag{1.3}$$

直观地说，注意到 t hat $\min\{x,\ (\log x)'\} = 1$，因此上述量彼此之间趋于反单调。

定理 1.4 （一般图的上界）。 对于任意无向图 G,

$$\alpha_{A,D}(G) \le \alpha_{min}(G).$$

我们证明，$\alpha_{min}(G)$ 始终不超过 $\le \min\left\{\mathrm{diam}(G),\ \kappa(L_{g_{\ell w}})\right\}$，其中 $\mathrm{diam}(G)$ 是G的直径，而 $\kappa(L_{g_{\ell w}})$ 是其 ℓ_∞-LW 加权拉普拉斯矩阵的条件数。虽然这已经能为低直径图提供良好的近似，但我们强调，α_{min} 通常能给出紧得多的上界

关于 $\alpha_{A,\ D}$：棒棒糖图（一个大小为n 的团与一条长度为n的路径相连）兼具 $\Omega(n)$ 的直径和 LW 条件数，但模拟表明 α_{min}(棒棒糖$_n$) $\leq O(\log n)$。在第 5 节中，我们展示了不同图族的近似比 α_{min}，并表明在实践中它以$O(1)$ 的速度增长。结合定理 1.3、1.4 以及我们的实证结果，我们猜想，对于任意图，ℓ_∞-LW 都能为基尔霍夫指数提供 $O(\log n)$ 近似：

猜想 1.5（Lewis 遇见 Kirchoff）。 $\forall G,\ \alpha_{A,\ D}(G) \leq O(\log n)$。

一个更强的猜想是 $\alpha_{\min}(G) \leq O(\log n)$；事实上，我们的分析和实验都表明，这个更强的猜想成立（见第 5 节）。我们没有发现 $\alpha_{\min}(G)$ 与$\alpha_{A,\ D}(G)$ 之间存在超常数差距，而 $\alpha_{\min}(G)$（1.3）是否为紧上界则是一个引人入胜的问题。

谱含义： **ℓ_∞-LW** 我们的第三组结果展示了 ℓ_∞-LW 重加权对图拉普拉斯矩阵特征值分布的有利影响。我们沿着这一方向给出两个基本结果：

定理 1.6（ℓ_∞-LW 重加权与混合时间，非正式表述）。

1. 我们将 *[Moh+91]* ($\lambda_2 \geq 4/nD$) 的经典谱隙界推广到加权图：$\lambda_2(G_w) \geq 2/(nD \cdot R_{\max}(G_w))$，其中 $R_{\max}$ 是加权图中一条边 的最大有效电阻。*2.* 我们揭示了图*[NES03]* 的平滑最大特征值的最优性条件（$\max_g \log(\boldsymbol{Tre}^{L_g^+})$）与 ℓ_∞-*LW* 的不动点条件（1.1）之间的紧密联系，这表明了 ℓ_∞-*LW* 重加权提高 *Softmax* 函数的一个条件。

谱稀疏树 **[AG15]** 在一项具有影响力的成果中，[Asa+10] 提出了一种针对非对称旅行商问题的舍入方案，突破了$O(\log n)$近似界，前提是底层图存在一棵谱稀疏 树（STT）（$L_T \leq \gamma L_G$，其中 $\gamma = \gamma_G(T)$ 表示该树的谱稀疏度）。Anari 等人 [AG15] 证明，对于任意图，$\gamma_G(T)$ 至少为 $R^{\max_e}(G)$，并且可以在多项式时间内找到一棵谱稀疏度为 $\sim\Theta(R^{\max_e}(G))$ 的最优树。我们将这一结果推广如下：对于任意图 G，我们都能找到一棵加权 树 T_w，其谱稀疏度为 $L_{T_w} \leq O((n-1)/m)L_G$：

引理 1.7。 **</style}对于任意连通图 G V,** $E\rangle$，存在一棵加权生成树 T_w，使得 T_w 的谱稀疏度为 $((n-1)/m) \cdot O(\log n/\log\log n)$。

我们证明，这始终不逊于 [AG15]树。值得探究的是，这种加权版的谱稀疏生成树能否用于非对称旅行商问题的舍入原语。我们相信，我们的研究成果将为谱图算法的设计与分析带来更多应用。

2 技术概述

这里我们将概述证明定理 1.3 和 1.4 所需的主要新思路。

ER 极化：定理 1.3 的技术概述我们用于界定树上的 ERMP 近似比（$\alpha_{A,\ D}(T)$）基于以下重要观察：树中第l条边的最优权重$g_l^*(T)$）与其拥塞$c_l(T)$）成正比（4.4）。因此，当树存在严重的“瓶颈”（即许多路径都经过的一条边）时，最优权重就会“远离”均匀分布（因而远离 LW 解（4.1））。基于这一观察，我们证明，对于任意树T，都可以构造一棵替代树T'，使其近似比更差，即 $\alpha_{A,\ D}(T') > \alpha_{A,\ D}(T)$。我们的构造是迭代进行的，利用上述观察——每次迭代都会加剧“瓶颈”

直到达到一个“不动点”（蝴蝶结图）。为了正式定义这一极化过程，我们引入树上的局部变换[2] (LT)这一概念。LT 是一种局部图操作，会改变树中单条边的拥塞。形式化地，我们用E_k 表示一个 LT，使得

$$c_l(E_k \circ T) = c_l(T),\ \forall\ l \neq k, \tag{2.1}$$

其中，$T' = E_k \circ T$ 是变换后的树。

增加树的瓶颈（即使式（4.4）中的分布进一步偏离均匀分布）有两种自然方式：(1) 减少不同路径的数量，或 (2) 增加更多入口点（即叶节点）。事实证明，这两种操作都可以表述为 LT。这样，我们就可以通过反复使用 LT 逐步构造新树。为此，我们将定义一个上部 LT 和一个下层 LT，分别记为$E^{\uparrow}{}_k$ 和$E^{\downarrow}{}_k$，使得对于任意树T，都有：

$$\begin{aligned} c_k(E_k^{\uparrow} \circ T) &> c_k(T), \\ c_k(E_k^{\downarrow} \circ T) &< c_k(T). \end{aligned} \tag{2.2}$$

（下文将会明确这一点。）上部 LT 对应于上述 (1)，即减少路径数量；下层 LT 对应于上述 (2)，即增加另一个入口点，这与我们最初的直觉一致。事实证明，决定下一步执行哪种局部操作的确切阈值，是最优权重的范数 – $\|g^*(T)\|_2^2$：

命题 2.1. 设 T 是一棵 n 阶树，E_k 是一个局部变换。那么 $\alpha_{A,\ D}(T) \leq \alpha_{A,\ D}(E_k \circ T)$ 满足以下任一条件时：

$$E_k \textit{ is a lower LT } \ \&\ g_k^*(T) \leq \|g^*(T)\|_2^2$$

or,

$$E_k \textit{ is an upper LT } \ \&\ g_k^*(T) > \|g^*(T)\|_2^2$$

根据这个命题，我们将树的边分成两组：

$$E_<(T) := \{l \mid g_l^*(T) \leq \|g^*(T)\|_2^2\},\ E_>(T) := \{l \mid g_l^*(T) > \|g^*(T)\|_2^2\}$$

并对每组进行相应的局部变换（直至饱和——$E_k \circ T = T$）。这一过程之所以可行，是因为这些集合有一个关键特性——它们在局部变换下保持不变。更准确地说，我们证明，对于任意树T和$k \in E_>$，都有$E_>(E^{\uparrow}{}_k \circ T) = E_>(T)$，反之对于$E_<$也成立。

我们反复应用这一过程（每次迭代都选择“正确的”局部变换 LT），并证明它必定会在有限步后终止，得到一棵拥塞分布尽可能“极化”的“不动点”树——我们将这类树称为某一阶的蝴蝶结图（见图 1）。我们证明，蝴蝶结图在所有树中使比值 $\alpha_{A,D}(T)$ 达到最大值，随后利用为一般图所开发的工具，对蝴蝶结图中的后一数量给出上界（为≈ 3.12）；相关内容将在下一段详述。

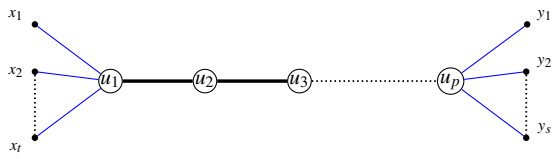


图 1：蝴蝶结图 - $\mathcal{B}_{t,\ p,\ s}$。一条长度为p 的路径，两端分别连接大小为t、s 的星图

[2]这个想法 受到矩阵初等变换的启发；事实上，在某种意义上，它们发挥着相似的作用 .

一般上界概述（定理 1.4） 对于为何ℓ_∞-LW(G) 能为任意图G上的 ERMP 问题提供一个O(直径)-近似，即 $\alpha_{A,D} \le$diam(G)，有一个简单直观的解释。事实上，由于 LW 使G 的边 中的最大有效电阻最小（见下一节），而有效电阻众所周知构成一个度量，因此三角不等式意味着，对于G中任意一对顶点，有效电阻$(i,j)_{lw}$ 至多等于连接i 和j的最短路径 上各边的有效电阻之和，而该和至多为 diam(G_{lw}) = diam(G)。我们利用算术几何平均数论证证明一个更强的上界：对于任意G,

$$\alpha_{A,D}(G) \le \alpha_1(g_{\ell w}), \tag{2.3}$$

其中， $\alpha_1(g) := \frac{2}{(n-1)^2}\mathbf{Tr}L_g^+$是一个与最优解 $\mathcal{K}_G(g)$ 密切相关的参数——事实上，引言中提到的调和平均数刻画表明，最优 ERMP 值可以改写为（见公式（3.7））

$$\mathcal{K}(g^*) = n\,\mathbf{Tr}\,L_{g^*}^+ = \frac{n(n-1)}{HM(\lambda(L_{g^*}))},$$

其中，调和平均数($\lambda(L_{g^*})$) 是加权拉普拉斯矩阵的 $n-1$ 个（正）特征值的调和平均数，这里的加权拉普拉斯矩阵采用最优 ERMP 权重。经过进一步推导，我们可以利用算术-几何平均不等式证明，近似比 $\mathcal{K}(g_{lw})/\mathcal{K}(g^*) \le \alpha_1(G)$ 始终不超过直径 diam(G)（根据三角不等式），但通常会更小，参见实验部分。

如果G 的直径很大怎么办？在这种情况下，我们证明 $\alpha_1(G)$ 的“对偶”函数通常很小。具体而言，我们证明以下量也是近似比的上界：

$$\alpha_{A,D}(G) \le \left\|-\nabla_g\left(\log \alpha_1(g_{\ell w})\right)\right\|_\infty. \tag{2.4}$$

这个上界在技术上更复杂，也不那么直观，但推导它的关键在于 [GBS08], 的 ERMP对偶间隙 ，它自然涉及关于权重的梯度次优条件。接着，利用微积分和 Courant-Fisher 原理，我们证明该值至多为 $\kappa(L(G_{lw}))$。

结合 (2.3) 和 (2.4) 可得定理 1.4：$\alpha_{A,\ D}(G) \le \min\left\{\alpha_1(g_{\ell w}),\ \left\|-\nabla_g\left(\log\ \alpha_1(g_{\ell w})\right)\right\|_\infty\right\}$。如前文引言中所述，前面提到的两个量的最小值 总是很小，其直觉源自标量不等式 $\min\{x、(\log x)'\} = 1$。事实上，我们的所有~模拟都证实最小值为O(1)，因此 $\alpha_{min}(G)$ 通常比 $\min\{$diam(G)、 $\kappa(L(G_{lw}))\}$ 给出的上界紧得多。猜想 4.6 提出，最小值始终为O(log$n$)。

本文结构： 在第 3 节中，我们介绍刘易斯权重、拉普拉斯矩阵和图的有效电阻的背景知识，并证明我们对杠杆分数和有效电阻的新刻画（引理 1.2）。在第 4 节中，我们证明定理 1.3 和 1.4。第 5 节总结实验结果。在第 6 节中，我们证明定理 1.6，并给出其在谱稀疏树中的应用。最后，在第 7 节中，我们重述计算刘易斯权重的算法，并说明低精度的刘易斯权重足以满足我们的结果。第 8 节中，我们探讨 A-最优设计与 D-最优设计之间的关系，以毕达哥拉斯平均数给出几何解释，并证明定理 1.1。

3 预备知识

3.1 符号

我们将$\boldsymbol{S}_+^n$ 记为 $\mathbb{R}^{n\times n}$的对称矩阵子空间。单位向量e_i是除i'位置的分量为 1 之外、所有分量均为 0 的向量。我们将矩阵$M \in \mathbb{R}^{n\times n}$ 的特征值（EV）记为

$$\lambda(M) = \lambda_1(M) \le \lambda_2(M) \le \cdots \le \lambda_n(M).$$

Fo对于任意矩阵$A \in \mathbb{R}^{n\times d}$，我们将*projection* 矩阵记为投影到$A$ a的列空间上的矩阵 s

$$\Pi_A = A(A^TA)^+A^T.$$

为便于后续讨论，我们在此重述算术-几何平均不等式。对于任意由x组成的序列，其包含n 个正数，定义：

$$AM(X) = \frac{1}{n}\sum_i x_i\,, GM(X) = \sqrt[n]{x_1\dots x_n}\,,\ HM(X) = \frac{n}{\sum_{i=1}^n x_i^{-1}}\,,$$

满足以下不等式：

$$HM(X) \le GM(X) \le AM(X),$$

当且仅当X 为常数序列时，等号成立。

我们采用以下椭球定义——椭球由 $\{v \in \mathbb{R}^n \mid v^TAv \le 1\}$定义，其中$A$是对称半正定矩阵。椭球的半轴由 $\boldsymbol{\sigma} = \{\sigma_i(\mathcal{E})\}_{i=1}^n$ 给出，且等于

$$\sigma_i = \lambda_i(A)^{-1/2}. \tag{3.1}$$

由此可 以清楚地看出，椭球的体积等于（相差一个常数因子）– $\mathrm{vol}(\mathcal{E}) \propto |A|^{-1/}$ ^{2}.

3.2 杠杆分数与 Lewis 权重

对于矩阵 $A \in \mathbb{R}^{n\times d}$，我们将第 i 行A的杠杆分数 (LS) 定义为

$$\tau_i(A) = a_i^T(A^TA)^+a_i,$$

其中，a_i 是A的第 i 行。我们称 $\tau(A) = \{\tau_i(A)\}_{i=1}^n$ 为A的杠杆分数。注意， $\tau(A)$ 恰好是投影矩阵 Π_A的对角线。由于 $0 \le \Pi_A \le I$，因此 $0 \le \tau_i(A) \le 1$，并且：

事实 3.1。（福斯特引理， [Fos53]） $\sum \tau_i(A) =$ **迹** $\Pi_A =$ 秩(A)。$_i$

杠杆分数传统上用作草图化和 ℓ_2-子空间嵌入 [CW13]的工具。事实上，基于 LS 对矩阵进行行采样，已知可以得到良好的谱近似 [CW13；Dri+11]。这也是谱图稀疏化 [SS11]背后的核心思想。

如引言中所述，我们关注的是 ℓ_∞-LW，它是杠杆分数的一种推广：

定义 3.2. (ℓ_∞-LW) 对于矩阵 $A \in \mathbb{R}^{n\times d}$，$A$ 的 ℓ_∞-LW 记作 $w_\infty(A) \in \mathbb{R}^n$，它是唯一的权重向量，使得对于每一行$i \in [n]$，都有

$$(w_\infty)_i = \tau_i(W_\infty^{1/2}A),$$

或者等价地，

$$a_i^T(A^TW_\infty A)^+a_i = 1,$$

其中 $W_\infty =$ =diag(w_∞)。从现在起，我们用 LW 表示 ℓ_∞-LW$(A) := w_\infty(A)$。

注意，定义（3.2）是循环的，而且 ℓ_∞-LW 是否存在并非先验显而易见。证明其存在且唯一的一种方法是反复对所得矩阵进行最小二乘计算，直到达到不动点，见 §7。注意，对于任意矩阵A，利用事实（3.1），我们有$w_\infty(A) \le \mathbf{1}$；并且，由于$W_\infty$是一个满对角矩阵，即 rank$(W_\infty^{1/2}A) =$ rank(A)，因此 $\sum_{i=1}^m (w_\infty(A))_i =$ rank(A)。

3.3 图的有效电阻

本节将正式定义图的有效电阻（ER），并推导一些后续将用到的重要性质。更多详情参见 [GBS08]。设 G = ⟨V, E⟩ 是一个有 n 个顶点和 m 条边的无向图。不失一般性，V = [n]。我们始终假设 G 是连通的且没有自环（如果 G 不连通，我们可以分别处理每个连通分量）。u ∈ V 的度记为 d(u) := d_G(u)。G 的邻接矩阵记为 A = A_G。对于每条边 l = (i, j)，若 i < j，则定义一个（任意的）定向，使得 $b_l \in \mathbb{R}^n$ 的 $(b_l)_i = 1$、$(b_l)_j = -1$，其余所有元素均为 0。$B = \{b_l\}_{l=1}^m \in \mathbb{R}^{n\times m}$ 称为 G 的边关联矩阵。

本文探讨了对图的边进行重新加权对各种图函数的影响。重新加权后的图 $G = \langle V, E, w\rangle$由相应的权重向量 $g \in \mathbb{R}^m$ 定义，使得 $g_l = w(e_l)$。

相应的 加权拉普拉斯矩阵$L_g = L_g(G)$（相对于权重向量 $g \in \mathbb{R}^m$）是半正定矩 ix

$$L_g := B \cdot \mathrm{diag}(g) \cdot B^T.$$

根据上下文需要，我们有时将$W = W_g :=$ 记为 diag(g)。等价地，$L_g = D_g - A_G(g)$，其中$A_G(g)$是G 的加权邻接矩阵，而D_g 是对角矩阵，其$d_g(u) = \sum_{v\sim u} w(u,v)$位于第$u$个对角项。注意，无权情形对应于$g = \mathbf{1} \in \mathbb{R}^m$。

关于拉普拉斯矩阵的一个重要事实是，L 的秩为$n - 1$，且$L\mathbf{1} = 0$。容易验证$L + (1/n)\mathbf{1}\mathbf{1}^T$ 可逆，且其逆矩阵等于，

$$(L + (1/n)\mathbf{1}\mathbf{1}^T)^{-1} = L^+ + (1/n)\mathbf{1}\mathbf{1}^T, \tag{3.2}$$

其中 L^+ 是 L 的伪逆。

在全文中，我们始终假设 g 已归一化，使得 $\mathbf{1}^T g = 1$。这种归一化的一个好处是，$AM(\lambda(L_g)) = {}^2_{n-1}$。要理解原因，请回想 $L_g = D_g - A(g)$。因此

$$\mathbf{Tr}\, L_g = \mathbf{Tr}\, D_g - \mathbf{Tr}\, A(g) = \mathbf{Tr}\, D_g = \sum_i \sum_{j\sim i} w(i,j) = 2 \cdot \sum_{e\in E} w(e) = 2 \cdot \mathbf{1}^T g = 2.$$

L_g 只有 $n - 1$ 个正特征值，因此

$$AM(\lambda(L_g)) = \frac{1}{n-1}\sum_{i=2}^{n} \lambda_i = \frac{1}{n-1}\mathbf{Tr}\, L_g = \frac{2}{n-1}. \tag{3.3}$$

我们稍后会用到这一点。

在全文中，我们将使用以下简写表示给定图G的 ℓ_∞-Lewis 权重

$$g_{lw} := \frac{1}{n-1} w_\infty(B^T). \tag{3.4}$$

这种归一化源于事实 (3.1) 以及 rank(B^T) $= n - 1$。

有了这些定义，我们现在可以正式定义图的有效电阻：

定义 3.3 (有效电阻)。 给定一个加权图 G，其拉普拉斯矩阵为 $L_g = BWB^T$，图中一对端节点（i，j）之间的有效电阻（ER）为：

$$ER(i,j) = b_{ij}^T L_g^+ b_{ij}$$

其中，$b_{ij} = e_i - e_j$，且L^+是L的伪逆。

在本文中，我们有时将两个顶点i 和j 之间、权重向量为g 时的有效电阻记作$R_{ij}(g)$。我们将$G(g)$中第l条边上的有效电阻记作

$$R_l(g) = b_l^T L_g^+ b_l \ , \ l = 1 \dots m$$

当权重 g 明确时，我们将写作 R_l。由定义和伪逆的性质可直接得出一个有用的性质：有效电阻是关于g 的 -1 次齐次函数，即

$$R_{ij}(\alpha g) = R_{ij}(g)/\alpha. \tag{3.5}$$

有效电阻的另一个重要性质是，它在图 [KR93], 上定义了一个度量，因此满足三角不等式（对任意权重 g）：

$$R_{ij} \le R_{ik} + R_{kj} \ \forall i, k, j \in V \tag{3.6}$$

我们的论文延续了 [GBS08] 关于最小化图的总有效电阻（ERMP）的研究，该问题也称为图G [LNT99] 的基尔霍夫指数 $\mathcal{K}_G(g)$。 $\mathcal{K}_G$ 也可以表示为

$$\begin{aligned} \mathcal{K}_G(g) :=& \sum_{i,j} R_{ij}(g) \\ =& n \, \mathbf{Tr} \, L_g^+ \\ =& n \, \mathbf{Tr}(L_g + (1/n)\mathbf{1}\mathbf{1}^T)^{-1} - n \end{aligned} \tag{3.7}$$

其中，第一个等式使用了定义和迹的运算法则，第二个等式使用了公式（3.2）。利用上述表达式，我们可以将 ERMP 重写为如下半定规划：

$$\begin{array}{ll} \text{minimize} & \mathbf{Tr} \, X^{-1} \\ \text{subject to} & X = \sum_{l=1}^m g_l b_l b_l^T + (1/n)\mathbf{1}\mathbf{1}^T \\ & \mathbf{1}^T g = 1 \end{array} \tag{ERMP}$$

这是 A-最优设计的一个特例（见（A-最优设计）），即针对拉普拉斯矩阵 的设计（其中关联矩阵B作为实验矩阵）。

3.4 杠杆分数与有效电阻的新刻画 s

在证明结果之前，我们先给出杠杆分数（LS）和有效电阻（ER）的一种新几何刻画。如引言中所述，D-最优设计的最优解恰好是实验矩阵 [CP14; Coh+20] 的 LW。我们的刻画源于对后一结果的一种不同且据我们所知全新的证明，该证明使用了基本的凸优化分析。为简洁起见，这里仅提供概述，详细内容留待附录 A.3。设 B 为某个图的边关联矩阵，G 为该图，L_g 为$G(g)$ 的加权拉普拉斯矩阵，并将我们的目标（见备注 A.2）记为

$$LD_G(g) := \text{logdet}(L_g^+ + (1/n)\mathbf{1}\mathbf{1}^T).$$

在 (A.6) 中，我们证明第l条边上的有效电阻等于

$$R_l(g) = b_l^T L_g^+ b_l = b_l^T (B W_g B^T)^+ b_l = -\frac{\partial LD(g)}{\partial g_l}.$$

换言之，$LD_G(g)$（关于g）的梯度等于各边上的有效电阻（相差一个负号）。这也可以推广到任意矩阵的杠杆分数：令$V \in \mathbb{R}^{n \times m}$，并定义$C = W_g^{1/2} V^T$，即

加权实验矩阵。定义$E_V(g) = VW_gV^T$，以及$LD_V(g) = \text{logdet}(E_V(g)^{-1}) = -\text{logdet}(E_V(g))$。那么$l$'个$C$的最小二乘等于

$$\tau_l(C) = -g_l \cdot \frac{\partial LD_V(g)}{\partial g_l} \qquad \text{see (A.8)} \tag{3.8}$$

也就是说，杠杆分 数是加权矩阵的加权梯度，即$LD_V(g)$的加权梯度（相差一个负号）。

这一刻画有如下几何解释：我们知道，D-最优设计等价于在V上寻找最小体积包围椭球。这意味着$LD_V(g)$的梯度与体积的梯度成正比$\mathcal{E}$。因此，(3.8) 意味着：

引理 3.4。矩阵 V 的杠杆分数等于由 $E_V(g)$ 所诱导椭球的体积梯度（相差一个常数因子）。特别地，对于拉普拉斯矩阵，加权图各边上的有效电阻等于由 L_g 所诱导椭球的体积梯度。

事实上，对于拉普拉斯矩阵，有一个更强的结论，即LD_G 的最优性准则变为

$$-\frac{\partial LD_G(g)}{g_l} = R_l(g) \le \text{rank}(B) = n-1. \qquad \text{see (A.10)}$$

此外，在$g = g_{\ell w}$处，我们使这一条件达到等号（见公式 (A.12)），这意味着

$$R_l(g_{\ell w}) = n-1\ ,\ \forall\ l = 1 \ldots m$$

现在，根据$LD(g)$最优解的唯一性，我们知道，对于任何其他归一化权重向量g，最优性准则都不成立。换言之，对于任意 $\widehat{\ }g \neq g_{\ell w} \in \mathbb{R}^m$，使得$\mathbf{1}^T\widehat{\ }g = 1$，都存在$l \in [m]$，使得

$$R_l(\widehat{g}) > n-1.$$

由此可得，ℓ_∞–LW 在所有（归一化）权重向量中，使边上的最大有效电阻最小。

4 通过 ℓ_∞-Lewis 权重进行 ERMP 近似

回顾一下，图的基尔霍夫指数 $\mathcal{K}_G^*$ 是 ERMP 问题（ERMP）的最优解。本节中，我们用 $\mathcal{K}_G^{\ell w}$表示 ERMP 在$g_{\ell w}$处的值。ℓ_∞-LW 所达到的近似比如下定义：

$$\alpha_{A,D} := \frac{\mathcal{K}_G^{\ell w}}{\mathcal{K}_G^*}. \tag{4.1}$$

回顾一下，$g_{\ell w}$ 是 D-最优设计的最优解，因此 $\alpha_{A,\ D}$ 准确刻画了拉普拉斯矩阵上的 A-最优设计与 D-最优设计之间的差距。由于我们提出将 ℓ_∞-LW 用作 ERMP 的近似算法，$\alpha_{A,\ D}$ 便是我们算法的近似比。

本节大致分为两个相互独立的结果。我们首先只关注树，从而证明定理 1.3。对于树，ERMP 具有更简单的形式，我们采用了一种专门的技术来证明定理。第二个结果尝试将结论推广到一般图。我们给出更一般的分析，并展示近似比的两个不同上界。我们猜想它们能够给出一个$O(\log n)$ 上界，进而推出猜想 1.5。最后，我们通过展示一些与该上界有关的公式来结束本节。

4.1 ℓ_∞-LW 是树的$O(1)$-近似

我们将本节分为两部分。首先，我们利用 [GBS08] 推导出的最优权重公式，特别针对树重新改写近似比。然后，我们给出定理 1.3 证明的技术细节。

我们首先证明，对于树$g_{\ell w} = g_{uni} = (1/m)\mathbf{1}$:

断言 4.1。 对于任意树 T，都有 $g_{\ell w} = g_{uni}$。

证明。 为简单起见，我们先考察非归一化权重 $w_\infty = w_\infty(B^T)$。我们知道$w_\infty \le \mathbf{1}$(公式（3.1）)。此外，由（3.1）可得

$$\sum_{l=1}^{m}(g_{\ell w})_l = \underbrace{n-1=m}_{\text{for trees}}.$$

因此，有m 个正数，它们均至多为 1 且总和为m，必然全部等于 1。因此，$(g_{\ell w})_l = 1, \forall\, l = 1。\ .\ .\ m$，所以除以$n - 1 = m$进行归一化后，我们得到$g_{\ell w} = g_{uni}$，如断言所述。

□

接下来，我们为树上的 ERMP 重述 [GBS08] 中的一些关键公式。首先，当图是一棵树时，对于任意两个顶点i、j，它们之间的有效电阻为

$$R_{ij}(g) = \sum_{e_l \in P(i,j)} \frac{1}{g_l}$$

从电网络的角度来看，这一点显而易见[3]。换句话说，第l条边会对所有经过它的路径贡献 $(1/g_l)$。基于此，我们将第l条边的拥塞度 定义如下:

定义 4.2 （边的拥塞度）。 边$e_l \in T$的拥塞度记为$c_l(T)$，表示T 中经过该边的路径数量。

容易看出

$$c_l = n_l(n - n_l),$$

其中n_l 是e_l 一侧子树 中的节点数（另一侧同理）。注意，由函数$f(x) = x(n-x)$的凹性可知，最小拥塞度在T的叶节点处取得（$1(n-1) = m$），而最大拥塞度为$n2\ (n - n2) = (n/2)^2$。更一般地，

$$f(x) = x(n-x) < y(n-y) = f(y) \iff \min(x, n-x) < \min(y, n-y). \tag{4.2}$$

根据此定义及上述事实，对于任意树T和任意权重向量$g \in \mathbb{R}^m$,

$$\mathcal{K}_T(g) = \sum_{l=1}^{m} \frac{c_l}{g_l}\ .$$

此外， [GBS08] （第 5.1 节）证明，对于树，最优 ERMP 解为

$$\mathcal{K}_T^* = \left(\sum_l c_l^{1/2}\right)^2, \tag{4.3}$$

[3]也就是说，串联电阻器的有效电阻具有可加性。

且最优权重的闭式表达式为

$$g_l^* = (c_l/\mathcal{K}^*)^{1/2} = \frac{c_l^{1/2}}{\sum_k c_k^{1/2}} \quad \forall l = 1, \ldots, m. \tag{4.4}$$

利用前面的表达式，可以立即得到

$$c_l = \mathcal{K}^* \cdot (g_l^*)^2 \implies \sum_l c_l = \mathcal{K}^* \cdot \sum_l (g_l^*)^2 = \mathcal{K}^* \cdot \|g^*\|_2^2. \tag{4.5}$$

最后，我们的证明将用到最优权重至多为 1/2 这一事实。确实，

$$g_{max}^* \le \frac{c_{max}^{1/2}}{\min\{\sum_l c_l^{1/2}\}} \le \frac{n/2}{m\sqrt{m}} \le \frac{1}{\sqrt{m}} \ll \frac{1}{2}, \tag{4.6}$$

当我们使用了式 (4.2) 中的界时。

利用$g_{\ell w} = g_{uni}$ （断言 4.1）这一事实，我们可以将 $\mathcal{K}_T^{\ell w}$写成：

$$\mathcal{K}_T^{\ell w} = \sum_{l=1}^{m} \frac{c_l}{(g_{\ell w})_l} = \sum_{l=1}^{m} \frac{c_l}{1/m} = m \sum_l c_l$$

因此，对于任意阶为n的树T_n，我们可以用拥塞来表示近似比：

$$\alpha_{A,D}(T_n) = \alpha(T_n) = \frac{m \sum_l c_l}{\left(\sum_l c_l^{1/2}\right)^2} \tag{4.7}$$

在定理 1.3 的证明中，我们还将用到下面的断言：

断言 4.3。 设 T，T'是两棵树，满足 $T \subset T'$。那么 $\alpha(T) \le \alpha(T')$。

证明。 回顾 $\alpha(T)$ 的定义：

$$\alpha(T) = \frac{\mathcal{K}_T^{\ell w}}{\mathcal{K}_T^*} = \frac{m \cdot \sum_l c_l}{\mathcal{K}_T^*}$$

显然， $\mathcal{K}^{\ell}{}_{wT} < \mathcal{K}_{T'}^{\ell w}$（我们只添加正项）。此外， [GBS08]证明了

$$T \subset T' \implies \mathcal{K}_{T'}^* \le \mathcal{K}_T^* .$$

（这可以看作 T 是 T' 的可行最优权重。）由此不难看出

$$\alpha(T) \le \alpha(T')$$

□

4.1.1 定理 1.3 的证明

我们首先说明局部变换何时会提高近似比。

命题 *2.1* 的证明。 记$T' = E_k \circ T$、$c'_k = c_k(T')$（其余拥塞保持不变）。回忆一下， $\alpha = (m\sum_l c_l)/\left(\sum_l c^{1/2}{}_l\right)^2$，因此我们将改写后的分子和分母记下来。

$$\sum_l c'_l = \sum_l c_l + (c'_k - c_k).$$
$$\sum_l c'^{1/2}_l = \sum_l c^{1/2}_l + (c'^{1/2}_k - c^{1/2}_k).$$

我们得到：

$$\begin{aligned}\alpha(T) \le \alpha(T') &\iff \frac{\sum_l c_l}{\left(\sum_l c_l^{1/2}\right)^2} \le \frac{\sum_l c_l + (c'_k - c_k)}{\left(\sum_l c_l^{1/2} + (c'^{1/2}_k - c^{1/2}_k)\right)^2} \\ &\iff \frac{\sum_l c_l}{\left(\sum_l c_l^{1/2}\right)^2} \le \frac{\sum_l c_l + (c'_k - c_k)}{\left(\sum_l c_l^{1/2}\right)^2 + \left(c'^{1/2}_k - c^{1/2}_k\right)^2 + 2\left(\sum_l c_l^{1/2}\right)\left(c'^{1/2}_k - c^{1/2}_k\right)}\end{aligned}$$

利用以下事实，并假设分母为正（在我们的情形中确实如此），

$$\frac{A}{B} \le \frac{A+D}{B+C} \iff AC \le BD,$$

我们得到

$$\begin{aligned}\alpha(T) \le \alpha(T') &\iff \left(\sum_l c_l\right)\left[\left(c'^{1/2}_k - c^{1/2}_k\right)^2 + 2\left(\sum_l c_l^{1/2}\right)\left(c'^{1/2}_k - c^{1/2}_k\right)\right] \le (c'_k - c_k)\left(\sum_l c_l^{1/2}\right)^2 \\ &\iff \left(\sum_l c_l\right)\left(c'^{1/2}_k - c^{1/2}_k\right)\left(2\sum_l c_l^{1/2} + c'^{1/2}_k - c^{1/2}_k\right) \le (c'^{1/2}_k - c^{1/2}_k)(c'^{1/2}_k + c^{1/2}_k)\left(\sum_l c_l^{1/2}\right)^2\end{aligned}$$

我们将证明的其余部分分为两种情况。

情形1： E_k 是一个上部 LT。根据定义，$c'_k > c_k$。于是，$c'^{1/2}{}_k - c^{1/2}{}_k > 0$，因此我们可以除以后一项，得到：

$$\begin{aligned}\alpha(T) \le \alpha(T') &\iff \left(\sum_l c_l\right)\left(2\sum_l c_l^{1/2} + c'^{1/2}_k - c^{1/2}_k\right) \le (c'^{1/2}_k + c^{1/2}_k)\left(\sum_l c_l^{1/2}\right)^2 \\ &\iff \|g^*\|_2^2\left(2\frac{c_k^{1/2}}{g_k^*} + c'^{1/2}_k - c^{1/2}_k\right) \le (c'^{1/2}_k + c^{1/2}_k) && \text{使用 (4.5)，并对第}k\text{条边使用 (4.4)。} \\ &\iff 2\frac{c_k^{1/2}}{g_k^*} - c_k^{1/2}\left(1 + \frac{1}{\|g^*\|_2^2}\right) \le c'^{1/2}_k\left(\frac{1}{\|g^*\|_2^2} - 1\right) && \text{按 } c^{1/2}{}_k \text{ 和} c'^{1/2}{}_k \text{对项分组} .\end{aligned}$$

在我们的情形中，$c'_k/c_k > 1$，因此两边除以$c^{1/2}{}_k$后，只需满足

$$\frac{2}{g^*_k} - 1 - \frac{1}{\|g^*\|_2^2} < \frac{1}{\|g^*\|_2^2} - 1 \iff g^*_k > \|g^*\|_2^2.$$

因此我们得到，

$$g^*_k > \|g^*\|_2^2 \implies \alpha(T) \le \alpha(T').$$

情形2： E_k是一个下层 LT。使用相同的论证，我们得到，

$$g^*_k \le \|g^*\|_2^2 \implies \alpha(T) \le \alpha(T')$$

将两种情形统一起来，即完成证明。 □

请记住，我们定义了如下划分：

$$E_<(T) \coloneqq \{l \mid g^*_l(T) \le \|g^*(T)\|_2^2\},\ E_>(T) \coloneqq \{l \mid g^*_l(T) > \|g^*(T)\|_2^2\}$$

现在我们来展示 该划分的关键特征——$E_<$和$E_>$在上层和下层 LT 下均保持不变：

命题 4.4. 对于任意一棵树 T 和 $k \in E_>$，我们有：

$$E_>(E_k^{\uparrow} \circ T) = E_>(T).$$

对于 $E_<$，反之亦然。

证明。 我们只证明 $E_>$ 的情形，因为 $E_<$ 的证明是对称的。设 $k \in E_>(T)$。记$T' = E^{\uparrow}{}_k \circ T$。我们首先证明 $g^*_k(T') > g^*_k(T)$。由公式 (4.4) 可知

$$g^*_l(T) = \left(\frac{c_l(T)}{\mathcal{K}^*_T}\right)^{1/2},\ g^*_l(T') = \left(\frac{c_l(T')}{\mathcal{K}^*_{T'}}\right)^{1/2}$$

此外，根据上部 LT 的定义，第k条边上的拥塞会增加，而其他任何边上的拥塞都不会改变，这意味着：

$$\mathcal{K}^*_{T'} = \left(\sum_l c_l(T')^{1/2}\right)^2 > \left(\sum_l c_l(T)^{1/2}\right)^2 = \mathcal{K}^*_T.$$

结合前两个公式可知，对于所有$l \ne k$，有：

$$c_l(T') = c_l(T) \implies \frac{g^*_l(T)}{g^*_l(T')} = \frac{\mathcal{K}^*_{T'}}{\mathcal{K}^*_T} > 1 \implies g^*_l(T') < g^*_l(T). \tag{4.8}$$

回想一下 $\mathbf{1}^T g^* = 1$，因此，

$$g^*_k(T') = 1 - \sum_{l \ne k} g^*_l(T') > 1 - \sum_{l \ne k} g^*_l(T) = g^*_k(T). \tag{4.9}$$

我们需要证明 $k \in E_>(T')$，也就是说

$$g^*_k(T') > \|g^*(T')\|_2^2$$

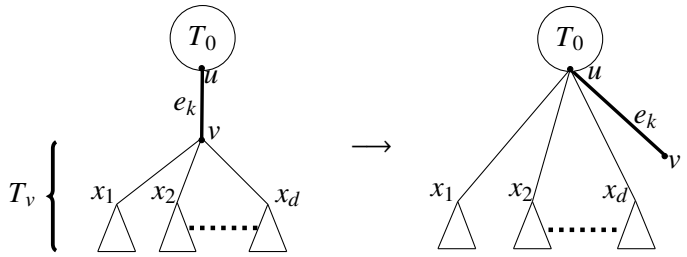


图2：下层 LT 的描述

利用（4.9），我们可以写成：

$$
\begin{aligned}
g_k^*(T') &= g_k^*(T) + (g_k^*(T') - g_k^*(T)) \\
&> \sum_{l \neq k} g_l^*(T)^2 + g_k^*(T)^2 + (g_k^*(T') - g_k^*(T)) && \text{since } k \in E_>(T) \\
&> \sum_{l \neq k} g_l^*(T')^2 + g_k^*(T)^2 + (g_k^*(T') - g_k^*(T)) && \text{using (4.8)} \\
&= \sum_{l \neq k} g_l^*(T')^2 + g_k^*(T')^2 + \left(g_k^*(T') - g_k^*(T) + g_k^*(T)^2 - g_k^*(T')^2\right) \\
&> \sum_{l} g_l^*(T')^2 = \|g^*(T')\|_2^2
\end{aligned}
$$

最后一个不等式成立，是因为 $\forall l$、$g_l^* < 1/2$（见上文 (4.6)），并且对于任意两个数 a、b，若 $a - b > 0$，则 $a + b < 1$ 成立：

$$a - b > (a-b)(a+b) = a^2 - b^2 \implies (a-b) - (a^2 - b^2) > 0.$$

将此事实应用于 $a = g_k^*(T')$、$b = g_k^*(T)$，即可完成证明。 □

到目前为止，我们已经证明：若存在如 (2.2) 中定义的 LT，则集合 $E_<$、$E_>$ 在匹配变换下保持不变，并且变换后的树的近似比大于原始树的近似比。利用这些性质，我们可以反复应用 LT（直至饱和），并保证最终的树是最难实例。因此，我们可以描述本节的主要过程了。我们从一棵阶为 n 的任意树 T开始。我们将其边划分为两个集合 $E_<(T)$、$E_>(T)$，如上文所定义。我们知道，这两个集合在下层和上层 LT 下保持不变——$E^\uparrow$、$E^\downarrow$ ——因此可以分别对两个集合反复应用这些变换，直到我们$\tilde{\ }$ $\tilde{\ }$得到 $E \circ T = T$。此外，根据命题 2.1，对于最终的树 T， $\alpha(T)$ 是原始树近似比的一个上界。剩下的工作是明确地定义 $E^\uparrow$、$E^\downarrow$，并计算 $\alpha(\tilde{T})$。

我们首先定义 $E^\downarrow$。假设这棵树以某个节点为根，因此“向上”和“向下”的方向都有明确的定义。令 $k \in E_<(T)$ 为一条非叶边（即其拥塞大于 m）。令 $e_k = (u, v)$。记 T_v 为以 v 为根的向下子树。于是，$E^\downarrow k$ 的作用是将 T_v “向上提升”至 u（见图2）。显然，对于任意 $l \neq k$，$c_l(E^\downarrow k \circ T) = c_l(T)$，因此 $E^\downarrow k$ 是一种局部变换。此外，$c_k(E^\downarrow k \circ T) = m < c_k(T)$（4.2），所以 $E^\downarrow k$ 确实是一种下层 LT。换言之，对所有 $k \in E_<(T)$ 应用 $E^\downarrow k$，只会将它们全部变成 T'中的叶节点。因此，我们首先对所有边 $k \in E_<(T)$ 应用 $E^\downarrow k$，得到 T'，使得对于任意 $l \in E_<(T')$，e_l都是一条叶边。因此，从现在起，我们假设 $E_<(T)$ 中的所有边都是叶边。

接下来，我们定义$E^\uparrow$。我们可以聚焦于由$E_>(T)$张成的子树$T_>$——所有其他边都是叶节点，因此两棵树的结构相同。令$k \in E_>(T)$。记$e_k = (u, v)$。如果$d_{T_>}(u)$、$d_{T_>}(v) \leq 2$，则

什么也不做。我们强调，$d(u)$是u 在子树**sub-tree** $T_>$ 中的度，也就是说，它是u 在$E_>$中的边数。现在我们将变换分为两种情况：

情形1：$d_{T_>}(u)$，$d_{T_>}(v) > 2$（见图5）。令 $n_v = |T_v|$。不失一般性，假设 $n_v \le n/2$。记 $l_i = (v, x_i)$，其中 $i = 1, ..., d(v)$，使得 $|T_{x_1}| = n_1 \le n_2 \le ... \le n_{x_{d(v)}}$。那么，$E^{\uparrow} k$ 是将 (v, x_2) 变为 (x_1, x_2)。显然，拥塞只在 $l'_1 = (x_1, x_2)$ 上发生变化。但由于 $n_1, n_2 \le n_1 + n_2 \le n_v \le n/2$，因此

$$\min(n_1, n - n_1) = n_1 < n_1 + n_2 = \min(n_1 + n_2, n - (n_1 + n_2))$$
$$\implies \underbrace{b_1(T) = n_1(n - n_1)}_{\text{congestion on } l_1} < \underbrace{(n_1 + n_2)(n - (n_1 + n_2)) = b_1(E_k^{\uparrow} \circ T)}_{\text{congestion on } l'_1}.$$

因此，确实 $E^{\uparrow} k$ 是一个上部 LT[4]。

情形2： 不失一般性，$d(u) = 2$，$d(v) > 2$（见图6）。与第一种情况类似，令 $l' = (u,y)$，$l_i = (v, x_i)$，其中 $i = 1, ...,d(v)$，使得 $|T_{x_1}| = n_1 \le n_2 \le ... \le n_{x_{d(v)}}$，$n' = n - |T_v|$。

情形2.1： 若 $n_1 \le n'$，那么同样需要将 $E^{\uparrow} k$ 中的(v,x_2)替换为(x_1,x_2)。拥塞仅在$l'_1 = (x_1, x_2)$上发生变化，但 $n_1 < n/(d(v) - 1) \le n/2$ 且 $n_1 < n_i, n'$，因此

$$\min(n_1, n - n_1) = n_1 < \min(n_1 + n_2, 1 + n' + n_3 + \dots)$$
$$\implies b_1(T) = n_1(n - n_1) < (n_1 + n_2)(n - (n_1 + n_2)) = b_1(E_k^{\uparrow} \circ T),$$

并且，事实上，$E^{\uparrow} k$ 是一个上部 LT。

情形2.2： 若 $n' < n_1$, ，那么$E^{\uparrow} k$ 需要取包含T_{x_1} 的(v,x_1)，并将其移至(u,x_1)（见图7）。拥塞仅在$e_k = (u,v)$上发生变化。但 $n' < n/(d(v) - 1) \le n/2$ ，$n' < n_1 \le n_i$ ，所以，

$$\min(n', n - n') = n' < \min(n_1 + n', 1 + n_2 + \dots)$$
$$\implies c_k(T) = n'(n - n') < (n' + n_1)(n - (n' + n_1)) = c_k(E_k^{\uparrow} \circ T),$$

因此，我们再次得到$E^{\uparrow} k$ 是一个上部 LT。

注4.5。请注意，在情形 **(2.2)** 结束时，我们要么转到情形**(1)**（如果 $d(v) > 3$），要么转到情形 **(2.1)**（如果$d(v) = 3$），因此我们只需关注第一类 upperLT（前两种情形）。对于这一类，经过变换后v 的度严格小于变换前的度，因此这一过程最终会终止，并且我们可以保证最终对所有满足$v \in E_>(T)$ 的 $d(v) = 2$，得到路径图 （由 $E_>$ 中的边构成）。

~现在我们终于可以描述最终的树 T。如前所述，我们应用下层 LT，直到$E_<$ 中的所有边都是叶节点。此外，我们应用上部 LT，直到$E_>$中的边构成一条单一路径。因此，最终的树具有“主干”结构（即路径），大小为 $p = |E_>(T)|$，沿途有 $s = |E_<(T)|$ 个“分支”（见图 3）。此时应该很清楚，将分支“推到一旁”、移至最近的一侧会增大近似比，因为这会增加树中的瓶颈。严格来说，将叶节点推到末端会增加路径上的拥塞，而这又是一次上部 LT，因此也会增大近似比~（记住，路径上的边来自$E_>$）。所以，最终的树 T 实际上就是

$$\widetilde{T} = \mathcal{B}_{s_1,p,s_2}$$

[4]虽然$E^{\uparrow} k$ 确实由k 定义，但它实际上会改变l_1上的拥塞。尽管这有些含糊，因为$l_1 \in E_>$，但这不影响证明，因此我们滥用了这一记号。

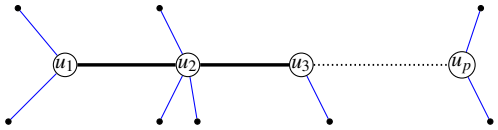


图 3：带有分支的茎图。“黑点”是叶节点，黑色边来自 $E_>$，蓝色边来自 $E_<$。

其中 $s_1 + s_2 = s$，而 $\mathcal{B}_{t,\ p,\ s}$ 是蝴蝶结图（见图 1）。

剩下的就是计算 $\alpha(\mathcal{B}_{t,\ p,\ s})$。我们知道 $T \subset T' \implies \alpha(T) \le \alpha(T')$（见断言 4.3），因此只需证明 $\mathcal{B}_{n,\ n,\ n}$（n 是原始树的大小）。虽然我们可以直接计算，但将采用一种简单得多的方法：利用后文为一般图推导的上界，得到 ~ 3 的上界，从而完成证明（详情见引理 4.10）。

4.2 一般图的 ERMP

现在我们继续讨论一般图的情形。与树不同，一般图的 ERMP 没有明确的“组合”形式，而是一个代数表达式。因此，我们的证明方法需要采取不同的思路。我们推导出近似比的两个不同上界，并猜想它们的最小值 $\widetilde{\ }$ $\widetilde{\ }$为 $O(1)$。事实上，我们将证明，对于蝴蝶结图，其最小值 为$O(1)$，这与定理 1.3 一致。最后，我们提供了大量支持这一猜想的模拟。

回顾一下，ERMP 可以表述为：

$$\begin{array}{ll} \text{minimize} & \mathbf{Tr}\, X^{-1} \\ \text{subject to} & X = \sum_{l=1}^{m} g_l b_l b_l^T + (1/n)\mathbf{1}\mathbf{1}^T \\ & \mathbf{1}^T g = 1 \end{array}$$

由于精确近似比的闭式表达式似乎难以计算，我们的目标是推导 ℓ_∞-LW 的近似比上界，下文将其记为 $\alpha_{A,\ D}$。结果表明，我们可以为 $\alpha_{A,\ D}$ 推导出两个不同的上界（它们之间存在有趣的关系）。第一个上界利用算术平均值 – 几何平均值不等式得到，第二个上界则通过拉格朗日对偶分析得到。

4.2.1 算术-几何平均上界

回顾一下，我们可以用“平均值”来表述 A-最优性和 D-最优性（见 §8）。因此，尝试使用算术平均值 – 几何平均值不等式是很合理的。我们知道，可以将 $\mathcal{K}_G(g)$ 写作（公式（3.7））

$$\mathcal{K}(g) = n\,\mathbf{Tr}\, L_g^+ = n\sum_{i=1}^{n-1}\lambda_i(L_g^+) = n\sum_{i=1}^{n-1}\lambda_i(L_g)^{-1} = \frac{n(n-1)}{HM(\lambda(L_g))} \tag{4.10}$$

（我们只有$n-1$ 个正特征值）。因此，我们得到

$$HM(\lambda(L_{g^*})) = \frac{n(n-1)}{\mathcal{K}_G^*},\ HM(\lambda(L_{g_{\ell w}})) = \frac{n(n-1)}{\mathcal{K}_G^{\ell w}}$$

利用算术平均值-几何平均值 不等式和方程（3.3），可得：

$$\alpha_{A,D} = \frac{\mathcal{K}_G^{\ell w}}{\mathcal{K}_G^*} = \frac{HM(\lambda(L_{g^*}))}{HM(\lambda(L_{g_{\ell w}}))} \le \frac{AM(\lambda(L_{g^*}))}{HM(\lambda(L_{g_{\ell w}}))} = \frac{2/(n-1)}{HM(\lambda(L_{g_{\ell w}}))} = \frac{2}{n(n-1)^2}\mathcal{K}_G^{\ell w}$$

我们将这个界称为算术平均值-几何平均值 界，记作 α_{AM}:

$$\alpha_{A,D} \le \alpha_{AM}(\ell w) = \frac{2}{n(n-1)^2}\mathcal{K}_G^{\ell w} = \frac{2}{(n-1)^2}\,\mathbf{Tr}\,L_{\ell w}^{+}$$

值得注意的是，由于 $\mathcal{K}_G^* \ge n(n-1)^2/2$（参见 [GBS08] 中关于最优解的下界），我们得到:

$$\alpha_1(g) = \frac{2}{n(n-1)^2}\mathcal{K}_G(g) \ge \frac{2}{n(n-1)^2}\mathcal{K}_G^* \ge 1,$$

当且仅当$g = g^*$时取等号，因此这是定义良好的。

4.2.2 对偶间隙上界

接下来，我们希望利用 ERMP 的对偶间隙 推导出第二个上界。利用对偶问题来界定可行解的次优性是一种众所周知的技巧，因此我们的动机很明确。我们首先重述 [GBS08] 关于 ERMP 问题对偶间隙的结果。由于这是我们结果的关键组成部分之一，为求完整，我们从头推导对偶间隙；不过请注意， [GBS08] 已经完成了这项工作。我们在此陈述最终结果，完整证明见附录 A.5。

ERMP 的对偶问题为（见公式 (A.15)）:

$$\begin{aligned} &\text{maximize } n(\mathbf{Tr}\,V)^2 \\ &\text{subject to } b_l^T V^2 b_l = \|Vb_l\| \le 1, \\ &\qquad V \succeq 0\ ,\ V\mathbf{1} = 0 \end{aligned}$$

我们知道， $\mathcal{K}_G(g)$ 关于 g 的梯度等于（见公式 (A.5)):

$$\frac{\partial L_g^+}{\partial g_l} = -\,\mathbf{Tr}\,L_g^+ b_l b_l^T L_g^+ = -\|L_g^+ b_l\|^2$$

对于任意可行的 g，令 $V = \frac{1}{\max_l \|L_g^+ b_l\|} L_g^+$（这显然是对偶可行的）。根据定义，$g$ 中的对偶间隙等于，

$$\begin{aligned} \eta = n\,\mathbf{Tr}\,L_g^+ - n(\mathbf{Tr}\,V)^2 &= n\,\mathbf{Tr}\,L_g^+ - \frac{n(\mathbf{Tr}\,L_g^+)^2}{\max_l \|L_g^+ b_l\|^2} \\ &= \frac{\mathcal{K}_G(g)}{-\min_l\ \partial\mathcal{K}_G(g)/\partial g_l}\left(-\min_l \frac{\partial\mathcal{K}_G(g)}{\partial g_l} - \mathcal{K}_G(g)\right) \end{aligned}$$

利用这一点，我们可以推导出第二个上界。根据对偶间隙的定义，

$$\mathcal{K}_G(g) - \mathcal{K}_G^* \le \eta = \frac{\mathcal{K}_G(g)}{-\min_l\ \partial\mathcal{K}_G(g)/\partial g_l}\left(-\min_l \frac{\partial\mathcal{K}_G(g)}{\partial g_l} - \mathcal{K}_G(g)\right) = \mathcal{K}_G(g) + \frac{(\mathcal{K}_G(g))^2}{\min_l\ \partial\mathcal{K}_G(g)/\partial g_l}$$

请记住， $\mathcal{K}_G(g)$ 的导数为负，经过简单整理，我们得到 g et,

$$\frac{\mathcal{K}_G(g)}{\mathcal{K}_G^*} \le \frac{-\min_l \frac{\partial\mathcal{K}_G(g)}{\partial g_l}}{\mathcal{K}_G(g)} = \frac{-\min_l(-n\|L^+ b_l\|_2^2)}{n\,\mathbf{Tr}\,L^+} = \frac{\max_l \|L^+ b_l\|_2^2}{\mathbf{Tr}\,L^+}$$

代入$g_{\ell w}$后，我们将这个界（即对偶间隙界）记为 α_{dual},

$$\alpha_{A,D} \le \alpha_{dual}(\ell w) := \frac{\max_l \|L_{\ell w}^+ b_l\|_2^2}{\mathbf{Tr}\,L_{\ell w}^+}$$

4.2.3 ERMP 的近似比上界

现在，我们可以定义近似比的推导上界 -$\alpha_{A,\ D}$:

$$\alpha_{A,D} \leq \alpha_{min} \coloneqq \min(\alpha_1, \alpha_2),$$

在全文中， α_{AM}, α_1 始终表示 $\alpha_{AM}(\ell w)$，而 α_{dual}, α_2 始终表示 $\alpha_{dual}(\ell w)$；除非另有说明，单独的 α 则表示 α_{min}。

现在我们可以提出以下更强的猜想:

猜想 4.6. 对于任意图 G， $\alpha_{min} \leq O(\log n)$

我们强调，这与猜想 1.5 并不等价，因为我们不知道 α_{min} 是否是紧的。不过，我们的模拟表明，α_{min} 已经足够（见下文）。

本节余下部分将推导关于 α_1 和 α_2 及其相互关系的一些公式。我们的第一个引理给出了这两个界之间的联系:

引理 4.7. $\alpha_2(g) = \left\|-\nabla_g(\log \alpha_1)\right\|_\infty$

证明。 回想一下， $\alpha_1(g) = \frac{2}{(n-1)^2}\,\mathbf{Tr}\, L_g^+$。利用链式法则和公式 (A.5)，我们得到

$$\begin{aligned}
-\frac{\partial}{\partial g_l}\log(\alpha_1(g)) &= -\frac{1}{\alpha_1(g)}\frac{\partial \alpha_1}{\partial g_l} = \frac{-(n-1)^2}{2\,\mathbf{Tr}\, L_g^+}\frac{2}{(n-1)^2}\,\mathbf{Tr}\,\frac{\partial L_g^+}{\partial g_l} \\
&= \frac{-1}{\mathbf{Tr}\, L_g^+}\,\mathbf{Tr}(-L_g^+ b_l b_l^T L_g^+) = \frac{\mathbf{Tr}\, b_l^T L_g^+ L_g^+ b_l}{\mathbf{Tr}\, L_g^+} = \frac{\|L_g^+ b_l\|^2}{\mathbf{Tr}\, L_g^+}
\end{aligned}$$

因此,

$$\left\|-\nabla_g(\log \alpha_1)\right\|_\infty = \max_l \frac{\|L_g^+ b_l\|^2}{\mathbf{Tr}\, L_g^+} = \alpha_2(g).$$

□

我们还可以对 α_1给出一种更具“组合性”的解释:

引理 4.8。 $\alpha_1 \leq D$，其中 D 是图的直径。

证明。 我们知道，在$g_{\ell w}$处，任意一条边上的有效电阻为$n-1$。令 $\mathrm{SP}(i,j)$ 表示 i 到 j 之间的最短路径。利用三角不等式（见公式 (3.6)），我们可以对$R_{ij}^{\ell w} = R_{ij}(g_{\ell w})$进行界定:

$$R_{ij}^{\ell w} \leq \sum_{l \in \mathrm{SP}(i,j)} R_l^{\ell w} = \sum_{l \in \mathrm{SP}(i,j)} (n-1) = (n-1)|\mathrm{SP}(i,j)| \leq (n-1)D,$$

其中D 是图的直径。因此，我们得到:

$$\mathcal{K}_G^{\ell w} = \sum_{i<j} R_{ij}^{\ell w} \leq \sum_{i<j}(n-1)D = \binom{n}{2}(n-1)D = \frac{n(n-1)^2}{2}D,$$

and,

$$\alpha_1 = \frac{2}{n(n-1)^2}\mathcal{K}_G^{\ell w} \leq \frac{2}{n(n-1)^2}\cdot\frac{n(n-1)^2}{2}D = D$$

□

相比之下， α_{dual} 与拉普拉斯矩阵的条件数有着密切联系：

引理 4.9. $\alpha_{dual} \le \kappa(L_{lw}) = \frac{\lambda_n(L_{lw})}{\lambda_2(L_{lw})}$

证明。 根据 Courant-Fisher 原理，

$$\lambda_n(L_{lw}^+) = \max_{x\in\mathbb{R}^n}\left\{\frac{x^T L_{lw}^+ x}{x^T x}\right\}$$

特别地，取 $x = (L_{lw}^+)^{1/2} b_l$ (请记住，L_g^+ 是对称半正定矩阵)可得，

$$\frac{b_l^T L_{lw}^+ L_{lw}^+ b_l}{b_l^T L_{lw}^+ b_l} \le \lambda_n(L_{lw}^+)$$

i.e,

$$\|L_{lw}^+ b_l\|^2 = b_l^T L_{lw}^+ L_{lw}^+ b_l \le \lambda_n(L_{lw}^+)\cdot b_l^T L_{lw}^+ b_l = (n-1)\lambda_n(L_{lw}^+).$$

此外，请记住，$\min(x_1,...,x_k) \le$算术平均值$(x_1,...x_k)$。因此，将其用于L_{lw}^+的特征值，并利用迹等于特征值之和这一事实，可得，

$$\lambda_2(L_{lw}^+) \le \frac{1}{n-1}\,\mathbf{Tr}\,L_{lw}^+$$

因此，

$$\alpha_2 = \max_l \frac{\|L_{lw}^+ b_l\|^2}{\mathbf{Tr}\,L_{lw}^+} \le \max_l \frac{(n-1)\lambda_n(L_{lw}^+)}{\mathbf{Tr}\,L_{lw}^+} \le \frac{(n-1)\lambda_n(L_{lw}^+)}{(n-1)\lambda_2(L_{lw}^+)} = \frac{\lambda_n(L_{lw}^+)}{\lambda_2(L_{lw}^+)} = \frac{\lambda_n(L_{lw})}{\lambda_2(L_{lw})} = \kappa(L_{lw})$$

这里用到了L^+ 的特征值是L的特征值的倒数这一事实 . □

如前所述，我们用 α_{dual} 推导树的近似比：

引理 4.10. 对于蝴蝶结图，$\mathcal{B}_{n,n,n}$, $\alpha_{dual} \approx 3.12$

证明. 回顾一下，对于树，

$$\mathcal{K}_T(g) = \sum_l \frac{c_l}{g_l},$$

and,

$$\frac{\partial\mathcal{K}(g)}{\partial g_l} = -\frac{c_l}{g_l^2}$$

因此，对于树，我们可以将 α_{dual} 写作

$$\alpha_{dual} = \max_l\left(-\frac{\partial\mathcal{K}(g)}{g_l}/\mathcal{K}(g)\Big|_{g_{\ell w}}\right) = \max_l \frac{m^2 c_l}{m\sum_k c_k} = \frac{m\cdot c_{\max}}{\sum_k c_k} \qquad g_{\ell w} = g_{uni}\text{, see claim 4.1}$$

显然， $c_{max} \le (3n/2)^2$ （中间的边）。至于分母，简单计算可得 u s,

$$\sum_k c_k = \underbrace{2n\cdot m}_{2n\text{ leafs edges}} + \underbrace{\sum_{k=n+1}^{2n-1} k(3n-k)}_{\text{path's edges}}$$

$$= 2n\cdot m + \sum_{k=1}^{2n-1} k(3n-k) - \sum_{k=1}^{n} k(3n-k) \approx 2n\cdot m + \frac{10}{3}n^3 - \frac{7}{6}n^3 = \frac{13}{6}n^3 + 6n^2.$$

因此，我们得到

$$\alpha_{dual}(\mathcal{B}_{n,n,n}) \leq \underbrace{\frac{m(3n/2)^2}{(13/6)n^3} \leq \frac{27/4}{13/6}}_{m=3n-1} \approx 3.12$$

□

需要注意的是，这是一个上界，而实验结果显示比值略好（见下文）。

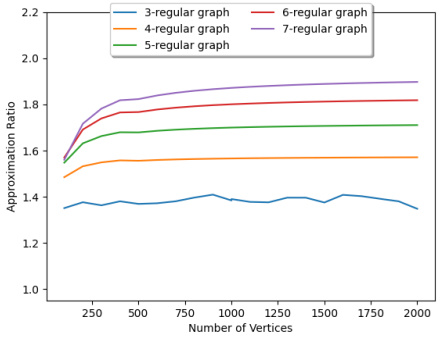
虽然我们尚未能证明一般图上的猜想，但可以利用上述公式证明该猜想适用于某些图族。例如，引理 4.8 证明了低直径图上的猜想。此外，我们知道优良的扩展图具有较低的条件数 [Spi11], , 因此引理 4.9 也证明了该猜想适用于这类图。对于更多图，我们通过各种模拟为猜想提供了有力证据，详情如下。

5 实验结果

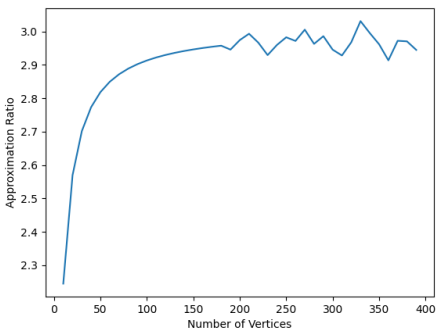
设置。 模拟使用 α_{min}计算近似比，从而与猜想 4.6 相符。所有随机图的结果均取 100 次运行中的最大值。我们使用算法 1 计算图的 LW，参数为 $\epsilon = 0.01$。若图不连通或不是简单图，则移除所有自环并取其最大连通分量（这只是技术处理）。我们对若干基本图族进行猜想模拟，例如k-正则图、棒棒糖图和网格图。表 1 总结了我们检验过的（有趣的）图及其近似比。图 4 展示了正则图和棒棒糖图的渐近行为。

数据集 type	图	节点	边	$\alpha_{min}(G)$
随机、高直径、d-正则	3-正则	400	600	1.55
	4-正则	400	800	1.17
	5-正则	400	1 000	1.11
	6-正则	400	1 200	1.08
小世界图	Watts–Strogatz小世界图[5]	400	800	1.64
grid	平衡网格[6]	400	760	1.35
	10倍网格[7]	400	750	1.94
扩展图	Margulis-Gabber-Galil图	400	1 520	1.06
	弦环图	400	1 196	1.64
稠密图	(400,400)–棒棒糖图	800	80 200	3.03
树	蝴蝶结图	3 000	2 999	2.5
真实世界图[8]	酵母	2 224	6 609	2.23
	Stif	17 720	31 799	4.08
	皇家	2 939	4 755	8.96

表 1：初等图近似界的总结。



(a) α_{min} 高直径正则图



(b) α_{min} 顶点数不超过 800 的棒棒糖图

图 4：初等图的渐近行为

[5]参数为 $-k = 4, p = 2/3$[6]2维正方形[7]矩形，宽度大小为 10[8]的图取自 [MSJ19]

6 ℓ_∞-LW 重加权的谱含义（定理 1.6 的证明）

本节研究 ℓ_∞-LW 如何影响重加权图拉普拉斯矩阵的特征值分布。我们将介绍若干相关结果，这些结果在数学和算法方面都具有意义。

6.1 代数连通度的界

[Moh+91] 的经典结果指出，对于无权图，代数连通度有如下界：

$$\lambda_2 \geq 4/nD.$$

我们将这一界推广至加权图，具体如下：

引理 6.1. 给定一个图 $G(g)$，记 $R_{max}(g) = \max_l R_l(g)$ 为最大边 ER，R_{max} 中的最大两两 ER 记为 $\widetilde{pairwiseER}$。则，

$$\lambda_2(L_g) \geq \frac{2}{n\widetilde{R_{max}}} \geq \frac{2}{nD \cdot R_{max}(g)} \tag{6.1}$$

and

$$\lambda_2(L_g) \geq \frac{4}{n\sum_l R_l(g)} \tag{6.2}$$

证明。 我们知道有效电阻是一种度量，因此对于任意两个顶点i、j，

$$R_{ij}(g) \leq \sum_{l \in P_{ij}} R_l(g) \leq D \cdot R_{max}(g)$$

其中 $\widetilde{P_{ij}}$ 是i 与j之间最短（长度）的路径。特别地，$R_{max}(g) \leq D \cdot R_{max}$。因此 s

$$\mathcal{K}_G(g) = \sum_{i<j} R_{ij}(g) \leq \binom{n}{2} \cdot \widetilde{R_{max}} = \frac{n(n-1)}{2}\widetilde{R_{max}}(g)$$

利用 $\mathcal{K}_G(g) = n\,\mathbf{Tr}\,L_g^+$这一事实，可得

$$\mathbf{Tr}\,L_g^+ \leq \frac{(n-1)}{2}\widetilde{R_{max}}(g)$$

我们知道 $\lambda_{max}(L_g^+) \leq \mathbf{Tr}\,L_g^+$，因此

$$\lambda_{max}(L_g^+) \leq \frac{(n-1)}{2}\widetilde{R_{max}}(g) \implies \lambda_2 \geq \frac{2}{(n-1)\widetilde{R_{max}}(g)} \geq \frac{2}{n\widetilde{R_{max}}(g)} \geq \frac{2}{nD \cdot R_{max}(g)}$$

对于第二个界，定义P_{ij} 的特征函数如下

$$\chi_{ij}(e) = \begin{cases} 1 & \text{if } e \in P_{ij} \\ 0 & \text{if } otherwise \end{cases}$$

现在，我们可以写成

$$\begin{aligned} \mathcal{K}(g) &= \frac{1}{2}\sum_{i,j} R_{ij}(g) \\ &\leq \frac{1}{2}\sum_i \sum_j \sum_{l \in P_{ij}} R_l = \frac{1}{2}\sum_i \sum_j \sum_{l \in E} R_l \cdot \chi_{ij}(e) \\ &= \frac{1}{2}\sum_l R_l \sum_i \sum_j \chi_{ij}(e) \leq \frac{n^2}{4}\sum_l R_l \end{aligned}$$

其中，在最后一行中，我们用到了以下事实：任意一条边 e 至多属于 $n^2/4$ 条路径 P_{ij} [Moh+91]。沿用之前相同的论证，我们得到

$$\lambda_{max}(L_g^+) \leq \frac{n}{4}\sum_l R_l \implies \lambda_2 \geq \frac{4}{n\sum_l R_l}$$

□

用于检查合理性，我们应当看看无权图的情形。事实上，对于$g = \mathbf{1}$，我们有

$$\begin{aligned} L_g = BB^T &\implies L_g^+ = (BB^T)^+ \\ &\implies BL_g^+B^T = B(BB^T)^+B^T = BB^+ \\ &\implies R_l = \mathrm{diag}(BL_g^+B^T)_l = \mathrm{diag}(BB^+)_l \leq \mathbf{1} \end{aligned}$$

最后一行之所以成立，是因为BB^+ 是投影矩阵。因此，我们得到$R_{max} \leq 1$，于是（使用第一个界）

$$\lambda_2 \geq \frac{2}{nD \cdot R_{max}} \geq \frac{2}{nD}$$

这与前一个界几乎相同（并且优于较弱的 $1/nD$ 版本）.

注 6.2。 这里我们只讨论了代数连通度，但还有许多其他图性质和界仅适用于无权图。上述技术以及可能的其他方法，可用于将这些界推广到加权图，并用有效电阻来表示。

6.2 最小化图的混合时间

回顾一下，图的谱隙是拉普拉斯矩阵的最小非负特征值。由于我们假设图是连通的，因此它等于 $\lambda_2(L_g) = \lambda_2$。较短（更快）的混合时间等价于更大的谱隙，因此我们的目标是最大化 λ_2。众所周知，可以使用对数和指数 （LSE）函数 [NES03] 对n个变量的最大值进行“平滑”：

$$\mathrm{LSE}(x_1, ..., x_n) = \log\left(\sum_i e^{x_i}\right)$$

LSE 是可微凸函数，对于任意n个数，LSE 满足

$$x_{max} \leq \mathrm{LSE}(x_1, ..., x_n) \leq x_{max} + \log(n)$$

因此，为了分析 $\lambda_2 = 1/\lambda_n(L^+)$，考虑$L^+$ 个特征值的 LSE 是很自然的（最大化λ_2 = 最小化 $\lambda_n(L^+)$）。利用以下事实：

$$\mathbf{Tr}\exp(M) = \sum_i e^{\lambda_i(M)}$$

我们将 softmaxEV（SEV）函数定义为：

$$\mathrm{SEV}(g) = \log(\mathbf{Tr}[\exp(L_g^+)])$$

为了模拟 简化计算，利用log 是单调函数这一事实，我们将分析 ：

$$f(g) = \mathbf{Tr}[\exp(L_g^+)]$$

注意，由于指数运算保留凸性，$f(g)$ 仍然是凸的。在附录 A.7 中，我们证明了 SEV 的最优性准则为：

$$\mathbf{Tr}[e^{L_g^+} L_g^+ (I - b_l b_l^T L_g^+)] \geq 0. \tag{6.3}$$

在下一节中，我们将证明，当 (6.3) 是逐点非负的 （即所有加项都是半正定的）时，判据 (6.3) 与 ℓ_∞-LW 最优性条件紧密相关，因此给出了 LW 重加权能够改善图G混合时间的条件。

最优性的一个充分条件 注意，如果$\mathbf{I} - b_l b_l^T L_g^+$是半正定矩阵，那么方程 (6.3) 很容易满足。此外，该矩阵的谱相当优美——$b_l b_l^T L_g^+$是一个秩一矩阵，只有一个非零特征值，等于$R_l(g)$（对应的特征向量为b_l）。因此$\mathbf{I} - b_l b_l^T L_g^+$的谱为

$$\underbrace{1, ..., 1}_{n-1}, 1 - R_l(g)$$

因此，g 达到最优的一个充分条件是每条边上的有效电阻都小于 1。注意，这个条件与归一化 LW 条件相似，在该条件下，每条边上的有效电阻至多为$n - 1$。

遗憾的是，不存在满足该条件的可行解。要了解原因，请注意，对于任意可行的g（方程 (A.9)）

$$-\nabla \operatorname{logdet}(g)^T \cdot g = \sum_l R_l(g) \cdot g_l = n - 1$$

但是，如果对于某个g'，我们有$R_l(g') \leq 1$，那么，

$$n - 1 = \sum_l R_l(g) g_l \leq \sum_l g_l = 1$$

矛盾[9]。总之，不存在使$I - b_l b_l^T L_g^+$为半正定的可行解。不过，这并不意味着不存在最优解——即使矩阵不是半正定，迹也可能为正。遗憾的是，除了上述条件之外，我们不知道如何推导出最优解的闭式表达式。

相对最优性 虽然 SEV 函数的最优解尚不明确，但我们可以利用凸性，尝试证明“解 1 优于解 2”。形式上，对于凸函数f

$$f(x) \leq f(y) \iff \nabla f(x)^T (y - x) \geq 0$$

例如，取$x = g_{lw} = \widehat{g}$，$y = g_{uni} = (1/m)\mathbf{1}$，我们得到以下条件

$$\nabla f(\widehat{g})^T ((1/m)\mathbf{1} - \widehat{g}) = (1/m) \sum_l \frac{\partial f}{\partial g_l} - \nabla f(\widehat{g})^T \widehat{g} \geq 0$$

利用附录 A.7 中推导出的公式，可得到如下形式

$$-(1/m) \sum_l \mathbf{Tr}[e^{L_{lw}^+} L_{lw}^+ b_l b_l^T L_{lw}^+] + \mathbf{Tr}[e^{L_{lw}^+} L_{lw}^+] = \mathbf{Tr}[e^{L_{lw}^+} L_{lw}^+ (\mathbf{I} - \sum_l (1/m) b_l b_l^T L_{lw}^+)] = \mathbf{Tr}[e^{L_{lw}^+} L_{lw}^+ (\mathbf{I} - L_{uni} L_{lw}^+)] \geq 0$$

[9] $n \leq 2$ 是可忽略的情形

注记 6.3. 最后一个条件并非显而易见。要理解原因，请注意，若 $\mathbf{I} - L_{uni}L_{lw}^{+}$ 是半正定的，则最后一个条件成立。不难看出，

$$\mathbf{I} - L_{uni}L_{lw}^{+} \succeq 0 \iff L_{uni}L_{lw}^{+} \preceq \mathbf{I} \iff (1/m)B^{T}L_{lw}^{+}B \preceq \mathbf{I} \iff B^{T}L_{lw}^{+}B \preceq m\mathbf{I}.$$

两边分别乘以$W = 1/(n-1)\cdot \mathrm{diag}(w_{\infty})$，得到

$$W^{1/2}B^{T}(BWB^{T})^{+}BW^{1/2} = \Pi_{W^{1/2}B^{T}} \preceq mW$$

众所周知，投影矩阵的特征值不是 0 就是 1。另一方面，i 个mW 的特征值显然为$m(w_{\infty})_i/(n-1)$。换言之，如果$w_{\infty} \geq (n-1)/m$，那么我们可以得出结论：重加权图的混合时间比无权图更短。遗憾的是，这种情况不可能发生——请记住，LW 的总和为 $\mathrm{rank}(B) = n-1$，因此如果 $(w_{\infty})_{\min} > (n-1)/m$，我们得到

$$n-1 \sum_{i}(w_{\infty})_i \geq m\cdot(w_{\infty})_{\min} > m\cdot((n-1)/m) = n-1$$

矛盾，除非$g_{\ell w} = g_{uni}$。我们强调，这并不意味着 LW 无法改善混合时间，因为$\mathbf{I} - L_{uni}L_{lw}^{+} \succeq 0$ 只是一个充分条件，而模拟确实表明 LW 能改善混合时间。不过，我们没有关于这种改善的任何闭式条件。

6.3 最优谱稀疏树

LW 的最后一项应用是谱稀疏树。若生成树 T 属于G，且满足 $L_T \preceq \gamma L_G$，则称其为 γ-谱稀疏树。在 [Asa+10] 中，研究者证明了非对称旅行商问题与谱稀疏树之间存在深刻联系：找到一棵低谱稀疏度的树，就可以将其与 ASTP 的分数解结合，从而找到一个近似解。这一直是 [Asa+10] 及其后续研究的关键要素。在 [HO13] 中，作者证明，对于连通图 $G = \langle V,E\rangle$，任意生成树 T 属于 G 的谱稀疏度至少为 $O(\max_{e\in T} R_G(e))$（即T 在G中的最大边有效电阻）。此外，可以在多项式时间内找到达到该下界的树。现在我们知道（§3.4），LW 会使后者最小。因此，自然会问：如何利用 LW 找到图的最优谱稀疏树?

沿着 [HO13], 的工作，我们证明以下引理：

引理 6.4。 对于任意连通图 $G = \langle V,E\rangle$，存在一棵加权生成树 T_g，使得 T_g 的谱稀疏度为 $((n-1)/m)\cdot O(\log n/\log\log n)$。

证明与 [HO13] 非常相似，并将使用以下引理：

引理 6.5。 设 $w_1,\ldots,w_m \in \mathbb{R}^n$ 为具有单位范数的 n 维向量。设 $p_1,\ldots,p_m$ 为定义在这些向量上的概率分布，使得协方差矩阵为 $\sum i\, p_i w_i w_i^{T} = (1/n)\mathbf{I}$。那么，存在一个多项式算法，可计算出一个子集 $S \subseteq [m]$，使得 $\{w_i \mid i \in S\}$ 构成 $\mathbb{R}^n$的一个基，并且$\| \sum i\in S\, w_i w_i^{T}\| \leq \alpha$ 其中为 $\alpha = O(\log n/\log\log n)$。

命题 *6.4* 的证明。 令w_{∞} 为图的 LW。定义$g_0 = n-1\ m\ \mathbf{1}$、$w = n-1\ m\ w_{\infty}$。对于每个$e_l \in E$，定义$w_l = \sqrt{(g_0)_l}\cdot(L_w)^{\pm/2}b_l$、$\boldsymbol{p} = \frac{1}{n-1}w_{\infty}$。事实上，由 (3.1) 以及伪逆的乘法性质可得

$$\sum_{l} p_l = 1\ ,\ \|w_l\|^2 = (g_0)_l\cdot\frac{m}{n-1}b_l^{T}(BW_{\infty}B^{T})^{+}b_l = \frac{n-1}{m}\cdot\frac{m}{n-1} = 1$$

此外，

$$\begin{aligned}\sum_{e_l\in E} p_l w_l w_l^T &= \frac{1}{n-1}(L_{\overline{w}})^{+/2}\left(\sum_l (w_\infty)_l (g_0)_l b_l b_l^T\right)(L_{\overline{w}})^{+/2}\\ &= \frac{1}{n-1}(L_{\overline{w}})^{+/2}\left(\sum_l \overline{w}_l b_l b_l^T\right)(L_{\overline{w}})^{+/2}\\ &= \frac{1}{n-1}\mathbf{I}_{\mathrm{im}L_G}\end{aligned}$$

因此，我们可以将向量 $\{w_l \mid l \in E\}$ 视为（$n-1$）维向量（位于B的线性张成中），并应用引理 6.5，得到一个大小为$n-1$ 的集合$T \subset E$（T 在 $\mathbb{R}^{n-1}$ 中构成一个基），使得 $\{w_e \mid e \in T\}$ 线性无关，并且

$$\sum_{e_l\in T} w_l w_l^T \preceq O(\log n/\log\log n)\mathbf{I}_{\mathrm{im}L_G}$$

前两个条件意味着T 诱导出G的一棵生成树。然后，最后一个条件给出（简单整理后）

$$\sum_{e_l\in T}(g_0)_l b_l b_l^T = L_{T_g} \preceq O(\log n/\log\log n)L_{\overline{w}}$$

现在，我们知道 LW 至多为 1，因此，$\overline{w} = \frac{n-1}{m} w_\infty \leq \frac{n-1}{m}\mathbf{1}$，这意味着

$$L_{\overline{w}} \preceq \frac{n-1}{m}BB^T = \frac{n-1}{m}L_G$$

因此，我们得到

$$L_{T_g} \preceq ((n-1)/m)\cdot O(\log n/\log\log n)L_G$$

□

ATSP 角度 探讨加权 *STT* 能否以及如何用于 ATSP 舍入方案（仿照 [Asa+10]）将会很有意思。尽管如此，我们强调 LW 加权 STT 的两项优势：首先，与无权情形 [AG15], 不同，我们保证无论原始图是什么，都能找到一个$O((n-1)/m)$-谱稀疏树 无论 如何；这是最优的——要理解原因，请回想 LW 是最大边 ER 的最优极小化器。取$g_{\ell w}$ 和g_{uni}，二者均已归一化，我们有

$$R_{max}(g_{\ell w}) = n-1 \leq R_{max}(g_{uni}) = m\cdot R_{max}(\mathbf{1}) \implies R_{max}(\mathbf{1}) \geq \frac{n-1}{m}.$$

其次， [AG15] 保证的无权谱稀疏树至多只能达到这一界限，其总权重为 $n-1$。相比之下，我们的加权谱稀疏树的总权重为 $\frac{(n-1)^2}{m} \leq n-1$ （对于大多数图，$\ll$）。从这个意义上说，引理 6.4 中的树总是更“便宜”。

7 计算 Lewis 权重

本节概述近期通过重复计算杠杆分数来计算矩阵的 ℓ_∞-LW 的加速方法（在我们的情形中，即拉普拉斯线性方程组）。第一种方法由 [Coh+20] 提出（基于[CP14]），其运行时间为输入稀疏时间，实用性很强，但只能提供低精度解。第二种方法由 [Faz+21]提出，为 ℓ_p-LW 提供了一种高精度算法，需要进行 $\widetilde{O}(p^3\log(1/\epsilon))$ 次杠杆分数计算（在我们的情形中为拉普拉斯线性方程组）。幸运的是，正如下文所示，低精度对于有效电阻矩阵素描（ERMP）应用已经足够，因此我们重点讨论 [Coh+20]的算法。不过，我们指出，在大多数优化应用中，只要计算满足 $p = n^{o(1)}$ 的 ℓ_p-LW 即可（见 [Faz+21；最小二乘14]），此时 [Faz+21] 的算法可为拉普拉斯矩阵提供一种高精度的$O(m^{1+o(1)}\log(1/\epsilon))$ 时间算法。

7.1 以低精度 [Coh+20] 计算 ℓ_∞-刘易斯权重

算法 1计算 ℓ_∞-刘易斯权重

输入： 一个秩为 k 的矩阵 $A \in \mathbb{R}^{m\times n}$，$T$—迭代次数。

结果： ℓ_∞-刘易斯权重，$w \in \mathbb{R}^m$。

Initialize: $w_l^{(1)} = \frac{k}{m}$ for $l = 1,\dots,m$.

for $t = 1,\dots,T-1$ **do**

$w_l^{(t+1)} = w_l^{(t)} \cdot a_l^T (A^T \operatorname{diag}(w^{(t)})A)^+ a_l$; for $l = 1,\dots,m$ ▷ // We can use here a Laplacian LS solver.

$(w)_i = \frac{1}{T}\sum_{t=1}^{T} w_i^{(t)}$ for $i = 1,\dots,m$

return w

由于我们将 [Coh+20] 作为黑盒使用，因此这里只对算法 1作高层次概述。[Coh+20；CP14] 的基本思路是利用如下观察：对于（精确的）ℓ_∞-LW，我们有

$$w_i = \tau_i(W^{1/2}A). \tag{7.1}$$

[Coh+20] 使用（7.1）定义一个如算法 1 所述的不动点迭代。换言之，该算法反复计算加权矩阵的杠杆分数，并根据以往迭代的平均值更新权重，直到达到一个（近似）不动点。该算法的性能如下：令 $\epsilon > 0$，并用 $\widetilde{}w$ 表示算法对输入 $A \in \mathbb{R}^{m\times n}$ 的输出。我们的目标是计算近似刘易斯权重，$\widetilde{}w$，使得

$$\widetilde{w} \approx_\epsilon w_\infty \tag{7.2}$$

其中，$a \approx_\epsilon b$ 当且仅当$a = (1 \pm \epsilon)b$。然而，算法 1 的近似保证是从最优性意义上给出的，即

$$\tau_i(\widetilde{W}^{1/2}A) \approx_\epsilon \widetilde{w}_i, \quad \forall i \in [m]. \tag{7.3}$$

幸运的是，[Faz+21] 最近证明，ϵ-近似最优 LS (7.3) 可推出 ϵ-近似 LW (7.2)：

定理 7.1。*(LS 近似 $\Rightarrow$ LW 近似，[Faz+21])*。对于所有 $\epsilon \in (0,1)$，算法 *1* 在 $T = O(\epsilon^{-1}\log m\, n)$ 次迭代后输出 LW 的 $(1\mp\epsilon)$-近似。

近似 **刘易斯权重：** 由于算法 1 的输出满足 $\widetilde{}w \approx_\epsilon w_\infty$，这意味着 s

$$\widetilde{L_w} = \sum_l (\widetilde{g_w})_l b_l b_l^T \approx_\epsilon \sum_l (g_{\ell w})_l b_l b_l^T = L_{\ell w}$$

其中 $\widetilde{}g_w$,$g_{\ell w}$ 分别是由 $\widetilde{}w$、w_∞定义的权重向量。因此，

$$\mathbf{Tr}\,\widetilde{L_w} \approx_\epsilon \mathbf{Tr}\, L_{\ell w}$$

换句话说，计算近似 LW 可为迹 $L_{\ell w}$ 提供 a-$(1+\epsilon)$ 因子。由于我们的目标函数仅为迹，这意味着近似 LW 将为我们的近似比提供 a-$(1+\epsilon)$ 因子。由于我们对 ERMP 只要求 $O(1)$ 近似，后者已经足够，因此算法 1 可用于生成定理 1.3 和 1.4 中所述的近似 ERMP 权重。请注意，这也适用于其余结果，因为 (7.2) 中的邻近性概念也意味着谱近似。

8 一般半正定矩阵的 A 与 D 设计之间存在 Ω(n/log² n) 差距

最后，我们说明为何聚焦于拉普拉斯矩阵。我们证明，对于一般半正定矩阵，A-最优设计与 D-最优设计问题实际上截然不同，具体而言， ℓ_∞-LW 对 A-最优设计问题所能达到的近似程度不可能优于 $\Omega(n/\log^2 n)$。

这一差距的动机源于将 A-最优设计和 D-最优设计重新表述为调和平均数 (HM) 与几何平均数 (GM) 的最小化问题。我们证明，对于一般序列（在某些约束下，稍后详述），HM 与 GM 之间的差距无界。随后，我们构造一个简单的实验矩阵（即空间中的一个点集），使得 D-最优解（由 ℓ_∞-LW 给出）是 A-最优解的 $\Omega(n/\log^2(n))$ 乘法近似。

将 A-最优设计表述为调和平均数优化：我们从 A-最优设计开始。给定一个实验矩阵 $V = (v_1, \ldots, v_m) \in \mathbb{R}^{n\times m}$，A-最优设计的对偶问题为

$$\begin{aligned} \text{maximize} \quad & \mathbf{Tr}(W^{1/2})^2 \\ \text{subject to} \quad & v_i^T W v_i \le 1\ ,\ i = 1, \ldots, m \end{aligned}$$

其中 $W \in S_+^n$（另见 (A.15)）。注意，该约束定义了一个椭球 $\mathcal{E}(W) = \{x \in \mathbb{R}^n \mid x^T W x \le 1\}$，它包围了点集 $\{v_i\}_{i=1}^m$。我们可以将目标函数写为（忽略平方项，因为它不会改变最优解，只会改变其值）（见 (3.1)）：

$$\mathbf{Tr}\ W^{1/2} = \sum_{i=1}^n \lambda_i(W)^{1/2} = \sum_{i=1}^n (\sigma_i)^{-1} = \frac{n}{HM(\boldsymbol{\sigma})},$$

其中 $\boldsymbol{\sigma}$ 是椭球 $\mathcal{E}$（W）的半轴长度。

换言之，A-最优设计问题旨在最小化包围椭球 $\mathcal{E}$（W）半轴长度的调和平均数。因此，A-最优设计等价于以下问题

$$\begin{aligned} \text{minimize} \quad & HM(\boldsymbol{\sigma}(W)) \\ \text{subject to} \quad & v_i^T W v_i \le 1\ ,\ i = 1, \ldots, m \end{aligned}$$

D-最优设计即几何平均数最小化我们可以对 D-最优设计进行同样的处理。D-最优设计的对偶问题为

$$\begin{aligned} \text{maximize} \quad & \text{logdet}(W) \\ \text{subject to} \quad & v_i^T W v_i \le 1\ ,\ i = 1, \ldots, m \end{aligned}$$

其中$W \in \boldsymbol{S}_+^n$。与之前一样，我们可以将目标函数写为

$$\log|W| = \log(\lambda_1(W)\cdots\lambda_n(W)) = \log((\sigma_1\cdots\sigma_n)^{-2}),$$

由此可得，

$$\text{logdet}\ W = \log|W| = -2n\log((\sigma_1\cdots\sigma_n)^{1/n}) = -2n\log(GM(\boldsymbol{\sigma})),$$

其中使用了与上文相同的符号。

换句话说，最大化函数 $\text{logdet}(W)$，等同于最小化包围椭球 $\mathcal{E}(W)$ 的半轴长度的几何平均数 。因此，D-最优设计等价于以下问题

$$\begin{aligned} \text{minimize} \quad & GM(\boldsymbol{\sigma}(W)) \\ \text{subject to} \quad & v_i^T W v_i \le 1\ ,\ i = 1, \ldots, m \end{aligned}$$

注 8.1. D-最优设计的几何解释，即最小体积包围椭球，也称为外约翰椭球，是一个经过深入研究的问题。更多详情另见 [KT93；Coh+20；ZF20；Tod16；Puk06]。

8.1 调和平均数与几何平均数的比较

正如我们刚才所见，A-最优性和 D-最优性分别与同一组数值——某个包围椭球的半轴长度——的调和平均数和几何平均数相关。我们知道，对于任意向量 $\boldsymbol{x} \in \mathbb{R}^n$，$HM(\boldsymbol{x}) \leq GM(\boldsymbol{x})$，且仅当向量均匀时等号成立；此外，它们都是$\boldsymbol{x}$的连续函数。因此，自然会认为它们在某种意义上是“单调的”。例如，人们可能会认为，对于两个向量$\boldsymbol{x}$、$\boldsymbol{x}' \in \mathbb{R}^n$

$$GM(\boldsymbol{x}) \leq \alpha\, GM(\boldsymbol{x}') \iff HM(\boldsymbol{x}) \leq \alpha\, HM(\boldsymbol{x}'),$$

对于某个常数因子 α。请注意，这类性质会使这两个问题等价（至多相差某些常数因子）。我们将证明情况并非如此。下面的技术性命题表明，对于任意由$\boldsymbol{x}$ 个n 正数组成的序列，都可以构造一个相关序列$\boldsymbol{x}'$，使得$GM(\boldsymbol{x}') = GM(\boldsymbol{x})$，但$HM(\boldsymbol{x}') \ll HM(\boldsymbol{x})$。证明推迟到附录 A.4。

命题 8.2. 设 $n \in \mathbb{N}$，$\boldsymbol{x} \in \mathbb{R}^n_+$。对于任意 $0 < t < 1 \in \mathbb{R}$，都存在 $\boldsymbol{x}' \in \mathbb{R}^n_+$，使得 $GM(\boldsymbol{x}') = GM(\boldsymbol{x})$，但 $HM(\boldsymbol{x}') = t \cdot HM(\boldsymbol{x})$

请注意，这个命题比我们所需的条件稍强——为了尽可能降低 HM，我们甚至不必在 GM 上付出“代价”（尽管在算术平均值上确实需要付出代价）。

8.2 一般半正定矩阵的 A-最优设计与D-最优设计的差距

正如我们刚才所说，对于一般序列，HM 和 GM 并不近似相等（也就是说，HM 与 GM 之间的差距无界）。由于我们可以将 A-最优设计与D-最优设计解释为“HM 与 GM”的问题，因此自然引出以下命题

命题 8.3. 存在 $V \in \mathbb{R}^{n\times m}_+$，使得

$$\boldsymbol{Tr}(Vg_D^* V^T)^{-1} \gg \boldsymbol{Tr}(Vg_A^* V^T)^{-1}$$

其中，g_D^*，g_A^* 分别是 D-最优设计和 A-最优设计的最优解 *ly.*

虽然我们可以利用命题 8.2 构造这个“反例”，但我们提供了一个简单得多的实例，其差距为$\widetilde{\Omega}(n)$，这已足以展示差距。

证明。 设V为一个对角矩阵，其对角线上为 $[1, \ldots, n]$。我们已经提到，D-最优设计的最优解是实验矩阵V的 LW，即

$$(g_D^*)_l = (w(V)_\infty)_l\,.$$

根据 LW 的定义，g_D^* 满足公式（3.2）：

$$v_l^T (V^T W_\infty V)^{-1} v_l = 1,$$

其中v_l 是l的第l行V。利用V 是一个对角矩阵，且其第l个对角元素等于 ℓ这一事实，可以将最后一个方程写为

$$\ell(\ell^2 (w_\infty)_l)^{-1}\ell = 1 \implies (w_\infty)_l = 1,$$

结合归一化可得

$$g_D^* = (1/n)\mathbf{1}.$$

在g_D^*处，A-最优的值等于，

$$\mathbf{Tr}(Vg_D^*V^T)^{-1} = \sum_{i=1}^{n}(i^2/n)^{-1} = n\sum_{i=1}^{n}\frac{1}{i^2} \approx \frac{\pi^2}{6}n \approx 1.5n.$$

接下来，我们来看 $(g_0)_i = 1/i$（在归一化意义下）。我们知道

$$\sum_{i=1}^{n}\frac{1}{i} \approx log(n),$$

因此，$(g_0)_i \approx \frac{1}{log(n)\cdot i}$。在$g_0$处，A-最优的值为

$$\mathbf{Tr}(Vg_0V^T)^{-1} \approx \sum_{i=1}^{n}\left(\frac{i^2}{i\cdot log(n)}\right)^{-1} = log(n)\sum_{i=1}^{n}\frac{1}{i} \approx log^2(n).$$

因此，我们得到

$$\mathbf{Tr}(Vg_D^*V^T)^{-1} = \theta(n) \gg \mathbf{Tr}(Vg_0V^T)^{-1} = \theta(log^2(n)) \geq \mathbf{Tr}(Vg_A^*V^T)^{-1}$$

最后一个不等式源于g_A^*的最优性。因此，我们可以得出结论：A-最优设计与D-最优设计之间存在一个自然的差距，至少为*at least* $\Omega(n/log^2 n)$。□

致谢

我们感谢 Peilin Zhong 富有洞见的讨论，以及他提出无限 Lewis 权重与随机游走混合时间之间的联系。

参考文献

[Fos53] F. G. Foster.《与某些排队过程相关的随机矩阵》。载于：《数学统计年鉴》24（1953），第 355–360 页。[Lew78] D. Lewis.《L_p 的有限维子空间》。英文。载于：《数学研究》63.2（1978），第 207–212 页。网址：http://eudml.org/doc/218208。[Mat88] Peter Matthews.《球面上布朗运动的覆盖问题》。载于：《概率年鉴》16.1（1988），第 189–199 页。doi：10.1214/aop/1176991894。网址：https://doi. org/10.1214/aop/1176991894。[Moh+91] B Mohar 等。《图的拉普拉斯矩阵谱》。载于：《图论、组合数学与应用》2（1991 年 1 月），第 5364 页。[KT93] Leonid G. Khachiyan 和 Michael J. Todd.《关于多面体最大内接椭球近似的复杂度》。载于：《数学规划》61.1（1993 年 8 月），第 137–159 页。issn：0025-5610, 1436-4646。doi：10.1007/BF01582144。网址：http://link.springer.com/10.1007/BF01582144。[KR93] D. J. Klein 和 M. Randi´c.《电阻距离》。载于：《数学化学杂志》12.1（1993 年 12 月），第 81–95 页。issn：1572-8897。doi：10.1007/BF01164627。网址：https://doi. org/10.1007/BF01164627。[Cha+96] Ashok Chandra 等。《图的电阻反映其往返时间和覆盖时间》。载于：Comput. Compl 6（1996 年 12 月），第 312–340 页。doi：10.1007/BF01270385。[LNT99] I. Lukovits、S. Nikoli´c 和 N. Trinajsti´c.《正则图中的电阻距离》。载于：Interna-tional Journal of Quantum Chemistry 71.3（1999），第 217–225 页。doi：https://doi.org/10. 1002/(SICI)1097-461X(1999)71:3<217::AID-QUA1>3.0.CO;2-C。[Nem99] Arkadi Nemirovski.《关于自协调凸-凹函数》。载于：Optimization Meth-ods and Software 11.1-4（1999），第 303–384 页。doi：10.1080/10556789908805755。[NES03] Yu NESTEROV.《非光滑函数的光滑最小化》。载于：鲁汶天主教大学运营研究与计量经济学中心（CORE），CORE 讨论论文 103（2003 年 1 月）。doi：10.1007/s10107-004-0552-5。[ST03] Daniel A. Spielman 和 Shang-Hua Teng.《用于图划分、图稀疏化和求解线性系统的近线性时间算法》。2003。doi：10.48550/ARXIV.CS/ 0310051。网址：https://arxiv.org/abs/cs/0310051。[BV04] Stephen P. Boyd 和 Lieven Vandenberghe.《凸优化》。英国剑桥；纽约：剑桥大学出版社，2004。716 页。isbn：978-0-521-83378-3。[Puk06] Friedrich Pukelsheim.《试验的最优设计》。工业与应用数学学会，2006。doi：10.1137/1.9780898719109。网址：https://epubs.siam.org/doi/abs/10.1137/1.9780898719109。[GBS08] Arpita Ghosh、Stephen Boyd 和 Amin Saberi.《最小化图的有效电阻》。载于：《SIAM 评论》50.1（2008 年 1 月），第 37–66 页。issn：0036-1445，1095-7200。doi：10 . 1137/ 050645452。网址：http://epubs.siam.org/doi/10.1137/050645452。[Haz08] Elad Hazan.《半定规划的稀疏近似解》。载于：第八届拉丁美洲理论计算机科学会议论文集。LATIN'08。柏林，海德堡：Springer-Verlag，2008，第 306–316 页。isbn：3540787720。[KM09] Jonathan A. Kelner 和 Aleksander Madry.《更快地生成随机生成树》。载于：CoRR abs/0908.1448（2009）。ar Xiv：0908.1448。网址：http://arxiv.org/abs/0908.1448。

[Asa+10] Arash Asadpour 等。“一种用于非对称旅行商问题的 O(Log n/ Log Log n)-近似算法”。载于：《第二十一届 ACM-SIAM 离散算法研讨会论文集》。SODA ’10。德克萨斯州奥斯汀：工业与应用数学学会，2010年，第 379–389 页。isbn：9780898716986。[KMP10] Ioannis Koutis、Gary L. Miller 和 Richard Peng。“求解 SDD 系统的最优性逼近”。载于：CoRR abs/1003.2958（2010年）。arXiv：1003.2958。url：http://arxiv.org/ abs/1003.2958。[Chr+11] Paul Christiano 等。“无向图中的电流、拉普拉斯系统与更快的最大流近似”。载于：第四十三届 ACM 计算理论年度研讨会论文集。STOC ’11。美国加利福尼亚州圣何塞：计算机协会，2011年，第 273–282 页。isbn：9781450306911。doi：10.1145/1993636.1993674。url：https://doi.org/10.1145/1993636.1993674。[Dri+11] Petros Drineas 等。“矩阵相干性和统计杠杆值的快速近似”。载于：CoRR abs/1109.3843（2011年）。arXiv：1109.3843。url：http://arxiv.org/abs/1109.3843。[Spi11] Daniel A. Spielman。“拉普拉斯矩阵中的算法、图论与线性方程”。载于：2010年国际数学家大会论文集（ICM 2010）。印度海得拉巴，2011年6月，第 2698–2722 页。doi：10.1142/9789814324359_0164。url：http: // www.worldscientific.com/doi/abs/10.1142/9789814324359_0164。[SS11] Daniel A. Spielman 和 Nikhil Srivastava。“通过有效电阻进行图稀疏化”。载于：《SIAM 计算杂志》40.6（2011年），第 1913–1926 页。doi：10.1137/080734029。eprint：https://doi.org/10.1137/080734029。url：https://doi.org/10.1137/080734029。[CW13] Kenneth L. Clarkson 和 David P. Woodruff。《输入稀疏时间内的低秩近似与回归》。2013年4月5日。arXiv：1207.6365[cs]。url：http://arxiv.org/abs/1207. 6365。[Hac+13] Adam Hackett 等。“一种基于基尔霍夫指数优化的网络配置算法”。载于：2013 年 IEEE 第 27 届并行与分布式处理国际研讨会。2013年，第 407–417 页。doi：10.1109/IPDPS.2013.116。[HO13] Nicholas J. A. Harvey 和 Neil Olver。“Pipage 舍入、悲观估计量与矩阵集中”。载于：CoRR abs/1307.2274（2013年）。arXiv：1307.2274。url：http://arxiv. org/abs/1307.2274。[CP14] Michael B. Cohen 和 Richard Peng。 ℓ_p 基于刘易斯权重的行采样。arXiv：1412.0588。类型：文章。arXiv，2014年12月1日。url：http://arxiv.org/abs/1412.0588。[LS14] Yin Tat Lee 和 Aaron Sidford。“线性规划的路径搜索方法：在 O(√rank) 次迭代内求解线性规划及更快的最大流算法”。载于：2014 年 IEEE 第 55 届计算机科学基础年度研讨会。2014年，第 424–433 页。doi：10.1109/FOCS.2014.52。[AG15] Nima Anari 和 Shayan Oveis Gharan。“有效电阻降低流、谱稀疏树与非对称旅行商问题”。载于：2015 年 IEEE 第 56 届计算机科学基础年度研讨会论文集（FOCS）。FOCS ’15。美国：IEEE 计算机学会，2015年，第 20–39 页。isbn：9781467381918。doi：10.1109/FOCS.2015.11。url：https://doi.org/ 10.1109/FOCS.2015.11。[MST15] Aleksander Madry、Damian Straszak 和 Jakub Tarnawski。“随机生成树的快速生成与有效电阻度量”。载于：CoRR abs/1501.00267（2015年）。arXiv：1501.00267。url：http://arxiv.org/abs/1501.00267。

[AK16] Sanjeev Arora 和 Satyen Kale。“一种组合式原始-对偶半定规划方法”。载于：《*ACM*学报》63.2（2016年5月）。issn：0004-5411。doi：10.1145/2837020。url：https: // doi.org/10.1145/2837020。[Tod16] Michael J. Todd。《最小体积椭球》。宾夕法尼亚州费城：工业与应用数学学会，2016年。doi：10.1137/1.9781611974386。url：https://epubs.siam.org/doi/abs/10.1137/1.9781611974386。[Ale+17] Vedat Levi Alev 等。“使用有效电阻进行图聚类”。载于：*CoRR* abs/1711.06530（2017年）。arXiv：1711.06530。url：http://arxiv.org/abs/1711.06530。[Dur+17] David Durfee 等。“比矩阵乘法更快地采样随机生成树”。载于：第四十九届 *ACM SIGACT* 计算理论研讨会论文集。STOC 2017。加拿大蒙特利尔：计算机协会，2017年，第 730–742 页。isbn：9781450345286。doi：10.1145/3055399.3055499。url：https://doi.org/10.1145/ 3055399.3055499。[MSJ19] Silviu Maniu、Pierre Senellart 和 Suraj Jog。“真实世界图数据树宽的实验研究（扩展版）”。载于：*CoRR* abs/1901.06862（2019年）。arXiv：1901. 06862。url：http://arxiv.org/abs/1901.06862。[Coh+20] Michael B. Cohen 等。“一种用于近似约翰椭球的近最优算法”。载于：*arXiv:1905.11580 [cs]*（2020年2月18日）。url：http://arxiv.org/abs/1905.11580。[Jia+20] Haotian Jiang 等。“一种更快的半定规划内点法”。载于：*CoRR*abs/2009.10217（2020年）。arXiv：2009.10217。url：https://arxiv.org/abs/2009.10217。[Mas+20] Anum Masood等。“用于调度通信

-

异构计算环境中的密集图”。载于：集群计算23.2（2020），第1419–1438页。[ZF20] Renbo Zhao 和 Robert M. Freund。涉及对数齐次障碍函数的凸复合优化中 *Frank-Wolfe* 方法的分析。2020。doi: 10 .48550/ ARXIV.2010.08999。url: https://arxiv.org/abs/2010.08999。[Faz+21] Maryam Fazel 等。“高精度计算刘易斯权重”。载于：*CoRR* abs/2110.15563（2021）。url: https://arxiv.org/abs/2110.15563。[Hua+22] Baihe Huang 等。“更快地求解半定规划：鲁棒内点法框架与高效实现”。载于：第*63*届 *IEEE*计算机科学基础年会（*FOCS 2022*），美国科罗拉多州丹佛，*2022*年*10*月*31*日至*11*月*3*日。IEEE，2022，第233–244页。doi: 10.1109/FOCS54457.2022.00029。url: https://doi.org/10.1109/FOCS54457.2022.00029。

A 附录

A.1 背景：凸优化分析

本文讨论的所有优化问题都是凸优化问题。在分析最优解时，我们将采用以下形式问题的解题步骤：

$$\begin{array}{ll}\text{minimize} & f(g)\\ \text{subject to} & \mathbf{1}^T g = 1\ ,\ g \in \mathbb{R}_+^m\end{array}$$

当$f(g)$为凸函数时。这是一个众所周知的恒等式：当且仅当

$$\nabla f(g)^T \cdot (\widehat{g} - g) \geq 0\ ,\ \forall\, \widehat{g}\colon\ \mathbf{1}^T \widehat{g} = 1.$$

这与下式相同：

$$\nabla f(g)^T \cdot (e_l - g) \geq 0\ ,\ \forall\, l = 1 \ldots m$$

（将g 展开为初等基表示）。因此，g 的最优性准则为 ，

$$\frac{\partial f}{\partial g_l}(g) - \nabla f(g)^T \cdot g \geq 0\ ,\ \forall\, l = 1 \ldots m.$$

因此，对于上述形式的凸问题，我们将采用以下步骤：

1. 求 f 的偏导数，即 $^{\partial}f(g)\ _{\partial g_l}$。 ——

2. 求 $\langle \nabla f(g),\ g\rangle = \nabla f(g)^T \cdot g$。

3. 推导最优性准则：

$$\frac{\partial f}{\partial g_l} - \nabla f(g)^T \cdot g \geq 0\ ,\ \forall\, l = 1 \ldots m \tag{A.1}$$

A.2 背景：实验最优设计

这里简要概述实验最优设计。更多信息请参见 [BV04；Puk06]。

假设你想估计一个向量 $x \in \mathbb{R}^n$，使用以下实验：在每一轮 $i = 1$，.....，m，你可以从给定的 p 个可能选项中选择一个测试 向量——$v_1, \ldots\ldots$，v_p——然后得到 $y_i = v_{j_i}^T x + w_i$，其中 $j_i \in [p]$，而 w_i 是某种高斯噪声（即均值为零、方差为单位方差的独立高斯随机变量）。通常假设 $V = (v_1, \ldots$ ...，$v_p)$ 是一个满秩矩阵，但也可以很容易地推广。最优估计由最小二乘解给出：

$$\widehat{x} = \underbrace{\left(\sum_{i=1}^m v_{j_i} v_{j_i}^T\right)^{-1}}_{E} \cdot \sum_{i=1}^m y_i v_{j_i}.$$

对于任意估计x，我们定义估计误差$e = \widehat{} x - x$，其对应的协方差矩阵为E。最优设计的目标是相对于某个偏序（例如洛文纳序）最小化E。一般而言，当m较大时，最优设计可能是一个困难的组合问题。为了解释其原因，令m_j 表示我们选择v_j的轮数，并使用m_j表示E：

$$E = \left(\sum_{i=1}^m v_{j_i} v_{j_i}^T\right)^{-1} = \left(\sum_{j=1}^p m_j v_j v_j^T\right)^{-1}$$

这表明，最优设计可以用整数变量m_j来表示，并满足它们的和为m这一约束。如果m很大，那么我们可能需要处理相对较小的整数，从而使问题变得相当困难。不过，我们可以按以下方式对其进行松弛。定义$g_j = m_j/m$，并用g_j表示E：

$$E(g) = \frac{1}{m}\left(\sum_{j=1}^{p} g_j v_j v_j^T\right)^{-1}$$

现在，我们不再对整数进行搜索，而是令g_j为实数，并满足它们的和为 1 的约束。因此，放宽后的最优设计要求在某种偏序下最小化E $\mathbf{1}^T g = 1$。例如，A-最优设计和 D-最优设计分别按照迹和行列式进行最小化。我们将V称为实验矩阵，并将$E_V(g) = V \cdot \mathrm{diag}(g) \cdot V^T$记作$E_V(g)^{-1}$，其中它是误差协方差矩阵。

A.3 D-最优设计问题

这里我们通过凸分析方法介绍一种求解 D-最优设计问题的替代方案。我们重点讨论拉普拉斯矩阵的情形，但该方法可以推广到任意实验矩阵。我们采用 A.1 中所述的步骤。我们将拉普拉斯矩阵上的 D-最优设计问题定义为问题（D-最优设计）的一种松弛，具体如下：

定义 A.1。给定一个无向图G，记加权图$G(g)$的拉普拉斯矩阵为L_g。*LD* 最小化问题旨在确定如何重新赋予各边权重，以使L_g^+的对数行列式 最小：

$$\begin{array}{ll} \text{minimize} & \mathrm{logdet}(L_g^+ + (1/n)\mathbf{1}\mathbf{1}^T) \\ \text{subject to} & \mathbf{1}^T g = 1\ ,\ g \geq 0 \end{array}$$

注 A.2。你可能会想知道为什么我们选择这个目标函数。注意，L_g^+有一个零特征值，而我们只关注正特征值。不过，不难证明：

$$\lambda(L_g^+ + 1/n\mathbf{1}\mathbf{1}^T) = \begin{cases} \lambda(L_g^+) & \text{if } \lambda(L_g^+) \neq 0 \\ 1 & \text{if } \lambda(L_g^+) = 0 \end{cases}$$

因此，我们的目标函数就是L^+所有正特征值的乘积。我们还将这一记法稍作扩展，写作 $|L_g^+ + (1/n)\mathbf{1}\mathbf{1}^T|$，即 $|L_g^+|_+$。

我们将目标函数写作$LD_G(g) = \mathrm{logdet}(L_g^+)$。当上下文明确指的是图的情形时，我们将简写为$LD(g)$。注意，对于一般的实验矩阵$V$，目标函数等于$LD_V(g) = \mathrm{logdet}(E_V(g)^+)$。由此得到以下结果：

引理 A.3。给定一个具有边关联矩阵 $B \in \mathbb{R}^{n\times m}$ 的无向图 G，使 $LD_G(g)$ 函数最小化的权重向量是 $LW(B^T)$，差一个归一化因子。更一般地，对于实验矩阵 $V \in \mathbb{R}^{n\times m}$，$V$ 的 D-最优性解是 $LW(V^T)$，差一个归一化因子。

我们先来计算偏导数。我们将用到以下事实：设可逆对称矩阵 $X(t)$ 是参数$t \in \mathbb{R}$ 的可微函数。那么我们有

$$\frac{\partial X^{-1}}{\partial t} = -X^{-1}\frac{\partial X}{\partial t}X^{-1} \tag{A.2}$$

我们还会用到雅可比公式——给定可微矩阵$X(t)$：

$$\frac{\partial |X|}{\partial t} = |X|\,\mathbf{Tr}(X^{-1}\frac{\partial X}{\partial t}) \tag{A.3}$$

现在，如果我们写出$L_g = \sum_{l=1}^{m} g_l b_l b_l^T$的一秩分解，显然

$$\frac{\partial L_g}{\partial g_l} = b_l b_l^T \tag{A.4}$$

利用上述结果和公式（3.2），可得:

$$\begin{aligned}
\frac{\partial L_g^+}{\partial g_l} &= \frac{\partial (L_g + 1/n\mathbf{1}\mathbf{1}^T)^{-1}}{\partial g_l} = -(L_g + 1/n\mathbf{1}\mathbf{1}^T)^{-1}\frac{\partial L_g}{\partial g_l}(L_g + 1/n\mathbf{1}\mathbf{1}^T)^{-1} && \text{using (A.2)} \\
&= -(L_g^+ + 1/n\mathbf{1}\mathbf{1}^T)\frac{\partial L_g}{\partial g_l}(L_g^+ + 1/n\mathbf{1}\mathbf{1}^T) && \text{from (3.2)} \\
&= -(L_g^+ + 1/n\mathbf{1}\mathbf{1}^T) b_l b_l^T (L_g^+ + 1/n\mathbf{1}\mathbf{1}^T) && \text{using (A.4)} \\
&= -L_g^+ b_l b_l^T L_g^+ && \text{(A.5)}
\end{aligned}$$

此处用到了$b_l^T \mathbf{1} = 0$这一事实。

现在，利用雅可比公式(A.3)和基本的链式法则，可得:

$$\begin{aligned}
&\frac{\partial |L_g^+|}{\partial g_l} = |L_g^+| \operatorname{\mathbf{Tr}}((L_g^+)^{-1}\frac{\partial L_g^+}{\partial g_l}) \\
&\implies \frac{\partial \log |L_g^+|}{\partial g_l} = |L_g^+|^{-1}\frac{\partial |L_g^+|}{\partial g_l} = \operatorname{\mathbf{Tr}}((L_g^+)^{-1}\frac{\partial L_g^+}{\partial g_l})
\end{aligned}$$

结合前两个公式可得:

$$\begin{aligned}
\frac{\partial \log |L_g^+|}{\partial g_l} &= -\operatorname{\mathbf{Tr}}((L_g^+)^{-1} L_g^+ b_l b_l^T L_g^+) \\
&= -\operatorname{\mathbf{Tr}}((b_l b_l^T L_g^+) = -\operatorname{\mathbf{Tr}}(b_l^T L_g^+ b_l) = -R_l(g)
\end{aligned} \tag{A.6}$$

注释 A.4. 对于一般情形，可得:

$$\begin{aligned}
\frac{\partial LD_V(g)}{\partial g_l} &= -\operatorname{\mathbf{Tr}}(E_V(g)^{-1} v_l v_l^T) \\
&= -\operatorname{\mathbf{Tr}}((v_l^T E_V(g))^{-1} v_l) = v_l^T (V \cdot \operatorname{diag}(g) \cdot V^T)^{-1} v_l
\end{aligned} \tag{A.7}$$

并且，用最小二乘表示为:

$$\frac{\partial LD_V(g)}{\partial g_l} = -g_l^{-1} \cdot \tau_l(W_g^{1/2} V^T) \tag{A.8}$$

接下来，我们需要计算 $\langle \nabla LD(g)、\, g \rangle = \nabla LD(g)^T \cdot g$ 。我们将使用以下“技巧”——利用行列式的乘法性以及伪逆的性质:

$$\begin{aligned}
LD(\alpha g) &= \log |L_{\alpha g}| = \log |(B \cdot \operatorname{diag}(\alpha g) \cdot B^T)^+| \\
&= \log |\alpha^{-1} L_g^+| \\
&= \log(\alpha^{-(n-1)} |L_g^+|) && \text{because } \operatorname{rank}(L_g) = n-1 \\
&= LD(g) - (n-1)\log(\alpha) && \text{logarithm laws}
\end{aligned}$$

对两边关于 α求导，我们得到:

$$\nabla LD(\alpha g)^T \cdot g = \frac{\partial}{\partial \alpha}(LD(g) - (n-1)\log(\alpha)) = -\frac{n-1}{\alpha}$$

在 $\alpha = 1$处计算该表达式，得到

$$\nabla LD(g)^T \cdot g = -(n-1) \tag{A.9}$$

注记 A.5. 不难看出，对于秩为 d 的一般矩阵 V，$rank(E_V) = rank(V) = d$，且$|\alpha^{-1}E_V^{-1}| = \alpha^{-d}|E_V^{-1}|$。因此我们有，

$$LD_V(\alpha g) = LD_V(g) - d\log(\alpha),$$

相应地，

$$\nabla LD_V(g)^T \cdot g = -d.$$

现在，我们可以将上述所有内容代入最优性准则（公式（A.1）），得出g 是$LD(g)$的最优解，当且仅当：

$$\frac{\partial LD(g)}{g_l} + n - 1 = -R_l(g) + n - 1 \geq 0\ ,\ l = 1, \ldots, m$$

or,

$$R_l(g) \leq n - 1\ ,\ \forall\ l = 1, \ldots, m. \tag{A.10}$$

对于一般的V,

$$g_l^{-1} \cdot \tau_l(W_g^{1/2}V^T) \leq \text{rank}(V)\ ,\ l = 1, \ldots, m \tag{A.11}$$

最后，我们可以证明 ℓ_∞-LW 确实是$LD(g)$的最优解。我们将对一般情形进行证明。设$C = W_g^{1/2}V^T$。假设V的秩为d。令$g_{\ell w} := (1/d)w_\infty(V^T)$。首先，根据公式（3.1）

$$\sum_{l=1}^{m}(g_{\ell w})_l = (1/d)\sum_{l=1}^{m} w_\infty(V^T)_l = (1/d) \cdot d = 1$$

因此$g_{\ell w}$ 是可行的。接下来，我们需要证明它符合最优性准则。根据公式（3.2），对于未归一化 LW,

$$v_l^T(VW_\infty V^T)^+ v_l = 1,\ l = 1, \ldots, m.$$

由于我们按 $(1/d)$ 进行归一化，因此对于$g = g_{\ell w}$，有：

$$v_l^T(VW_g V^T)^+ v_l = d,\ l = 1, \ldots, m.$$

因此，对于$g = g_{\ell w}$，我们有：

$$\tau_l(W_\infty^{1/2}V^T) = g_l^{1/2} \cdot v_l^T(V \cdot W_\infty V^T)^+ v_l \cdot g_l^{1/2} = d \cdot g_l,\ \forall\ l = 1, \ldots, m.$$

将后者代入公式（A.11）可得

$$g_l^{-1} \cdot \tau_l(W_g^{1/2}V^T) = d \cdot g_l \cdot g_l^{-1} = d \leq d. \tag{A.12}$$

这恰好符合最优性准则，并证明了引理 A.3。注意，事实上$g_{\ell w}$ 满足最优性准则。

A.4 命题 8.2 的证明

证明。 设$\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$。定义$\boldsymbol{x}' = (px1, x2/p, x3, \ldots, x_n)$。事实上，

$$GM(\boldsymbol{x}') = (\Pi_i\, x_i')^{1/n} = (\Pi_i\, x_i)^{1/n} = GM(\boldsymbol{x})$$

令 $S = \sum_{i=1}^{n} \frac{1}{x_i}$，对 S' 也作类似定义。显然，调和平均数$(\boldsymbol{x}) = n/S$，因此若要使调和平均数增加 t，我们应将S 减少t。

具体来说，我们需要满足

$$S' = \sum_{i=1}^{n} \frac{1}{x_i'} = S/t = (1/t) \sum_{i=1}^{n} \frac{1}{x_i}.$$

根据$\boldsymbol{x}'$ 的定义可得

$$S' = \frac{1}{px_1} + \frac{p}{x_2} + \sum_{i=3}^{n} \frac{1}{x_i} = \frac{1}{px_1} + \frac{p}{x_2} + S - \frac{1}{x_1} - \frac{1}{x_2} = S + (1/p - 1)\frac{1}{x_1} + (p - 1)\frac{1}{x_2} = S/t,$$

这是关于p的二次方程:

$$(1 - 1/t)S - \left(\frac{1}{x_1} + \frac{1}{x_2}\right) + \frac{1}{x_1}\frac{1}{p} + \frac{1}{x_2}p = 0. \tag{A.13}$$

我们可以作变量替换$y_1 = \frac{1}{x_1}, y_2 = \frac{1}{x_2}, t' = (1/t - 1)S$，得到:

$$y_2 \cdot p^2 - (t' + y_1 + y_2) \cdot p + y_1 = 0,$$

其解为:

$$\begin{aligned} p &= \frac{1}{2y_2}\left[(t' + y_1 + y_2) + \sqrt{(t' + y_1 + y_2)^2 - 4y_2y_1}\right] \\ &= \frac{1}{2y_2}\left[(t' + y_1 + y_2) + \sqrt{t'^2 + (y_1 + y_2)^2 + 2t'(y_1 + y_2) - 4y_2y_1}\right] \\ &= \frac{1}{2y_2}\left[(t' + y_1 + y_2) + \sqrt{t'^2 + 2t'(y_1 + y_2) + (y_1 - y_2)^2}\right] \end{aligned} \tag{A.14}$$

注意，t'、y_1,y_2 均为非负数，因此p 有定义。

由此可知，对于任意$t < 1$，利用方程（A.14）可得到p_t，进而可以构造$\boldsymbol{x}'$，使得$GM(\boldsymbol{x}) = GM(\boldsymbol{x}')$，且$HM(\boldsymbol{x}') = t \cdot HM(\boldsymbol{x})$。

值得一提的是，在这一构造中，算术平均值如何变化。事实上，

$$AM(\boldsymbol{x}') = \frac{1}{n}\sum_{i=1}^{n} x_i' = \frac{1}{n}(\sum_i x_i - x_1 - x_2 + px_1 + x_2/p) = AM(\boldsymbol{x}) + \frac{1}{n}((p_t - 1)x_1 + (1/p_t - 1)x_2).$$

有趣的是， 我们得到算术平均数总是增加。要了解原因，请考虑方程（A.13 ）

$$(1/p - 1)\frac{1}{x_1} + (p - 1)\frac{1}{x_2} = S(1/t - 1)$$

将两边同时乘以 x_1x_2，得到

$$(1/p - 1)x_2 + (p - 1)x_1 = S(1/t - 1)x_1x_2 > 0.$$

□

注释 A.6. 值得注意的是， Δ算术平均值 $=$ 算术平均值(x') – 算术平均值(x) 的行为。当调和平均数$'$ 调和平均数$= t$时，算术平均值的差为（见上一个证明的最后一个等式）

$$\Delta AM = S\,x_1x_2\frac{1}{n}(1/t-1) = \theta\left(\frac{1}{n}(1/t-1)\right) \quad \text{(we refer to } \boldsymbol{x} \text{ as constant.)}$$

例如，如果$t = \theta(1/n)$，那么，

$$\frac{HM'}{HM} = \theta\left(\frac{1}{n}\right),\ \Delta AM = \theta(1).$$

因此，虽然HM 的值乘以了n，但算术平均值 的值变化相对较小。换言之，HM 值的下降超过了算术平均值值的轻微上升。

A.5 ERMP 的拉格朗日对偶问题

我们首先将 ERMP 问题改写为:

$$\begin{array}{ll} \text{minimize} & n\,\mathbf{Tr}\,X^{-1} - n \\ \text{subject to} & X = \sum_{l=1}^m g_l b_l b_l^T + (1/n)\mathbf{1}\mathbf{1}^T \\ & \mathbf{1}^T g = 1, g \geq 0 \end{array}$$

变量为$g \in \mathbb{R}^m$ 和$X = X^T \in \mathbb{R}_+^{n\times n}$。对于等式约束，定义对偶变量$Z \in \boldsymbol{S}_+^n$ 、$v \in \mathbb{R}$；对于$g \geq 0$ 的非负约束，定义对偶变量 $\lambda \in \mathbb{R}^m$。拉格朗日函数为

$$L(X,g,Z,v,\lambda) = n\,\mathbf{Tr}\,X^{-1} - n + \mathbf{Tr}[Z(X - \sum_{l=1}^m g_l b_l b_l^T + 1/n\mathbf{1}\mathbf{1}^T)] + v(\mathbf{1}^T g - 1) - \lambda^T g.$$

于是，对偶函数为，

$$\begin{aligned} h(Z,v,\lambda) &= \inf_{X\geq 0,g} L(X,g,Z,v,\lambda) \\ &= \inf_{X\geq 0} \mathbf{Tr}[nX^{-1} + ZX] + \inf_g \left(\sum_{l=1}^m g_l(b_l^T Z b_l + v - \lambda_l)\right) - n - v - (1/n)\mathbf{1}^T Z\mathbf{1} \\ &= \begin{cases} -v - (1/n)\mathbf{1}^T Z\mathbf{1} + 2\,\mathbf{Tr}(nZ)^{1/2} - n & \text{if } b_l^T Z b_l + v = \lambda_l,\ Z \geq 0 \\ -\infty & \text{otherwise} \end{cases} \end{aligned}$$

最后一行可以解释如下。左侧项$\mathbf{Tr}[nX^{-1} + ZX]$ 无下界，除非$Z \geq 0$ （否则，我们可以取X_1 为 Z 对应于负特征值的特征向量乘以任意大的常数，并补全$X_{2\ldots n}$ ，使得X 为半正定矩阵）。现在，如果$Z > 0$ ，使其最小化的唯一X 为$X = (Z/n)^{-1/2}$ （通过对X求导很容易证明），因此其最优值为，

$$\mathbf{Tr}[nX^{-1} + ZX] = \mathbf{Tr}[n(Z/n)^{1/2} + Z(Z/n)^{-1/2}] = 2\,\mathbf{Tr}(nZ)^{1/2}.$$

如果Z 是半正定但不是正定矩阵，这仍然是最优值，因为我们可以取一列正定矩阵Z_i ，使其收敛到Z ；并且由于这是Z 上的线性函数，定义在闭集上，因此得到相同的最优值。

此外，右侧项 $\sum_{l=1}^m \mathrm{g}_l(\mathrm{b}_l^T\mathrm{Z}\mathrm{b}_l + \mathrm{v} - \lambda_l)$ 也无下界，除非 $\forall \mathrm{l}, \mathrm{b}_l^T\mathrm{Z}\mathrm{b}_l + \mathrm{v} - \lambda_l = 0$。否则，取 $\mathrm{g} = \alpha \mathrm{e}_k$，使得 $\mathrm{b}_k^T\mathrm{Z}\mathrm{b}_k + \mathrm{v} - \lambda_k = \mathrm{c} \neq 0$，从而得到 $\sum_{l=1}^m \mathrm{g}_l(\mathrm{b}_l^T\mathrm{Z}\mathrm{b}_l + \mathrm{v} - \lambda_l) = \alpha \cdot \mathrm{c}$，该项作为 α 的函数无下界。

拉格朗日对偶问题为:

$$\begin{aligned}&\text{maximize } h(Z, v, \lambda)\\&\text{subject to } b_l^T Z b_l \le v, Z \succeq 0, \lambda > 0\end{aligned}$$

使用上文推导出的 $h(Z,v,\lambda)$ 公式，并消去 λ（因为它不参与目标函数），我们得到以下形式:

$$\begin{aligned}&\text{maximize } -v - (1/n)\mathbf{1}^T Z\mathbf{1} + 2\,\mathbf{Tr}(nZ)^{1/2} - n\\&\text{subject to } b_l^T Z b_l \le v, Z \succeq 0\end{aligned}$$

这是另一个以 Z、v 为变量的凸优化问题。可以验证，如果$X^* = L^* + 1/n\mathbf{1}\mathbf{1}^T$（其中 $L^* = \sum_{l=1}^m g_l^* b_l b_l^T$是最优加权拉普拉斯矩阵）是原始 ERMP 的最优解，那么

$$Z^* = n(X^*)^{-2} = n(L^* + 1/n\mathbf{1}\mathbf{1}^T)^{-2}\ ,\ v^* = \max_l b_l^T Z^* b_l$$

就是对偶问题的最优解（这可由我们的分析轻易得出）。我们可以利用它进一步化简对偶问题。对于任意 g，都有 $(L_g + 1/n\mathbf{1}\mathbf{1}^T)^{-2}\mathbf{1} = \mathbf{1}$（见公式（3.2））。因此，$Z^*\mathbf{1} = n\mathbf{1}$，这意味着最优解满足 $Z\mathbf{1} = n\mathbf{1}$。所以，我们可以添加这一约束而不改变最优值（即 $\mathcal{K}^*$）。得到的对偶问题现在为,

$$\begin{aligned}&\text{maximize } -v - (1/n)\mathbf{1}^T Z\mathbf{1} + 2\,\mathbf{Tr}(nZ)^{1/2} - n\\&\text{subject to } b_l^T Z b_l \le v,\\&\qquad Z \succeq 0\ ,\ Z\mathbf{1} = n\mathbf{1}\end{aligned}$$

令$Y = Z - \mathbf{1}\mathbf{1}^T$进行变量替换，可得到以下形式,

$$\begin{aligned}&\text{maximize } -v + 2\,\mathbf{Tr}(nY)^{1/2}\\&\text{subject to } b_l^T Y b_l \le v,\\&\qquad Y \succeq 0\ ,\ Y\mathbf{1} = 0\end{aligned}$$

令$W = Y/v$（注意，根据我们的约束v 为正数），则问题变为:

$$\begin{aligned}&\text{maximize } -v + 2(nv)^{1/2}\,\mathbf{Tr}\,W^{1/2}\\&\text{subject to } b_l^T W b_l \le 1,\\&\qquad W \succeq 0\ ,\ W\mathbf{1} = 0\end{aligned}$$

现在，对于正数常量C，$f(v) = -v + 2C\cdot v^{1/2}$的最大值为$v = C^2$。取$C = n^{1/2}\,\mathbf{Tr}\,W^{1/2}$，问题便变为

$$\begin{aligned}&\text{maximize } n(\mathbf{Tr}\,W^{1/2})^2\\&\text{subject to } b_l^T W b_l \le 1,\\&\qquad W \succeq 0\ ,\ W\mathbf{1} = 0\end{aligned}$$

最后，通过变量替换为 $V = W^{1/2}$，我们可以将对偶问题表示为:

$$\begin{aligned}&\text{maximize } n(\mathbf{Tr}\,V)^2\\&\text{subject to } b_l^T V^2 b_l = \|Vb_l\| \le 1,\\&\qquad V \succeq 0\ ,\ V\mathbf{1} = 0\end{aligned} \tag{A.15}$$

A.6 ERMP 的半正定规划表述

如 [GBS08; BV04], 所示, ERMP 可以表述为一个半定规划:

$$\begin{array}{ll}\text{minimize} & n\,\mathbf{Tr}\,Y \\ \text{subject to} & g \geq 0,\ \mathbf{1}^T g = 1 \\ & \begin{bmatrix} L_g + (1/n)\mathbf{1}\mathbf{1}^T & I \\ I & Y \end{bmatrix} \geq 0\end{array}$$

定义系数矩阵

$$C \coloneqq n\begin{bmatrix} 0 & 0 \\ 0 & -I \end{bmatrix}$$

以及约束条件

$$A \coloneqq \begin{bmatrix} I & 0 \end{bmatrix}, b \coloneqq \begin{bmatrix} L_g + (1/n)\mathbf{1}\mathbf{1}^T & I \end{bmatrix}$$

则上述表述可写成如下形式:

$$\begin{array}{ll}\text{maximize} & C \bullet X \\ \text{subject to} & g \geq 0,\ \mathbf{1}^T g = 1 \\ & AX = b \\ & X \geq 0, X = X^T\end{array}$$

即为 [Haz08] 表述。

A.7 SEV 最优性准则

根据之前的分析，我们将使用熟悉的条件，推导$f(g) = \mathbf{Tr}\, e^{L_g^+}$的最优性准则,

$$g \text{ is optimal for } f \iff \nabla f(g)^T (e_l - g) \geq 0 \text{ , for } l = 1 \ldots m$$

我们先计算$f(g)$的梯度。根据链式法则，可得:

$$\frac{\partial f}{\partial g_l} = \mathbf{Tr}[e^{L_g^+} \frac{\partial L_g^+}{\partial g_l}] = -\,\mathbf{Tr}[e^{L_g^+} L_g^+ b_l b_l^T L_g^+]$$

其中，我们使用了公式 (A.5) 来计算L_g^+的梯度。

接下来，我们要计算 $\nabla f(g)^T g$。我们将使用标准的“技巧”:

$$\begin{aligned} & f(\alpha g) = \mathbf{Tr}[e^{L_{\alpha g}^+}] = \mathbf{Tr}[e^{L_g^+/\alpha}] \\ \implies & \frac{\partial}{\partial \alpha} f(\alpha g) = \frac{\partial}{\partial \alpha}\,\mathbf{Tr}[e^{L_g^+/\alpha}] = \mathbf{Tr}[e^{L_g^+/\alpha} \frac{\partial}{\partial \alpha}(L_g^+/\alpha)] = \mathbf{Tr}[e^{L_g^+/\alpha} L_g^+(-1/\alpha^2)] \end{aligned}$$

并令 $\alpha = 1$, 可得

$$\nabla f(g)^T g = \frac{\partial}{\partial \alpha} f(\alpha g)\Big|_{\alpha=1} = -\,\mathbf{Tr}[e^{L_g^+} L_g^+].$$

将它们代入最优性条件，可得以下形式:

$$-\,\mathbf{Tr}[e^{L_g^+} L_g^+ b_l b_l^T L_g^+] + \mathbf{Tr}[e^{L_g^+} L_g^+] \geq 0$$

利用迹的可加性，我们得到方程 (6.3):

$$\mathbf{Tr}[e^{L_g^+} L_g^+ (I - b_l b_l^T L_g^+)] \geq 0$$

A.8 上部 LT 的图

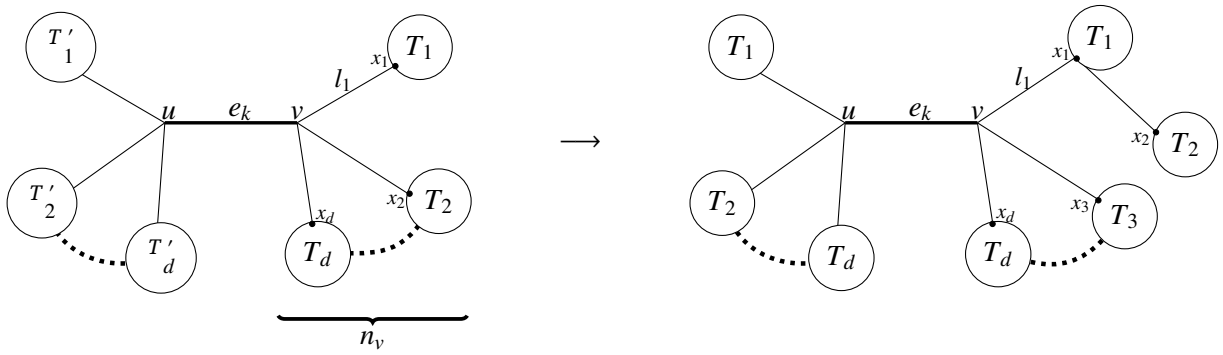


图5：第一种情况的上部 LT 描述

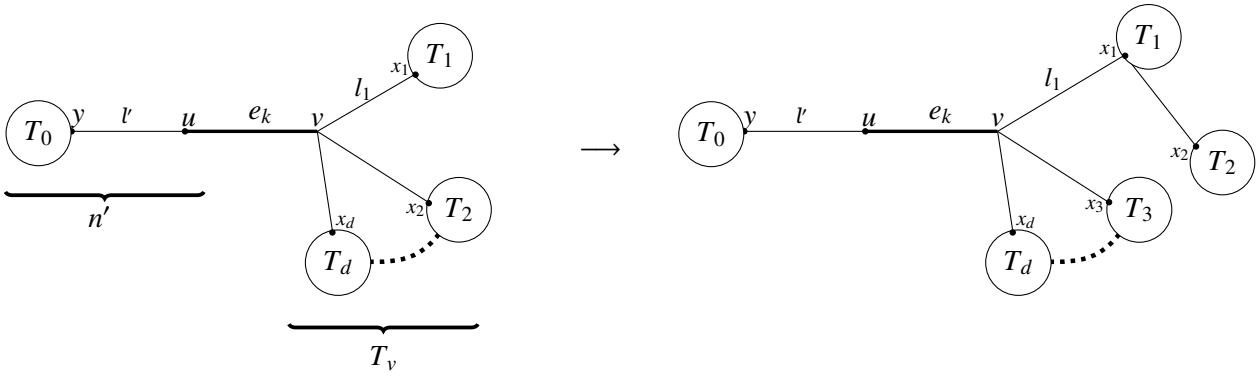


图6：情形2.1的上部 LT 描述

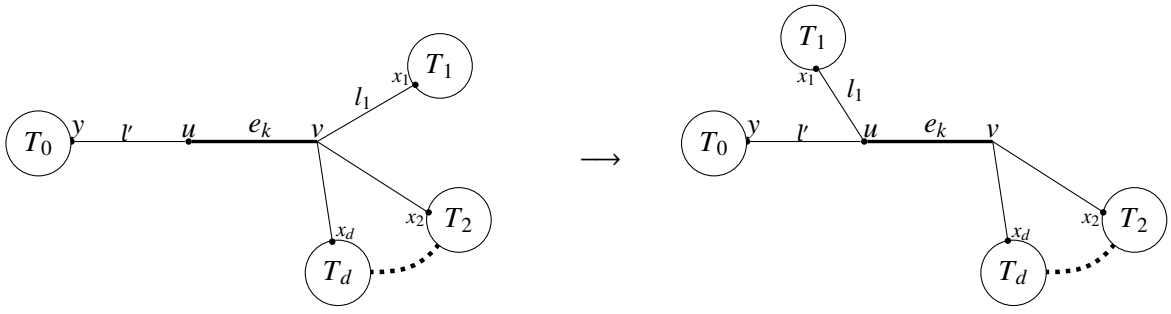


图7：情形2.2的上部 LT 描述