arXiv:2302.03143v1 [cs.DS] 6 Feb 2023

低曲率单调 k-次模函数的稀疏化*

ˇ Jannik Kudla Stanislav Zivn´y 牛津大学† 牛津大学
janniku@google.com standa.zivny@cs.ox.ac.uk

2023年2月8日

摘要

Bencz´ur 和 Karger 首创了图中的割稀疏化 [STOC'96], ，这是一个基础性课题，具有广泛的应用，已在图和超图、组合与谱情形，以及加性和乘性误差界等方面得到研究。Rafiey 和 Yoshida 最近研究了可分解次模函数的稀疏化 [AAAI'22]。我们扩展了他们的工作，提出了一种高效算法，用于构造低曲率单调 k-次模函数的稀疏化器。

1 引言

在保留所有割的值不变的同时"稀疏化图"（即减少边的数量）这一思想，源自一篇颇具影响力的论文[6]。其最初目的是加快割问题以及更一般的图问题的算法。事实证明，这一概念影响深远，并衍生出多种推广和扩展，从图割 [5, 7, 2] 到超图中割的稀疏化器 [1, 2], 、谱稀疏化 [25, 31], 、其他谓词的稀疏化 [42, 41, 40,18, 27, 39, 24], 以及加性稀疏化 [14], 。

图的割函数是次模函数的一个重要例子，我们现在对其进行定义。设 E 为一个有限集合。在 E 的子集上定义的（集合）函数 $F\colon 2^E \to \mathbb{R}$ 称为次模函数，如果

$$F(S \cap T) + F(S \cup T) \leq F(S) + F(T) \qquad \forall S, T \subseteq E. \tag{1}$$

次模性是组合优化中的一个基本概念，在计算机科学和经济学的各个领域都有应用 [30, 45, 38, 16]。次模函数的一个等价定义体现了边际收益递减这一思想。

$$F(T \cup \{e\}) - F(T) \leq F(S \cup \{e\}) - F(S) \qquad \forall S \subseteq T \subseteq E, e \in E \setminus T. \tag{2}$$

如果$F = \sum_{i=1}^{N} f_i$，则集合函数 F是可分解的，其中 $f_i\colon$ 对于每个 $i \in [N] =$，都有 $2^E \to \mathbb{R}\{1, \ldots, N\}$，且 E是一个大小为 $n = |E|$的有限集合。图中的割函数就是一个例子

*本研究完全由 UKRI EP/X024431/1 资助。为实现开放获取，作者已对因本次投稿产生的任何作者接受稿版本申请 CC BY 公共版权许可。所有数据均已在本文的结果部分完整提供。

†现就职于谷歌。

，它是一个可分解次模函数，其中单个函数的数量 N 等于图中边的数量。（即使图是有向图且边权重非负，这一点仍然成立。）

Rafiey 和 Yoshida [34] 考虑了如下针对 F 的自然稀疏化问题。[1]给定容差参数 ε，$\delta \in (0, 1)$，寻找一个向量 $w \in \mathbb{R}^N$，称为 ε-稀疏化器（或简称稀疏化器），使得函数 $F' = \sum_{i=1}^{N} w_i f_i$ 以至少 $1-\delta$ 的概率满足，

$$(1-\varepsilon)F'(S) \;\le\; F(S) \;\le\; (1+\varepsilon)F'(S) \qquad \forall S \subseteq E, \tag{3}$$

并且 $\mathsf{size}(w)$（即 w 的非零项集合）尽可能小。[34] 中的思想是，对于每个 $i \in [N]$，以与比率 κ_i 成比例的概率对函数 f_i 进行采样

$$p_i = \max_{\substack{S \subseteq E \\ F(S) \neq 0}} \frac{f_i(S)}{F(S)}. \tag{4}$$

如果对 f_i 进行采样，即决定将 f_i 纳入稀疏化器，则以权重 $1/\kappa_i$将其纳入稀疏化器，从而使其期望权重等于 $\mathbb{E}[w_i] = \kappa \cdot 1/\kappa_i = 1$——即其在初始分解中的权重。从统计学术语来看，该采样过程是无偏的。 [34] 的作者证明了如下结论。

定理 1 （[34]）。设 $F = \sum_{i=1}^{N} f_i$，其中每个 $f_i : 2^E \to \mathbb{R}$ 都是次模的。对于任意ε，$\delta \in (0, 1)$，存在一个向量 $w \in \mathbb{R}^N$，使得

(i) $\mathbb{P}[w$ 是一个 ε稀疏化器$] \ge 1-\delta$;

(ii) $\mathbb{E}[\mathsf{size}(w)] = \mathcal{O}\left(\frac{n}{\varepsilon^2}\sum_{i=1}^{N} p_i\right)$, *where* $p_i = \max_{\substack{S \subseteq E \\ F(S) \neq 0}} \frac{f_i(S)}{F(S)}$.

计算 p_i，以及在许多有趣的情况下对其进行近似，毫无疑问是构造稀疏化器过程中最困难的一步。我们将 p_i称为峰值贡献——因为 p_i 描述了在 0 到 1 的范围内，当选择集合 $S \subseteq E$ 而非 f_i时， f_i 对 F 的最大贡献。

设 $F = \sum_{i=1}^{N} f_i$ 如定理 1中所述，即所有 f_i均为次模。设 $|\mathsf{EX}(\mathcal{B}(f_i))|$ 表示 f_i [16]的基多面体中的极点数，并设 $B = \max_{i \in [N]} |\mathsf{EX}(\mathcal{B}(f_i))|$（精确定义见附录B.1 ）。 [34] 的作者声称

$$\sum_{i=1}^{N} p_i \;\le\; Bn, \tag{5}$$

由此，根据定理 1，可知存在一个期望大小为 $\mathcal{O}(\frac{Bn^2}{\varepsilon^2})$ 的稀疏化器。正如我们稍后将看到的，只有当 f_i 单调时，这一结论才成立。利用椭球法对峰值贡献进行 $\mathcal{O}(\sqrt{n})$ 近似后，[3], 中进一步证明了：若所有 f_i 不仅是次模函数，而且是单调的，那么在假设（5）成立的情况下，可以在随机多项式时间内找到一个期望大小为 $\mathcal{O}(Bn^{2.5}\log_{\varepsilon^2} n)$ 的稀疏化器。~~这里，~~若对任意 $S \subseteq T \subseteq E$，函数 F: $2^E \to \mathbb{R}$ 满足 $F(S) \le F(T)$，则称其为单调函数。

[1]每个 f_i 都由一个预言机表示，该预言机针对任意 $S \subseteq E$返回值 $f_i(S)$。

贡献 我们的主要贡献是，为低曲率的可分解单调 k-次模函数提出了一种稀疏化算法。作为切入点，我们在第 2 节（并在附录A中给出证明）观察到， [34] 中用于证明定理 1 的采样算法在很大程度上独立于次模性，因此可以得到一种适用于更一般的可分解函数的稀疏化算法。在此过程中，我们建立了一个集中界，表明所得稀疏化器的大小超过期望大小的 3/2倍的可能性极低。具体而言，考虑一个有限定义域 $\mathcal{D}$；在集合函数的情形下，它就是幂集 $\mathcal{D} = 2^E$ 。再假设 $F\colon \mathcal{D} \to \mathbb{R}$ 可分解为 $F = \sum_{i=1}^N f_i$，其中$f_i\colon \mathcal{D} \to \mathbb{R}$ 对每个 $i \in [N]$ 均成立。[2]

定理 2 （定理 3的非正式版本）。设 $F = \sum_{i=1}^N f_i$，其中 $f_i\colon \mathcal{D} \to \mathbb{R}$。对于每个$\varepsilon$， $\delta \in (0, 1)$ 都存在一个向量 $w \in \mathbb{R}^N$使得

(i) $\mathbb{P}\left[w \text{ 是一个 } \varepsilon\text{稀疏化器}\right] \geq 1 - \delta$;

(ii) $\mathbb{E}\left[\mathsf{size}(w)\right] = \mathcal{O}\left(\frac{\log|\mathcal{D}|+\log\frac{1}{\delta}}{\varepsilon^2}\sum_{i=1}^N p_i\right)$, *where* $p_i = \max\limits_{\substack{A\in\mathcal{D}\\F(A)\neq 0}} \frac{f_i(A)}{F(A)}$;

(iii) $\mathbb{P}\left[\text{大小}(w) \leq \frac{3}{2}\mathbb{E}\left[\text{大小}(w)\right]\right] \geq 1 - 4\varepsilon^2$。

作为我们的主要贡献，我们利用定理 2 在第 3 节为可分解的低曲率单调 k-次模函数给出一个稀疏化器。作为次要贡献，我们澄清了 [34]中的一些次模函数稀疏化结果。首先，我们通过给出一个反例，指出 5) 中声称的不等式（ [34] 不正确。不过，我们证明，在附加单调性假设下，不等式（5）成立。相关证明见附录 B。其次，在附录 C 中，我们为一类有界元数的可分解单调次模函数给出一个稀疏化器。

对于一个自然数 $k \geq 1$，定义在 k 上由两两不相交子集构成的 k-元组上的函数 $F\colon (k+1)^E \to \mathbb{R}$ ，若 f 满足类似于不等式 (1) 中次模不等式的关系，则称其为 k-次模函数。具体而言，设 $\mathbf{A} = (A_1, \ldots, A_k) \in (k+1)^E$ 是一个由 E 的两两不相交子集构成的 k-元组，$\mathbf{B} = (B_1, \ldots, B_k) \in (k+1)^E$ 也类似。那么，若 $F\colon (k+1)^E \to \mathbb{R}$ 满足

$$f(\mathbf{A} \sqcap \mathbf{B}) + f(\mathbf{A} \sqcup \mathbf{B}) \leq f(\mathbf{A}) + f(\mathbf{B}), \tag{6}$$

其中

$$\mathbf{A} \sqcap \mathbf{B} = (A_1 \cap B_1, \ldots, A_k \cap B_k), \tag{7}$$

and

$$\mathbf{A} \sqcup \mathbf{B} = \Big((A_1 \cup B_1) \setminus \bigcup_{i\in\{2,\ldots,k\}} (A_i \cup B_i), \ldots, (A_k \cup B_k) \setminus \bigcup_{i\in\{1,\ldots,k-1\}} (A_i \cup B_i)\Big). \tag{8}$$

根据此定义，1-次模性与不等式（1）中定义的集合函数次模性的标准概念完全对应；同样，2-次模性与双次模性 [8, 9]完全对应。 k-次模函数这一类函数由[19] 引入，并在所谓的有限值约束满足问题 [20, 26]研究中发挥了重要作用。

[2]E每个 f_i 都由一个求值预言机表示，该预言机需要 $\mathcal{O}(\mathsf{EO}_i)$ 时间来返回任意 $S \in \mathcal{D}$ 的 $f_i(S)$.

通过求值预言机给出的 1-次模函数 [37, 22] 和 2-次模函数 [17] 可以高效地最小化[21], , 但 k-次模函数的 $k \geq 3$最小化问题的复杂性仍未解决。另一方面, k-次模函数 [47, 23, 32], 最大化问题的可近似性已得到充分理解, [35], 包括基数约束下的单调情形 [13], 、流式模型 [44, 43, 33]以及其他变体。

次模函数的单调性定义可以自然地扩展到 k-次模函数: $F\colon (k+1)^E \to \mathbb{R}$ 称为单调函数, 如果对于所有 $\mathbf{A} = (A_1, \ldots, A_k) \in (k+1)^E$ 和 $\mathbf{B} = (B_1, \ldots, B_k) \in (k+1)^E$, 且每个 $i \in [k]$都满足 $A_i \subseteq B_i$ 时, $F(\mathbf{A}) \leq F(\mathbf{B})$。

在次模函数的研究中, 一个重要概念是有界曲率 [12]。对于单调次模函数 $F\colon 2^E \to \mathbb{R}_{\geq 0}$, 其曲率（在 [46]中也称为总曲率） c_F 定义如下:

$$c_F = 1 - \min_{S \subseteq E, e \in E \setminus S} \frac{\Delta_e F(S)}{\Delta_e F(\emptyset)}, \tag{9}$$

其中, $\Delta_e f(S)$表示 e 相对于 S的边际增益, 即:

$$\Delta_e F(S) = F(S \cup \{e\}) - F(S). \tag{10}$$

换言之, 曲率比较的是向任意集合和空集添加基集元素时的边际增益。注意, $c_F \in [0, 1]$, 其上界由不等式 (2) 得出。此外, 当且仅当 f 是模函数时, $c_F = 0$ 成立; 也就是说, 不等式 (1) （等价地, 不等式 (2) ）取等号时, $c_F = 0$ 成立。若 f 满足 $c_F < 1$, 则称其具有低曲率。直观而言, 曲率 c_F 表示“函数弯曲的程度”。曲率的概念已从次模函数扩展到 k-次模函数, 见 [36], , 另见 [28]。为定义曲率, 我们首先需要介绍 k-次模函数的边际值概念, 这是 $k = 1$ 情形的自然推广。设 $F\colon (k+1)^E \to \mathbb{R}$ 是一个k-次模函数。对于 $\mathbf{A} = (A_1, \ldots, A_k)$、 $i \in [k]$和 $e \in E \setminus \cup_{j \in [k]} A_j$, 我们定义 e相对于 $\mathbf{A}$ 和 i 的边际增益如下:

$$\Delta_{e,i} F(\mathbf{A}) = F(A_1, \ldots, A_{i-1}, A_i \cup \{e\}, A_{i+1}, \ldots, A_k) - F(\mathbf{A}). \tag{11}$$

那么, c_F F 的曲率定义为

$$c_F = 1 - \min_{i \in [k], e \in E, \mathbf{A} \in (k+1)^{E \setminus \{e\}}} \frac{\Delta_{e,i} F(\mathbf{A})}{\Delta_{e,i} F(\emptyset)}. \tag{12}$$

和之前一样, 如果 $c_F < 1$, 我们就说 f 具有低曲率。

作为我们的主要贡献, 我们将证明, 在单调性和低曲率的假设下, 可以高效地近似峰值贡献, 从而高效执行第 2 节中的采样算法。除了在技术上颇具挑战性之外, 我们还将这项工作视为对稀疏化领域的一项概念性贡献, 因为我们探索了比以往工作更一般的设定。

2 TheCore 算法

本节将介绍我们所有稀疏化算法的核心——Rafiey 和 Yoshida [34] 最初为可分解次模

函数所描述的一种随机采样程序。正如第 1 节所述，我们发现它在很大程度上独立于次模性，因此可以得到一种适用于可分解函数的更一般的稀疏化算法。本文所介绍的内容基本沿袭了 [34] 中的第 3 节，详细内容见附录A。

本文提出的算法以概率方式为任意可分解函数$F = \sum_{i=1}^{N} f_i$ 构造一个 ε-稀疏化器：$\mathcal{D} \to \mathbb{R}$ 。与 [34]，一样，对于每个 $i \in [N]$，它按照与比值成比例的概率对函数进行采样，

$$p_i = \max_{\substack{A \in \mathcal{D} \\ F(A) \neq 0}} \frac{f_i(A)}{F(A)}. \tag{13}$$

具体过程详见算法 1。

算法 1 核心稀疏化算法

要求： 函数 $F = f_1 + \cdots + f_N$，其 f_i： $\mathcal{D} \to \mathbb{R}$ 由求值预言机给出；误差容忍参数 ε、$\delta \in$ （0, 1） **确保：** 向量 $w \in \mathbb{R}^N$ 满足

- $\mathbb{P}[w \text{ is an } \varepsilon\text{-sparsifier}] \geq 1 - \delta$;
- $\mathbb{E}[\mathsf{size}(w)] = \mathcal{O}\left(\frac{\log|\mathcal{D}|+\log\frac{1}{\delta}}{\varepsilon^2} \sum_{i=1}^{N} p_i\right)$, where $p_i = \max_{\substack{A \in \mathcal{D} \\ F(A) \neq 0}} \frac{f_i(A)}{F(A)}$;
- $\mathbb{P}\left[\mathsf{size}(w) \leq \frac{3}{2}\mathbb{E}[\mathsf{size}(w)]\right] \geq 1 - 4\varepsilon^2$.

$w \leftarrow (0, \ldots, 0)$
$\kappa \leftarrow 3 \log\left(\frac{2|\mathcal{D}|}{\delta}\right)/\varepsilon^2$
对于 $i = 1, \ldots, N$ **do**
  $p_i \leftarrow$最大值 $_{\substack{A \in \mathcal{D} \\ F(A) \neq 0}} \frac{f_i(A)}{F(A)}$ ▷计算峰值贡献（此处：采用朴素方法）
  $\kappa_i \leftarrow$ 最小值$\{1, \kappa p_i\}$ ▷ 由于 κ_i 是概率，因此上限设为 1
  $w_i \leftarrow \begin{cases} 1/\kappa_i & \text{以概率 } \kappa_i \\ 0 & \text{以概率 } 1 - \kappa_i \end{cases}$ ▷对 f_i进行加权抽样
结束循环
返回 w

定理 3. 算法 *1* 输出一个向量 $w \in \mathbb{R}^N$ 使得

(i) $\mathbb{P}[w \text{ is an } \varepsilon\text{-sparsifier}] \geq 1 - \delta$;

(ii) $\mathbb{E}[\mathsf{size}(w)] = \mathcal{O}\left(\frac{\log|\mathcal{D}|+\log\frac{1}{\delta}}{\varepsilon^2} \sum_{i=1}^{N} p_i\right)$, *where* $p_i = \max_{\substack{A \in \mathcal{D} \\ F(A) \neq 0}} \frac{f_i(A)}{F(A)}$;

(iii) $\mathbb{P}\left[\mathsf{size}(w) \leq \frac{3}{2}\mathbb{E}[\mathsf{size}(w)]\right] \geq 1 - 4\varepsilon^2$.

备注 4. 算法1 可通过 $\delta = \mathcal{O}(1/n^c)$ 调用，从而以高概率得到一个 ε-稀疏化器。这只会使运行时间增加一个常数因子 c，因为其依赖于 $\log\frac{1}{\delta}$。

备注 5. 如果主要关注的是稀疏化器的大小，那么多次运行算法 1 并取最小的向量 w （就 $\mathsf{size}(w)$ 而言），即可得到一种过程：对于任意固定的 $\varepsilon > 0$，经过对数次迭代后，返回大小为 $\mathcal{O}\left(^{\log|\mathcal{D}|+}{}_{\log_{\varepsilon^2}} \ ^{\frac{1}{\delta}} \sum_{i=1}^{N} p_i\right)$ 的稀疏化器 ——————

。这是定理 3(iii) 的推论。请注意，可能需要适当选择 δ，以确保该解确实以高概率成为一个 ε-稀疏化器。

推论 6。 在算法 *1*的设置下，设 $\widehat{p}_1$、……、$\widehat{p}_N \in \mathbb{R}_{\neq 0}$ 满足 $\widehat{p}_i \geq p_i$ 对所有 $i \in [N]$均成立。如果算法 *1* 在第 *4* 行中使用 $\widehat{p}_i$而不是 $p_i = \max_{A \in \mathcal{D}} \frac{f_i(A)}{F(A)}$，则返回向量 $w \in \mathbb{R}^N$，使得

(i) $\mathbb{P}[w \text{ 是一个 } \varepsilon\text{稀疏化器}] \geq 1 - \delta$;
(ii) $\mathbb{E}[\text{大小}(w)] = \mathcal{O}\left(\frac{\log |\mathcal{D}| + \text{对数}}{\varepsilon^2} \cdot \frac{1}{\delta} \sum_{i=1}^N \widehat{p}_i\right)$;
(iii) $\mathbb{P}\left[\text{大小}(w) \leq \frac{3}{2}\mathbb{E}[\text{大小}(w)]\right] \geq 1 - 4\varepsilon^2$。

注意，推论 6 表明，对 p_i进行任何常数因子近似都能完成任务，并得到相同的渐近界。然而，对于一般的函数f_1、……、f_N：$\mathcal{D} \to \mathbb{R}$，不可能以远快于求精确 p_i的方式获得这些结果。即使所有 f_i都是次模函数，p_i也很难进行近似。

注 7。 一般而言，我们所知的峰值贡献最佳上界为 $p_i \leq 1$。因此，推论 6 告诉我们，对所有 $\widehat{p}_i = 1$，均可正确地调用算法 1，并令 $i \in [N]$。由于 $\kappa > 1$ 当 $\varepsilon \in (0, 1)$ 时，我们便有 $\kappa_i = \min\{1, \kappa p_i\} = 1$。这就得到了初始分解，也就是说，算法 1 实际上什么也没做——稀疏化器可不是免费的!

注 8。 算法 1有多种实现方式，因此会得到不同的运行时间界。这里涉及的函数如何表示起着关键作用。在最一般的情形下，若不作进一步假设，合理的做法是假设每个 f_i都由一个求值预言机表示，其响应时间为 $\mathcal{O}(\mathsf{EO}_i)$。我们还可以假设有一个额外的预言机用于计算 F，其响应时间为 $\mathcal{O}(\mathsf{EO}_\Sigma)$——事实上，在许多应用中，$F(S)$ 的计算速度远快于将所有 $f_i(S)$ 对 $i \in [N]$求和。

算法1 成功构造一个（小型）稀疏化器所需完成的主要工作，是计算或近似峰值贡献。在最一般的情形下，我们必须通过遍历所有 p_i 来计算 $A \in \mathcal{D}$，这至少需要 $\Omega(|\mathcal{D}|)$ 的时间。因此，朴素实现的运行时间为 $\mathcal{O}\left(N|\mathcal{D}| \sum_{i=1}^N \mathsf{EO}_i\right)$。我们的工作的主要贡献在于，展示如何针对一些有趣的情形高效近似 p_i，其中最值得关注的是低曲率的单调 k-次模函数。

3 低曲率单调 k-次模函数

设 $F\colon (k+1)^E \to \mathbb{R}_{\geq 0}$ 可分解为 $F = \sum_{i=1}^N f_i$，其中每个 $f_i\colon (k+1)^E \to \mathbb{R}_{\geq 0}$都是非负的低曲率单调 k-次模函数。根据定义可知，F 同样是非负、单调的 k-次模函数，且具有低曲率。

我们将展示如何近似峰值贡献 $p_i = \max_{\mathbf{A} \in (k+1)^E} \frac{f_i(\mathbf{A})}{F(\mathbf{A})}$。

为此，只需近似两个低曲率单调 k-次模函数 f, g 的~~比~~ $\max_{\mathbf{A} \in (k+1)^E} \frac{f}{g}(\mathbf{A})$：$(k+1)^E \to \mathbb{R}_{\geq 0}$。给定 $\mathsf{A} = (A_1, \ldots, A_k) \in (k+1)^E$，我们

定义

$$S_f(\mathbf{A}) := \sum_{i=1}^{k} \sum_{e \in A_i} \Delta_{e,i}\left(f \mid \varnothing, \ldots, \varnothing\right) \tag{14}$$

and

$$S_g(\mathbf{A}) := \sum_{i=1}^{k} \sum_{e \in A_i} \Delta_{e,i}\left(g \mid \varnothing, \ldots, \varnothing\right). \tag{15}$$

事实证明，$\frac{S_f(\mathbf{A})+f(\varnothing,\ldots,\varnothing)}{S_g(\mathbf{A})+g(\varnothing,\ldots,\varnothing)}$ 能够很好地近似 $\frac{f(\mathbf{A})}{g(\mathbf{A})}$。

引理 9. 设 $\mathbf{A} \in (k+1)^E$ 是 $1-\varepsilon$-近似最大化器，针对 $\frac{S_f(\mathbf{A})+f(\varnothing,\ldots,\varnothing)}{S_g(\mathbf{A})+g(\varnothing,\ldots,\varnothing)}$。那么

$$\frac{f(\mathbf{A})}{g(\mathbf{A})} \geq (1-\varepsilon)(1-c_f)(1-c_g)\frac{f(\mathbf{A}^*)}{g(\mathbf{A}^*)}$$

对于任意 $\mathbf{A}^* \in (k+1)^E$。

令 $\varepsilon = 1/2$ 代入引理 9 可得一个 $\frac{1}{2}(1-c_f)(1-c_g)$-近似；当 c_f 和 c_g 均视为常数时，这是一个常数因子近似。接下来需要说明如何将 $\max_{\mathbf{A}\in(k+1)^E} \frac{S_f(\mathbf{A})+f(\varnothing,\ldots,\varnothing)}{S_g(\mathbf{A})+g(\varnothing,\ldots,\varnothing)}$ 近似到因子 $1-\varepsilon$ 以内（在我们的应用场景中，具体为 $1/2$）。具体做法是将问题归约到下文定义的 MODULAR-RATIO-MAX 问题，[3] 并针对该问题设计一个全多项式近似方案（全多项式时间近似方案）。

> MODULAR-RATIO-MAX 给定：$x_1, \ldots, x_n, y_1, \ldots, y_n \in \mathbb{R}_{>0}$ 和 A，$B \in \mathbb{R}_{\geq 0}$。目标：找出指标集 $\varnothing \neq I \subseteq [n]$，使 $\varrho(I) := \frac{A+\sum_{i\in I} x_i}{B+\sum_{i\in I} y_i}$ 最大。

现在很容易描述如何归约到 MODULAR-RATIO-MAX。回顾一下

$$\frac{S_f(\mathbf{A}) + f(\varnothing,\ldots,\varnothing)}{S_g(\mathbf{A}) + g(\varnothing,\ldots,\varnothing)} = \frac{A + \sum_{i=1}^{k}\sum_{e\in A_i}\Delta_{e,i}\left(f \mid \varnothing,\ldots,\varnothing\right)}{B + \sum_{i=1}^{k}\sum_{e\in A_i}\Delta_{e,i}\left(g \mid \varnothing,\ldots,\varnothing\right)} \tag{16}$$

其中 $A := f(\varnothing,\ldots,\varnothing)$ 且 $B := g(\varnothing,\ldots,\varnothing)$。如果我们按任意顺序将这些对 $(e,i) \in E \times [k]$ 编号为 $(e_1,i_1), \ldots, (e_{nk},i_{nk})$，就会发现，最大化 $\frac{S_f(\mathbf{A})+f(\varnothing,\ldots,\varnothing)}{S_g(\mathbf{A})+g(\varnothing,\ldots,\varnothing)}$ 等价于最大化

$$\frac{A + \sum_{\ell\in I}\Delta_{e_\ell,i_\ell}\left(f \mid \varnothing,\ldots,\varnothing\right)}{B + \sum_{\ell\in I}\Delta_{e_\ell,i_\ell}\left(g \mid \varnothing,\ldots,\varnothing\right)} \tag{17}$$

对所有指标集 $\varnothing \neq I \subseteq [nk]$（$I = \varnothing$ 的情形可以手动检验，也可以忽略，因为它对应于 $\mathbf{A} = (\varnothing,\ldots,\varnothing)$）。由于 f 和 g 是单调的，边际增益始终非负，因此对所有 $\ell \in [nk]$，$\Delta_{e_\ell,i_\ell}(f \mid \varnothing,\ldots,\varnothing) \geq 0$ 以及 $\Delta_{e_\ell,i_\ell}(g \mid \varnothing,\ldots,\varnothing) \geq 0$。为满足 MODULAR-RATIO-MAX 定义中要求的严格正性，我们可以删去等于0的边际增益。这只会缩小问题规模。

总而言之，计算 $p_i = \max_{\mathbf{A}\in(k+1)^E} \frac{f_i}{F}(\mathbf{A})$（$\mathbf{A}$）的算法可概述如下：

[3] 请注意，我们允许 A 和 B 等于零，而 x_i 和 y_i 则严格大于零。这是一项略显技术性的要求，旨在避免除以零。

• 通过1/2-近似算法 $\mathbf{A}^*$ 来求解 $\max_{\mathbf{A}\in(k+1)^E} S_f(\mathbf{A}) + f(\varnothing,\ldots,\varnothing) S_g(\mathbf{A}) + g(\varnothing,\ldots,\varnothing)$，采用 MODULAR-RATIO-MAX归约和算法 2。

• 令 $\widehat{p}_i := \frac{2(1-c_f)(1-c_g)}{f(\mathbf{A}^*)\, g(\mathbf{A}^*)}$。

根据引理 $\widehat{p}_i \geq p_i$ 可保证9。此外，$\widehat{p}_i = \mathcal{O}(1)\cdot p_i$，因此，通过应用核心算法，我们可以得到一个期望大小与存在性结果相符的稀疏化器。此外，算法在多项式时间内运行，正如我们将在引理 10中看到的。

3.1 全多项式时间近似方案

如果 $A = B = 0$, MODULAR-RATIO-MAX 有一个非常简单的解：我们只需取使 x_i/y_i最大化的索引 i 对应的 $I = \{i\}$。然而，正如下面的例子所示，这通常并非最优。令 $A = 1$, $B = 100$ 和 $x_1 = 2$, $y_1 = 3$, $x_2 = 1$, $y_2 = 1$。显然，比率x_i/y_i 在 $i = 2$时达到最大，从而总的 ϱ-值为

$$\varrho(\{2\}) = \frac{1+1}{100+1} = \frac{2}{101}$$

而非

$$\varrho(\{1,2\}) = \frac{1+2+1}{100+3+1} = \frac{4}{104},$$

显然更大。事实上，不难看出，x_i/y_i 的最大化解甚至无法提供常数因子近似；与最优解相比，它最终可能任意差。这表明我们需要采取其他方法。

该解是一种基于二分搜索的全多项式时间近似方案，如算法2所示。这是可行的，因为我们可以轻松求解相关的判定问题：给定目标值$\lambda \in \mathbb{R}$，是否存在一个指标集 $\varnothing \neq I \subseteq [n]$ 使得 $\varrho(I) \geq \lambda$?

通过代数等价变换，可以简化该判定问题：

$$\varrho(I) \geq \lambda \iff \frac{A + \sum_{i\in I} x_i}{B + \sum_{i\in I} y_i} \geq \lambda \iff A - B\lambda + \sum_{i\in I}(x_i - \lambda y_i) \geq 0$$

要判断最后一个表达式是否对任意非空指标集 I均非负，我们按数量 $x_i - \lambda y_i$从大到小排列各指标。必须先选取该顺序中的第一个指标（因为要求 $I \neq \varnothing$），然后选取所有满足 $x_i - \lambda y_i$ 为正的其余指标 i。这样便能使左端在所有 $I \neq \varnothing$中达到最大值。如果该最大值非负，根据上述等价关系，我们便知道 $\varrho(I) \geq \lambda$。

令 $m := \min\{x_1,\ldots,x_n,y_1,\ldots,y_n\}$ 和 $M := \max\{x_1,\ldots,x_n,y_1,\ldots,y_n\}$。对于任意非空指标集 I, 我们总有

$$\varrho(I) = \frac{A + \sum_{i\in I} x_i}{B + \sum_{i\in I} y_i} \geq \frac{A+m}{B+nM} \tag{18}$$

以及

$$\varrho(I) = \frac{A + \sum_{i\in I} x_i}{B + \sum_{i\in I} y_i} \leq \frac{A+nM}{B+m}, \tag{19}$$

因此，我们可以用 $\varrho^- = \frac{A+m}{B+nM}$ 和 $\varrho^+ = \frac{A+nM}{B+m}$ 初始化二分搜索。一旦区间 $[\varrho^-, \varrho^+]$ 的长度不超过 $\varepsilon\frac{A+m}{B+nM}$，我们就知道乘法误差至多为$\varepsilon$。由于区间长度在每一步都会减半，因此最多经过$\left\lceil \log \frac{1}{\varepsilon} + 2\left(\log n + \log \frac{M}{m}\right)\right\rceil$ 次迭代即可达到这一点，正如下述引理所示。

引理 10。 算法 2中的二分搜索会在 $k := \left\lceil \log\frac{1}{\varepsilon} + 2\left(\log n + \log \frac{M}{m}\right)\right\rceil$次迭代内终止。此外，最终集合 I 满足 $\varrho(I) \geq (1-\varepsilon)\varrho(I^*)$ 对任意 $\varnothing \neq I^* \subseteq [n]$成立。

证明。为确定迭代次数的上界，我们注意到 $|\varrho^+ - \varrho^-|$ 在每次迭代中都会缩小一半。因此，经过 k 次迭代后，有

$$\left|\varrho^+ - \varrho^-\right| \leq \frac{1}{2^k}\left|\frac{A+nM}{B+m} - \frac{A+m}{B+nM}\right| \leq \frac{1}{2^k}\frac{A+nM}{B+m}.$$

我们希望其为 $\leq \varepsilon\frac{A+m}{B+nM}$，这等价于

$$\frac{1}{2^k}\frac{A+nM}{B+m} \leq \varepsilon\frac{A+m}{B+nM} \iff 2^k \geq \frac{1}{\varepsilon}\frac{A+nM}{A+m}\frac{B+nM}{B+m}. \tag{20}$$

由于 $nM \geq m$，我们也有 $\frac{A+nM}{A+m} \leq \frac{nM}{m}$ 以及 $\frac{B+nM}{B+m} \leq \frac{nM}{m}$，因此

$$\frac{1}{\varepsilon}\frac{A+nM}{A+m}\frac{B+nM}{B+m} \leq \frac{1}{\varepsilon}\left(\frac{nM}{m}\right)^2,$$

因此，只需满足 $2^k \geq \frac{1}{\varepsilon}\left(\frac{nM}{m}\right)^2$，方程（20）便成立。对任意 $k \geq \log\frac{1}{\varepsilon} + 2\left(\log n + \log\frac{M}{m}\right)$，该条件确实成立，由此得到迭代次数的上界。

对于误差界，令 $\varnothing \neq I^* \subseteq [n]$ 为任意值。根据公式（18），我们知道$\varrho(I^*) \geq \frac{A+m}{B+nM}$。此外，二分搜索保持以下两个不变量：

(i) $\varrho^- \leq \varrho(I) \leq \varrho^+$ 对于算法 中的集合 2，以及

(ii) 不存在满足 $\varrho(I') > \varrho^+$ 的集合 $\varnothing \neq I' \subseteq [n]$ 。

结合这两点以及$|\varrho^+ - \varrho^-| \leq \varepsilon\frac{A+m}{B+nM}$ 在算法终止时这一事实，我们得出

$$\varrho(I) \geq \varrho^- = \varrho^+ - \left(\varrho^+ - \varrho^-\right) \geq \varrho^+ - \varepsilon\frac{A+m}{B+nM} \geq \varrho(I^*) - \varepsilon\varrho(I^*) = (1-\varepsilon)\varrho(I^*),$$

其中我们还利用了 $\varrho\ I' \geq \frac{A+m}{B+nM}$ 18 $\varrho\ I^* \leq$ ()（公式（））， 并使用不变量 (ii) 作为（）。 ϱ^+ □

我们指出，如果所有输入数字都是整数，则迭代次数上界为

$$k \leq \left\lceil \log\frac{1}{\varepsilon} + 2\left(\log n + \log U\right)\right\rceil = \mathcal{O}\left(\log\frac{1}{\varepsilon} + \log n + \log U\right),$$

其中， U 是输入中出现的最大数。至此，我们已经求解了 MODULAR-RATIO-MAX。

算法 2： MODULAR-RATIO-MAX的全多项式时间近似方案

要求： MODULAR-RATIO-MAX 实例 $x_1, \ldots, x_n, y_1, \ldots, y_n \in \mathbb{R}_{>0}$ 和 $A, B \in \mathbb{R}_{\geq 0}$；误差容差 $\varepsilon > 0$。

确保：指标集 $\varnothing \neq I \subseteq [n]$，使得对于所有 $\varnothing \neq I^* \subseteq [n]$，$\varrho(I) \geq (1-\varepsilon)\varrho(I^*)$。

过程 CHECK(λ)
  对指标排序，使得 $x_{i_1} - \lambda y_{i_1} \geq x_{i_2} - \lambda y_{i_2} \geq \cdots \geq x_{i_n} - \lambda y_{i_n}$
  $I \leftarrow \{i_1\}$
  $S \leftarrow (A - \lambda B) + (x_{i_1} - \lambda y_{i_1})$
  对于 $\ell = 2$ 到 n，执行
    如果 $x_{i_\ell} - \lambda y_{i_\ell} > 0$，则
      $I \leftarrow I \cup \{i_\ell\}$
      $S \leftarrow S + (x_{i_\ell} - \lambda y_{i_\ell})$
    结束如果
  结束循环
  返回 $\begin{cases} I & \text{如果 } S \geq 0 \\ \perp & \text{否则} \end{cases}$
结束过程
$I \leftarrow \{1\}$ ▷ 以任意可行解初始化
$m \leftarrow \min\{x_1, \ldots, x_n, y_1, \ldots, y_n\}$
$M \leftarrow$ 最大值 $\{x_1, \ldots, x_n, y_1, \ldots, y_n\}$
$\bar{\varrho} \leftarrow \frac{A+m}{B+nM}$
$\varrho^+ \leftarrow \frac{A+nM}{B+m}$
当 $|\varrho^+ - \varrho^-| \geq \varepsilon \frac{A+m}{B+nM}$ **do** ▷ 二分搜索，直到乘法误差为 $\leq \varepsilon$
  $\lambda \leftarrow \frac{1}{2}(\varrho^- + \varrho^+)$
  $I_\lambda \leftarrow$ CHECK(λ)
  如果 $I_\lambda = \perp$，则
    $\varrho^+ \leftarrow \lambda$
  else
    $I \leftarrow I_\lambda$
    $\varrho^- \leftarrow \lambda$
  结束如果
结束循环
返回 I

3.2 引理9的证明

在本节余下部分，我们将证明引理9。但首先，我们需要一个有用的引理 a.

Lemma 11. 设 $f, g : (k+1)^E \to \mathbb{R}_{\geq 0}$ 是低曲率单调 k-次模函数。那么

$$\frac{(1-c_f)S_f(\mathbf{A}) + f(\varnothing,\ldots,\varnothing)}{S_g(\mathbf{A}) + g(\varnothing,\ldots,\varnothing)} \leq \frac{f(\mathbf{A})}{g(\mathbf{A})} \leq \frac{S_f(\mathbf{A}) + f(\varnothing,\ldots,\varnothing)}{(1-c_g)S_g(\mathbf{A}) + g(\varnothing,\ldots,\varnothing)}$$

对所有 $\mathbf{A} \in (k+1)^E$ 均成立。

证明。固定 $\mathbf{A} \in (k+1)^E$，并以如下方式标记 E 中的元素：对于每个 $1 \leq i \leq k$，$A_i = \left\{e_1^{(i)}, \ldots, e_{a_i}^{(i)}\right\}$，其中 $a_i = |A_i|$。现在，

$$f(\mathbf{A}) - f(\varnothing,\ldots,\varnothing) = \sum_{i=1}^{k}\sum_{j=1}^{a_i} \Delta_{e_j^{(i)},i}\left(f \mid A_1,\ldots,A_{i-1},\{e_1^{(i)},\ldots,e_{j-1}^{(i)}\},\varnothing,\ldots,\varnothing\right).$$

$\Delta_{e_j^{(i)},i}\left(f \mid A_1,\ldots,A_{i-1},\{e_1^{(i)},\ldots,e_{j-1}^{(i)}\},\varnothing,\ldots,\varnothing\right)$ 项可以从两个方向进行估计。根据边际收益递减性质，我们知道

$$\Delta_{e_j^{(i)},i}\left(f \mid A_1,\ldots,A_{i-1},\{e_1^{(i)},\ldots,e_{j-1}^{(i)}\},\varnothing,\ldots,\varnothing\right) \leq \Delta_{e_j^{(i)},i}(f \mid \varnothing,\ldots,\varnothing),$$

而我们得出结论

$$\Delta_{e_j^{(i)},i}\left(f \mid A_1,\ldots,A_{i-1},\{e_1^{(i)},\ldots,e_{j-1}^{(i)}\},\varnothing,\ldots,\varnothing\right) \geq (1-c_f)\Delta_{e_j^{(i)},i}(f \mid \varnothing,\ldots,\varnothing)$$

来自 f 的曲率。由于

$$S_f(\mathbf{A}) := \sum_{i=1}^{k}\sum_{j=1}^{a_i} \Delta_{e_j^{(i)},i}(f \mid \varnothing,\ldots,\varnothing),$$

我们看到，在合并两个不等式后，$(1-c_f)S_f(\mathbf{A}) \leq f(\mathbf{A}) - f(\varnothing,\ldots,\varnothing) \leq S_f(\mathbf{A})$。类似地，我们可以推出 g 的 $(1-c_g)S_g(\mathbf{A}) \leq f(\mathbf{A}) - f(\varnothing,\ldots,\varnothing) \leq S_g(\mathbf{A})$。接下来，

$$\frac{f(\mathbf{A})}{g(\mathbf{A})} = \frac{f(\mathbf{A}) - f(\varnothing,\ldots,\varnothing) + f(\varnothing,\ldots,\varnothing)}{g(\mathbf{A}) - g(\varnothing,\ldots,\varnothing) + g(\varnothing,\ldots,\varnothing)} \leq \frac{S_f(\mathbf{A}) + f(\varnothing,\ldots,\varnothing)}{(1-c_g)S_g(\mathbf{A}) + g(\varnothing,\ldots,\varnothing)}$$

and

$$\frac{f(\mathbf{A})}{g(\mathbf{A})} = \frac{f(\mathbf{A}) - f(\varnothing,\ldots,\varnothing) + f(\varnothing,\ldots,\varnothing)}{g(\mathbf{A}) - g(\varnothing,\ldots,\varnothing) + g(\varnothing,\ldots,\varnothing)} \geq \frac{(1-c_f)S_f(\mathbf{A}) + f(\varnothing,\ldots,\varnothing)}{S_g(\mathbf{A}) + g(\varnothing,\ldots,\varnothing)},$$

将 $\frac{f(\mathbf{A})}{g(\mathbf{A})}$ 夹在 $\frac{(1-c_f)S_f(\mathbf{A})+f(\varnothing,\ldots,\varnothing)}{S_g(\mathbf{A})+g(\varnothing,\ldots,\varnothing)}$ 与 $\frac{S_f(\mathbf{A})+f(\varnothing,\ldots,\varnothing)}{(1-c_g)S_g(\mathbf{A})+g(\varnothing,\ldots,\varnothing)}$。 □

引理（引理 9 重述）。设 $\mathbf{A} \in (k+1)^E$ 是 $\frac{S_f(\mathbf{A})+f(\varnothing,\ldots,\varnothing)}{S_g(\mathbf{A})+g(\varnothing,\ldots,\varnothing)}$ 的 $(1-\varepsilon)$-近似最大化器。则

$$\frac{f(\mathbf{A})}{g(\mathbf{A})} \geq (1-\varepsilon)(1-c_f)(1-c_g)\frac{f(\mathbf{A}^*)}{g(\mathbf{A}^*)}$$

对于任意 $\mathbf{A}^* \in (k+1)^E$。

证明。 注意到 $f_{(\mathbf{A})}\ g(\mathbf{A})$ 以及 $f_{(\mathbf{A}^*)}\ g(\mathbf{A}^*)$ 均包含在引理 11所述的范围内，即

$$\frac{(1-c_f)S_f(\mathbf{A}^*)+f(\varnothing,\ldots,\varnothing)}{S_g(\mathbf{A}^*)+g(\varnothing,\ldots,\varnothing)} \leq \frac{f(\mathbf{A}^*)}{g(\mathbf{A}^*)} \leq \frac{S_f(\mathbf{A}^*)+f(\varnothing,\ldots,\varnothing)}{(1-c_g)S_g(\mathbf{A}^*)+g(\varnothing,\ldots,\varnothing)}$$

and

$$\frac{(1-c_f)S_f(\mathbf{A})+f(\varnothing,\ldots,\varnothing)}{S_g(\mathbf{A})+g(\varnothing,\ldots,\varnothing)} \leq \frac{f(\mathbf{A})}{g(\mathbf{A})} \leq \frac{S_f(\mathbf{A})+f(\varnothing,\ldots,\varnothing)}{(1-c_g)S_g(\mathbf{A})+g(\varnothing,\ldots,\varnothing)}.$$

结合 A 是一个 $(1-\varepsilon)$-近似最大化器，且 f 和 g均非负这一事实，我们得出

$$\begin{aligned}
\frac{f(\mathbf{A})}{g(\mathbf{A})} &\geq \frac{(1-c_f)S_f(\mathbf{A})+f(\varnothing,\ldots,\varnothing)}{S_g(\mathbf{A})+g(\varnothing,\ldots,\varnothing)} \\
&\geq (1-c_f)\frac{S_f(\mathbf{A})+f(\varnothing,\ldots,\varnothing)}{S_g(\mathbf{A})+g(\varnothing,\ldots,\varnothing)} \\
&\geq (1-c_f)(1-\varepsilon)\frac{S_f(\mathbf{A}^*)+f(\varnothing,\ldots,\varnothing)}{S_g(\mathbf{A}^*)+g(\varnothing,\ldots,\varnothing)} \\
&\geq (1-\varepsilon)(1-c_f)(1-c_g)\frac{S_f(\mathbf{A}^*)+f(\varnothing,\ldots,\varnothing)}{(1-c_g)S_g(\mathbf{A}^*)+g(\varnothing,\ldots,\varnothing)} \\
&\geq (1-\varepsilon)(1-c_f)(1-c_g)\frac{f(\mathbf{A}^*)}{g(\mathbf{A}^*)}.
\end{aligned}$$

□

参考文献

[1] Kook Jin Ahn 和 Sudipto Guha。半流式模型中的图稀疏化。载于 第 36 届自动机、语言与编程国际研讨会论文集（ICALP'09），第二部分，第 《计算机科学讲义》第 5556 卷，第 328–338 页。Springer，2009 年。doi:10.1007/978-3-642-02930-1_27。

[2] Alexandr Andoni、Jiecao Chen、Robert Krauthgamer、Bo Qin、David P. Woodruff 和 Qin Zhang。《二次型草图化》。载于《2016年ACM理论计算机科学创新会议论文集》（ITCS'16），第311–319页，2016年。doi:10.1145/2840728.2840753。[3] Wenruo Bai、Rishabh Iyer、Kai Wei 和 Jeff Bilmes。优化次模函数比值的算法。载于第33届国际机器学习会议论文集（ICML'16），《机器学习研究论文集》第48卷，第2751–2759页。PMLR，2016年6月20–22日。网址：https://proceedings.mlr.press/v48/baib16.html。[4] Nikhil Bansal、Ola Svensson 和 Luca Trevisan。图与超图稀疏化的新概念与构造。第60届IEEE计算机科学基础年会论文集（FOCS'19），第910–928页，2019年。doi:10.1109/FOCS.2019.00059。[5] Joshua D. Batson、Daniel A. Spielman 和 Nikhil Srivastava。双重拉马努金稀疏化器。SIAM计算杂志，41(6):1704–1721，2012年。doi:10.1137/090772873。[6] Andr´as A. Bencz´ur 和 David R. Karger。以 $\tilde{O}(n^2)$ 时间近似求解 s-t 最小割。载于第28届ACM计算理论年会论文集（STOC'96），第47–55页。ACM，1996年。doi:10.1145/237814.237827。

[7] Andr´as Bencz´ur 和 David Karger。容量图中割与流的随机近似方案。SIAM计算杂志，44(2):290–319，2015年。doi:10.1137/070705970。[8]Andr´e Bouchet。贪心算法与对称拟阵。数学规划，38(2):147–159，1987年。doi:10.1007/BF02604639。[9] Ramaswamy Chandrasekaran 和 Santosh N. Kabadi。拟拟阵。离散数学，71(3):205–217，1988年。doi:10.1016/0012-365X(88)90101-X。[10] Yu Chen、Sanjeev Khanna 和 Ansh Nagda。近线性规模的超图割稀疏化器。载于第61届IEEE计算机科学基础年会论文集（FOCS'20），第61–72页，2020年。arXiv:2009.04992，doi:10.1109/FOCS46700.2020.00015。[11] Michael B. Cohen、Jonathan Kelner、John Peebles、Richard Peng、Anup B. Rao、Aaron Sidford 和 Adrian Vladu。马尔可夫链的近线性时间算法及有向图的新谱原语。载于第49届ACM SIGACT计算理论研讨会年会论文集（STOC'17），第410–419页。ACM，2017年。arXiv:1611.00755，doi:10.1145/3055399.3055463。[12] Michele Conforti 和 G´erard Cornu´ejols。次模集合函数、拟阵与贪心算法：紧确的最坏情况界及 Rado-Edmonds 定理的一些推广。离散应用数学，7(3):251–274，1984年。doi:10.1016/0166-218X(84)90003-9。[13] Alina Ene 和 Huy L. Nguyen。具有基数约束的单调 k-次模最大化流式算法。载于国际机器学习会议论文集（ICML'22），《机器学习研究论文集》第162卷，第5944–5967页。PMLR，2022年。网址：https://proceedings.mlr.press/v162/ene22a.html。[14]Arnold Filtser 和 Robert Krauthgamer。二变量赋值约束满足问题的稀疏化。SIAM离散数学杂志，31(2):1263–1276，2017年。doi:10.1137/15M1046186。[15] Satoru Fujishige。次模系统与基多面体。载于《次模函数与优化》，《离散数学年刊》第47卷，第17–108页。Elsevier，1991年。doi:10.1016/S0167-5060(08)70127-4。[16] Satoru Fujishige。《次模函数与优化》。Elsevier，2005年。[17] Satoru Fujishige 和 Satoru Iwata。双次模函数最小化。SIAM离散数学杂志，19(4):1065–1073，2005年。doi:10.1137/S0895480103426339。[18] Wai Shing Fung、Ramesh Hariharan、Nicholas J. A. Harvey 和 Debmalya Panigrahi。图稀疏化的通用框架。载于第43届ACM计算理论研讨会论文集（STOC'11），第71–80页。ACM，2011年。doi:10.1145/1993636.1993647。[19]Anna Huber 和 Vladimir Kolmogorov。迈向最小化 k-次模函数。载于第2届组合优化国际研讨会论文集（ISCO'12），《计算机科学讲义》第7422卷，第451–462页。Springer，2012年。doi:10.1007/978-3-642-32147-4_40。[20] Anna Huber、Andrei Krokhin 和 Robert Powell。斜双次模性与赋值约束满足问题。SIAM计算杂志，43(3):1064–1084，2014年5月。doi:10.1137/120893549。[21] Satoru Iwata。次模函数最小化。数学规划，112(1):45–64，2008年。doi:10.1007/s10107-006-0084-2。[22] Satoru Iwata、Lisa Fleischer 和 Satoru Fujishige。最小化次模函数的组合强多项式算法。ACM学报，48(4):761–777，2001年。doi:10.1145/502090.502096。

[23]Satoru Iwata、Shin-ichi Tanigawa 和 Yuichi Yoshida。用于 k-子模函数最大化的改进算法。载于第 27 届 ACM-SIAM离散算法研讨会（SODA'16）论文集，第 404–413 页。SIAM，2016。arXiv:1502.07406，doi:10.1137/1.9781611974331.ch30。[24] Michael Kapralov、Robert Krauthgamer、Jakab Tardos 和 Yuichi Yoshida。近线性规模的谱超图稀疏化器。载于第 62 届 IEEE 计算机科学基础年会（FOCS'21）论文集，第 1159–1170 页。IEEE，2021。doi:10.1109/FOCS52979.2021.00114。[25] Dmitry Kogan 和 Robert Krauthgamer。图与超图中的割的草图化。载于 2015 年理论计算机科学创新会议（ITCS'15）论文集，第 367–376 页，2015。arXiv:1409.2391，doi:10.1145/2688073.2688093。[26] Vladimir Kolmogorov、Johan Thapper 和 Stanislav ˇZivn ´y。线性规划对一般值约束满足问题的作用。SIAM计算杂志，44(1):1–36，2015。arXiv:1311.4219，doi:10.1137/130945648。[27] Yin Tat Lee 和 He Sun。近线性时间内构造线性规模的谱稀疏化器。SIAM J. Comput.，47(6):2315–2336，2018。doi:10.1137/16M1061850。[28] Tatsuya Matsuoka 和 Naoto Ohsaka。具有有界曲率的单调 k次模函数及非k次模函数的最大化。载于第 13 届亚洲机器学习会议论文集，《机器学习研究论文集》第 157 卷，第 1707–1722 页。PMLR，2021 年 11 月 17–19 日。网址：https://proceedings.mlr.press/v157/matsuoka21b.html。

[29] Rajeev Motwani 和 Prabhakar Raghavan。随机算法。剑桥大学出版社，1995年。doi:10.1017/CBO9780511814075。

[30] George L. Nemhauser 和 Laurence A. Wolsey。《整数与组合优化》。约翰·威立父子出版公司，1988年。

[31] Ilan Newman 和 Yuri Rabinovich。关于乘法λ近似及其某些几何应用。SIAM计算杂志，42(3):855–883，2013年。doi:10.1137/100801809。

[32]Hiroki Oshima。改进的 k-次模函数最大化随机算法。SIAM离散数学杂志，35(1):1–22，2021年。arXiv:1907.12942，doi:10.1137/19M1277692。

[33] Canh V. Pham、Quang C. Vu、Dung K. T. Ha、Tai T. Nguyen 和 Nguyen D. Le。在预算约束下最大化 k-次模函数：应用与流式算法。*J. Comb.Optim.*，44:723–751，2022年。doi:10.1007/s10878-022-00858-x。

[34] Akbar Rafiey 和 Yuichi Yoshida。可分解次模函数的稀疏化。AAAI人工智能会议论文集，36(9):10336–10344，2022年6月。arXiv:2201.07289，doi:10.1609/aaai.v36i9.21275。[35] Shinsaku Sakaue。带拟阵约束的单调 k-次模函数最大化。离散优化，23:105–113，2017年。arXiv:1607.07957，doi:10.1016/j.disopt.2017.01.003。[36]Richard Santiago 和 F. Bruce Shepherd。多变量子模优化。载于第36届国际机器学习会议论文集（ICML'19），第97卷，收录于《机器学习研究论文集》，第5599–5609页。PMLR，2019年。网址：http://proceedings.mlr.press/v97/santiago19a.html。[37]Alexander Schrijver。强多项式时间内最小化次模函数的组合算法。组合理论杂志B辑，80(2):346–355，2000年。doi:10.1006/jctb.2000.1989。[38] Alexander Schrijver 等。组合优化：多面体与效率，第24卷。Springer，2003年。

[39] Tasuko Soma 和 Yuichi Yoshida。超图谱稀疏化。载于第30届年度 ACM-SIAM离散算法研讨会论文集（SODA'19），第2570–2581页。SIAM，2019年。arXiv:1807.04974, doi:10.1137/1.9781611975482.159。[40] Daniel A. Spielman 和 Nikhil Srivastava。基于有效电阻的图稀疏化。SIAM计算杂志，40(6):1913–1926，2011年。doi:10.1137/080734029。[41] 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。载于第36届年度 ACM 理论计算机科学研讨会论文集（STOC'04），第81–90页。ACM，2004年。doi:10.1145/1007352.1007372。[42] Daniel A. Spielman 和 Shang-Hua Teng。图的谱稀疏化。SIAM计算杂志，40(4):981–1025，2011年。doi:10.1137/08074489X。[43] Zhongzheng Tang、Chenhao Wang 和 Hau Chan。单调 k-次模秘书问题：基数与背包约束。理论计算机科学，921:86–99，2022年。doi:10.1016/j.tcs.2022.04.003。[44] Zhongzheng Tang、Chenhao Wang 和 Hau Chan。带背包约束的单调 k- 次模函数最大化。运筹学快报，50(1):28–31，2022 年。arXiv:2105.15159，doi:10.1016/j.orl.2021.11.010。[45] Donald Topkis。超模性与互补性。普林斯顿大学出版社，1998年。[46] Jan Vondr´ak。次模性与曲率：最优算法。RIMS Kˆokyˆuroku Bessatsu，23:253–266，2010年。网址：http://hdl.handle.net/2433/177046。[47] Justin Ward 和 Stanislav ˇZivn´y。k- 次模函数最大化及其扩展。ACM 算法汇刊，12(4)，2016 年 8 月。arXiv:1409.1399，doi:10.1145/2850419。

定理3和推论6的证明

定理（定理3 重述）。算法1 输出一个向量 $w \in \mathbb{R}^N$ 使得

(i) $\mathbb{P}[w$ 是一个 ε-稀疏化器$] \geq 1-\delta$；(ii) $\mathbb{E}[$大小$(w)] = \mathcal{O}\left(\frac{\log|\mathcal{D}|+\text{对数}_{\varepsilon^2}}{\delta}\sum_{i=1}^{N} p_i\right)$，其中 $p_i = \max_{A\in\mathcal{D}:F(A)\neq 0} \frac{f_i(A)}{F(A)}$；

(iii) $\mathbb{P}\left[\text{大小}(w) \leq \frac{3}{2}\mathbb{E}[\text{大小}(w)]\right] \geq 1-4\varepsilon^2$。

证明。我们首先证明部分(i)，方法是证明每个 $A \in \mathcal{D}$ 满足

$$\mathbb{P}\left[(1-\varepsilon)F(A) \leq F'(A) \leq (1+\varepsilon)F(A)\right] \geq 1-\frac{\delta}{|\mathcal{D}|}. \tag{21}$$

对定义域中的所有 $|\mathcal{D}|$ 元素[4] 应用并集界，即可证明该结论。注意，

$$\mathbb{E}\left[F'(A)\right] = \mathbb{E}\left[\sum_{i=1}^{N} w_i f_i(A)\right] = \sum_{i=1}^{N}\mathbb{E}[w_i]\, f_i(A) = \sum_{i=1}^{N} f_i(A) = F(A) \tag{22}$$

[4]遍历所有元素 $A \in \mathcal{D}$ 使得 $F(A) \neq 0$ 准确地说。如果 $F(A) = 0$，我们知道 $f_i(A) = 0$ 对所有$1 \leq i \leq N$都成立，因此方程（22）和方程（23）显然成立。我们在分析中不考虑这样的集合 A，并建议读者在接下来的分析中考虑 $A \in \mathcal{D}$，其中 $F(A) > 0$ 成立。

由于 $\mathbb{E}[w_i] = \kappa_i \cdot 1/\kappa_i + (1-\kappa_i) \cdot 0 = 1$。因此，上述表示“成功”的不等式等价于 $(1-\varepsilon)\mathbb{E}[F'(A)] \le F'(A) \le (1+\varepsilon)\mathbb{E}[F'(A)]$，即 $F'(A)$处于其期望值的 $1 \pm \varepsilon$ 倍范围内。下一步是证明

$$\mathbb{P}\left[\left|F'(A) - \mathbb{E}\left[F'(A)\right]\right| \ge \varepsilon \cdot \mathbb{E}\left[F'(A)\right]\right] \le 2e^{-\varepsilon^2\kappa/3}. \tag{23}$$

这基本上是通过注意到这些项 $w_i f_i(A)$可以视为取值范围有界的独立随机变量 $[0, a]$，对于某个 $a \le F(A)/\kappa$。然而，我们不能在此简单照搬Rafiey和Yoshida的证明[34]，因为其中有一步不成立。他们声称，$\max_{1\le i\le N} w_i f_i(A) = \max_{1\le i\le N} f_i(A)\,\kappa p_i$，但当 $\kappa_i = 1$ 时（注意 $\kappa_i = \min\{1, \kappa p_i\}$），这未必成立。在这种情况下，$w_i = 1/\kappa_i = 1$，所以 $w_i f_i(A) = f_i(A)$，它大于 $f_i(A)\,\kappa p_i$，对于 $\kappa p_i > 1$而言都是如此。事实上，这表明他们的不等式方向反了。

幸运的是，通过分别考虑满足 $\kappa p_i > 1$的指标 i，可以避开这个问题。令 $I := \{i \in \{1,\dots,N\} \mid \kappa p_i > 1\}$且 $\bar{I} := \{1,\dots,N\} \setminus I$。此外，对所有 $1 \le i \le N$, $Z := \sum_{i\in I} Z_i$ 和 $\bar{Z} := \sum_{i\in\bar{I}} Z_i$，定义$Z_i := w_i f_i(A)$。我们有

$$F'(A) = \sum_{i=1}^{N} Z_i = \sum_{i\in I} Z_i + \sum_{i\in\bar{I}} Z_i = Z + \bar{Z}$$

其中，第一部分 $\sum_{i\in I} Z_i = \sum_{i\in I} f_i(A)$ 是确定性的，因为我们有 $w_i = 1$，对每个 $i \in I$ 以概率 $\kappa_i = \min\{1, \kappa p_i\} = 1$。因此，

$$\begin{aligned}
F'(A) - \mathbb{E}\left[F'(A)\right] &= \sum_{i=1}^{N} w_i f_i(A) - \sum_{i=1}^{N} f_i(A) \\
&= \sum_{i\in I} w_i f_i(A) + \sum_{i\in\bar{I}} w_i f_i(A) - \sum_{i\in I} f_i(A) - \sum_{i\in\bar{I}} f_i(A) \\
&= Z + \sum_{i\in\bar{I}} w_i f_i(A) - Z - \sum_{i\in\bar{I}} f_i(A) \\
&= \sum_{i\in\bar{I}} w_i f_i(A) - \sum_{i\in\bar{I}} f_i(A) \\
&= \sum_{i\in\bar{I}} Z_i - \sum_{i\in\bar{I}} \mathbb{E}[Z_i] \\
&= \bar{Z} - \mathbb{E}\left[\bar{Z}\right]
\end{aligned}$$

根据公式（22），对 $\{1,\dots,N\} = I \cup \bar{I}$ 进行划分、Z_i的定义、 Z以及 $\bar{Z}$，再利用期望的线性性。我们得到，公式（23）等价于

$$\mathbb{P}\left[\left|\bar{Z} - \mathbb{E}\left[\bar{Z}\right]\right| \ge \varepsilon \cdot \mathbb{E}\left[F'(A)\right]\right] \le 2e^{-\varepsilon^2\kappa/3}. \tag{24}$$

由于对所有 $1 \le i \le N$ 都有 $f_i(A) \ge 0$，注意到

$$\mathbb{E}\left[\bar{Z}\right] = \sum_{i\in\bar{I}} f_i(A) \le \sum_{i\in\bar{I}} f_i(A) + \sum_{i\in I} f_i(A) = \sum_{i=1}^{N} f_i(A) = F(A) = \mathbb{E}\left[F'(A)\right], \tag{25}$$

其中最后一步是根据公式（22）得出的。现在，我们通过应用 $\mathbb{P}\left[\left|\bar{Z} - \mathbb{E}\left[\bar{Z}\right]\right| \ge \varepsilon \cdot \mathbb{E}[F'(A)]\right]$ Chernoff-Hoeffding 界[29]，采用 [34]中的定理 2.2 版本。一个

检查前置条件可知， Z_i 是独立随机变量（因为 w_i 是独立抽样的）。令 $a := \max_{i\in\bar{I}} Z_i$，所有满足 $i \in I$ 的 Z_i 的取值范围均为 $[0, a]$。取 $\mu := \mathbb{E}[F'(A)] \geq \mathbb{E}[\bar{Z}]$，我们得出

$$\begin{aligned} &\mathbb{P}\left[\left|\bar{Z} - \mathbb{E}\left[\bar{Z}\right]\right| \geq \varepsilon\mu\right] \leq 2e^{-\frac{\varepsilon^2\mu}{3a}} \\ \Longleftrightarrow\ &\mathbb{P}\left[\left|\bar{Z} - \mathbb{E}\left[\bar{Z}\right]\right| \geq \varepsilon \cdot \mathbb{E}\left[F'(A)\right]\right] \leq 2e^{-\frac{\varepsilon^2 F(A)}{3a}}. \end{aligned} \tag{26}$$

这与公式（24）非常接近。实际上，我们只需给出 a的一个上界。为此，注意到

$$Z_i = w_i f_i(A) = \frac{f_i(A)}{\kappa p_i} \leq \frac{f_i(A)}{\kappa \max_{S\subseteq E} \frac{f_i(S)}{F(S)}} \leq \frac{f_i(A)}{\kappa \frac{f_i(A)}{F(A)}} = \frac{F(A)}{\kappa}$$

对所有 $i \in I$均成立。因此，对所有 $i \in I$ 取最大值可得 $a = \max_{i\in\bar{I}} Z_i \leq F(A)/\kappa$。将其代入公式（26）可得公式（24）。当按照上文算法1 选择$\kappa = 3\log\left(\frac{2|\mathcal{D}|}{\delta}\right)/\varepsilon^2$ 时，右端 $2e^{-\varepsilon^2\kappa/3}$ 变为$\leq \frac{\delta}{|\mathcal{D}|}$，从而证明了公式（21）。

P部分（ii）是一个简单的计算。事件 $\{w_i \neq 0\}$ 发生的概率为 κ_i，因此

$$\mathbb{E}[\mathsf{size}(w)] = \mathbb{E}\left[\sum_{i=1}^{N}[w_i \neq 0]\right] = \sum_{i=1}^{N}\mathbb{P}[w_i \neq 0] = \sum_{i=1}^{N}\kappa_i \leq \sum_{i=1}^{N}\kappa p_i = \kappa\sum_{i=1}^{N}p_i,$$

其中 $\kappa = 3\log\left(\frac{2|\mathcal{D}|}{\delta}\right)/\varepsilon^2 = \mathcal{O}\left(\log|\mathcal{D}| + \log\frac{1}{\delta}\right)$，这是根据算法 1 的选择得出的。因此，总体上界为 $\mathbb{E}[\mathsf{size}(w)] = \mathcal{O}\left(\log|\mathcal{D}| + \log\frac{1}{\delta}\right)\sum_{i=1}^{N}p_i$。

第 (iii) 部分如下所示。令 $[w_i \neq 0]$ 为事件$\{w_i \neq 0\}$的指示随机变量，即当 $[w_i \neq 0] = 1$ 时 $w_i \neq 0$，当 $[w_i \neq 0] = 0$ 时 $w_i = 0$。独立随机变量之和的方差等于其方差之和。因此，

$$\mathbb{V}\mathrm{ar}[\mathsf{size}(w)] = \mathbb{V}\mathrm{ar}\left[\sum_{i=1}^{N}[w_i \neq 0]\right] = \sum_{i=1}^{N}\mathbb{V}\mathrm{ar}[[w_i \neq 0]]$$

其中，对于每个 $1 \leq i \leq N$，$\mathbb{V}\mathrm{ar}[[w_i \neq 0]] = \mathbb{E}\left[[w_i \neq 0]^2\right] - \mathbb{E}[[w_i \neq 0]]^2 = \kappa_i - \kappa_i^2$。由此可得 $\mathbb{V}\mathrm{ar}[\mathsf{size}(w)] = \sum_{i=1}^{N}\left(\kappa_i - \kappa_i^2\right) \leq \sum_{i=1}^{N}\kappa_i = \mathbb{E}[\mathsf{size}(w)]$。

因此，根据切比雪夫不等式，

$$\mathbb{P}\left[|\mathsf{size}(w) - \mathbb{E}[\mathsf{size}(w)]| \geq \frac{1}{2}\mathbb{E}[\mathsf{size}(w)]\right] \leq \frac{\mathbb{V}\mathrm{ar}[\mathsf{size}(w)]}{\left(\frac{1}{2}\mathbb{E}[\mathsf{size}(w)]\right)^2} \leq \frac{4}{\mathbb{E}[\mathsf{size}(w)]},$$

由于 $\mathbb{E}[\mathsf{size}(w)] \geq 1/\varepsilon^2$，这意味着 $\leq 4\varepsilon^2$：如果对某个 i 有 $\kappa_i = 1$（即 $I \neq \varnothing$），则我们知道 $\mathbb{E}[\mathsf{size}(w)] = \sum_{i=1}^{N}\kappa_i \geq 1 \geq 1/\varepsilon^2$（回顾 $\varepsilon \in (0, 1)$）。如果对所有 i 都有 $\kappa_i < 1$（即 $I = \varnothing$），则有 $\mathbb{E}[\mathsf{size}(w)] = \sum_{i=1}^{N}\kappa_i = \kappa\sum_{i=1}^{N}p_i \overset{(\star)}{\geq} \kappa \geq 1/\varepsilon^2$，其中（$\star$）由以下界限得出。

断言：我们始终有 $\sum_{i=1}^{N}p_i \geq 1$。
固定任意元素 $A^* \in \mathcal{D}$，使得 $F(A^*) \neq 0$。然后

$$\sum_{i=1}^{N}p_i = \sum_{i=1}^{N}\max_{\substack{A\in\mathcal{D}\\F(A)\neq 0}}\frac{f_i(A)}{F(A)} \geq \sum_{i=1}^{N}\frac{f_i(A^*)}{F(A^*)} = \frac{1}{F(A^*)}\sum_{i=1}^{N}f_i(A^*) = \frac{1}{F(A^*)}\cdot F(A^*) = 1.$$

注意到 $\mathbb{P}\left[\mathsf{size}(w) \geq \frac{3}{2}\mathbb{E}\left[\mathsf{size}(w)\right]\right] \leq \mathbb{P}\left[|\mathsf{size}(w) - \mathbb{E}\left[\mathsf{size}(w)\right]| \geq \frac{1}{2}\mathbb{E}\left[\mathsf{size}(w)\right]\right] \leq 4\varepsilon^2$（因为前者蕴含后者），证明得证。 □

推论（推论6 重述）。在算法1的设定中，令 $\widehat{p}_1,\ldots,\widehat{p}_N \in \mathbb{R}_{\geq 0}$ 对所有 $\widehat{p}_i \geq p_i$ 均满足 $1 \leq i \leq N$。若算法1在第 $\widehat{p}_i$ 行中使用 $p_i =$ 的 $F^{A\in\mathcal{D}}_{A}()\neq 0$ $F_{(A)\,(A)}$ 而非 max，则返回一个向量4，使得 $w \in \mathbb{R}^N$ 满足

(i) $\mathbb{P}\left[w \text{ 是一个 } \varepsilon\text{稀疏化器}\right] \geq 1-\delta$,
(ii) $\mathbb{E}\left[\text{大小（}w\text{）}\right] = \mathcal{O}\left(\frac{\log|\mathcal{D}|+\log}{\varepsilon^2}\frac{1}{\delta}\sum_{i=1}^{N}\widehat{p}_i\right)$,
(iii) $\mathbb{P}\left[\text{大小}(w) \leq \frac{3}{2}\mathbb{E}\left[\text{大小}(w)\right]\right] \geq 1-4\varepsilon^2$。

证明。定理3 的证明中所有步骤均可仿照进行。对于第(i)部分，关键在于估计 a这个量，其中

$$\frac{f_i(A)}{\kappa \cdot \max_{S\in\mathcal{D}} \frac{f_i(S)}{F(S)}} \leq \frac{f_i(A)}{\kappa \cdot \frac{f_i(A)}{F(A)}} = \frac{F(A)}{\kappa}$$

对所有 $i \in I$ 均成立。由于对每个 i，$\widehat{p}_i \geq p_i$，因此我们有

$$w_i f_i(A) = \frac{f_i(A)}{\kappa \cdot \widehat{p}_i} \leq \frac{f_i(A)}{\kappa \cdot p_i},$$

对每个 i 满足 $\kappa\widehat{p}_i \leq 1$，因此同样的界 $a \leq F(A)/\kappa$ 得以成立。指标 i 满足$\kappa\widehat{p}_i < 1$ 的情形无需任何修改即可处理。部分 (ii) 完全相同。部分 (iii) 基本上也相同，只有论断 $\sum_{i=1}^{N} p_i \geq 1$ 需要针对 $\widehat{p}_i$进行验证。不过这显然成立，因为 $\sum_{i=1}^{N}\widehat{p}_i \geq \sum_{i=1}^{N} p_i \geq 1$。 □

B （单调）次模函数

在第 B.1 节给出关于次模函数的一些预备结果后，我们将在第 B.2 节证明不等式（5），在此重述为

$$\sum_{i=1}^{N} p_i \leq Bn,$$

对一般的次模 f_i而言并不成立，因此否定了 [34]。随后，我们将在第 B.3 节证明（5）对单调次模 f_i成立。结合利用椭球法对峰值贡献进行的 $\mathcal{O}(\sqrt{n})$-近似 [3]，，可得定理 3(ii) 可简化为

$$\mathbb{E}\left[\mathsf{size}(w)\right] = \mathcal{O}\left(B\frac{n^2 + n\log\frac{1}{\delta}}{\varepsilon^2}\right) \tag{27}$$

如果所有的 f_i 都是单调次模的，那么

$$\mathcal{O}\left(\frac{Bn^2}{\varepsilon^2}\right) \tag{28}$$

对于固定的 δ。

B.1 预备知识

对于一个 次模函数 f： $2^E \to \mathbb{R}$，次模多面体 f [15] 定义为

$$\mathcal{P}(f) := \left\{x \in \mathbb{R}^E \mid x(S) \le f(S) \text{ for all } S \subseteq E\right\} \tag{29}$$

以及 f [15] 的基多面体定义为

$$\mathcal{B}(f) := \{x \in \mathcal{P}(f) \mid x(E) = f(E)\}, \tag{30}$$

其中， x 是一个 $|E|$维向量， $x(S)$ 表示量 $\sum_{e\in S} x_e$。

我们用 EX(P) 表示多面体 P 的极点集合。特别地，EX($\mathcal{B}(f)$) 是次模函数 f 的基多面体的极点集合。

我们将需要一个刻画次模函数的基多面体极点的结果。

定理 12（[15]）。设 f： $2^E \to \mathbb{R}$ 为次模函数。点 $y \in \mathcal{B}(f)$ 是 $\mathcal{B}(f)$ 的极点，当且仅当存在一条极大链（关于包含关系）

$$\mathcal{C} : \varnothing = S_0 \subset S_1 \subset \cdots \subset S_n = E \tag{31}$$

2^E，我们有 $y_{e_i} = f(S_i) - f(S_{i-1})$，其中 $i = 1$、.....、 n，且 $\{e_i\} = S_i \setminus S_{i-1}$。

推论 13。设 f： $2^E \to \mathbb{R}$ 为单调次模函数。则 EX($\mathcal{B}(f)$) $\subseteq \mathbb{R}^E_{\ge 0}$，即基多面体的任意极点都只有非负坐标。

证明。设 $y \in$ EX($\mathcal{B}(f)$) 是基多面体的一个极点。根据定理 12，存在一条极大链 $\varnothing = S_0 \subset S_1 \subset \cdots \subset S_n = E$使得 $y_{e_i} = f(S_i) - f(S_{i-1})$，其中$1 \le i \le n$，且 $\{e_i\} = S_i \setminus S_{i-1}$。由单调性可得，

$$y_{e_i} = f(S_i) - f(S_{i-1}) = f(S_{i-1} \cup \{e_i\}) - f(S_{i-1}) \ge f(S_{i-1}) - f(S_{i-1}) = 0$$

对于每个 $1 \le i \le n$。由于该链是极大的，每个元素 $e \in E$ 都作为某个 e_i 出现，这意味着对所有 $e \in E$， $y_e \ge 0$。 □

下一个引理将次模函数的值 $f(A)$ 表示为在基多面体上的最大化问题。

引理 14. 设 f： $2^E \to \mathbb{R}$ 为归一化次模。则

$$f(A) = \max_{y \in \text{EX}(\mathcal{B}(f))} \langle y, \mathbb{1}_A \rangle$$

对所有 $A \subseteq E$ 均成立。

证明。由于任意 $y \in \mathcal{B}(f)$ 都满足 $y(S) \le f(S)$ 对所有 $S \subseteq E$ 均成立，包括 $S = A$，因此不等式 $\max_{y\in \text{EX}(\mathcal{B}(f))} \langle y, \mathbb{1}_A \rangle \le f(A)$ 显然成立。于是

$$\langle y, \mathbb{1}_A \rangle = \sum_{e\in E} y_e \cdot \mathbb{1}_A(e) = \sum_{e\in A} y_e = y(A) \le f(A).$$

为了说明该值 $f(A)$ 确实能够取到，我们利用定理 12 构造一个合适的极点。假设 $A = \{e_1, \ldots, e_k\}$。任意排列其余元素，使得 $E \setminus A = \{e_{k+1}, \ldots, e_n\}$。现在定义一条极大链

$$\mathcal{C} : \varnothing = S_0 \subset S_1 \subset \cdots \subset S_n = E$$

通过令 $S_i := \{e_1, \ldots, e_i\}$ 用于 $0 \le i \le n$。接下来，我们设

$$y_{e_i} := f(S_i) - f(S_{i-1})$$

对于 $1 \le i \le n$，这定义了一个向量 $y \in \mathbb{R}^E$，根据构造，它恰好是 $\mathcal{B}(f)$ 的一个极点（由定理 12 得出）。此外，它满足

$$y(A) = \sum_{e \in A} y_e = \sum_{i=1}^{k} y_{e_i} = \sum_{i=1}^{k} (f(S_i) - f(S_{i-1})) = f(S_k) - f(\varnothing) = f(S_k) = f(A)$$

由于 f 已归一化。因此，y 证明了 $\max_{y \in \mathsf{EX}(\mathcal{B}(f))} \langle y, \mathbb{1}_A \rangle \ge f(A)$。 □

B.2 反例

现在，我们将通过给出一个反例，推翻 Rafiey 和 Yoshida 所声称的界限 $\sum_{i=1}^{N} p_i \le Bn$（参见附录中 [34], 证明里的断言 3.3）。

这个反例的灵感来自 Cohen 等人 [11], ，他们利用它证明了割稀疏器大小的下界。这正是我们推翻 $\sum_{i=1}^{N} p_i \le Bn$所需要的：能够隔离边，从而使峰值贡献达到 1。考虑有向完全二部图 $G = (V, E)$，其二部划分为 $V = L \cup R$，其中 $L = \{u_1, \ldots, u_5\}$，$R = \{v_1, \ldots, v_5\}$ 且 $E = L \times R$，如图 1 所示。对于每条边 $e = (u, v) \in E$，我们定义一个割函数

$$f_e : 2^V \to \mathbb{R}, \quad S \mapsto \begin{cases} 1 & \text{if } u \in S \text{ and } v \notin S, \\ 0 & \text{otherwise.} \end{cases}$$

众所周知（而且不难看出），f_e是次模的。

和 $F := \sum_{e \in E} f_e$ 称为 G 的（有向）割函数，因为 $F(S)$ 是从 S 中的顶点连接到 S 外部顶点的割边数。令 $p_e = \max_{S \subseteq V} \frac{f_e(S)}{F(S)}(S)$ 表示这些 f_e 的峰值贡献。

现在我们将证明这个例子如何违反 $\sum_{e \in E} p_e \le Bn$。首先，

$$B = \max_{e \in E} |\mathsf{EX}(\mathcal{B}(f_e))| = 2$$

因为每个基多面体 $\mathcal{B}(f_i)$ 都有两个极点。根据定理 12，可以很容易地看出这一点。对于任意极大链 $\varnothing = V_0 \subset V_1 \subset \cdots \subset V_n = V$，我们有

$$f_e(V_{i+1}) - f_e(V_i) = \begin{cases} 1 & \text{if } V_{i+1} \setminus V_i = \{u\} \text{ and } v \notin V_i, \\ -1 & \text{if } V_{i+1} \setminus V_i = \{v\} \text{ and } u \in V_i, \\ 0 & \text{otherwise.} \end{cases}$$

因此，唯一的极点是 y 和 y'，其中 $y_e = 0$ 对所有 $e \in E$均成立，且 $y'_u = 1$、$y'_v = -1$、$y'_e = 0$ 对所有 $e \in E \setminus \{u, v\}$均成立。由于 G 有 10 个顶点，且顶点集 $V = L \cup R$ 是

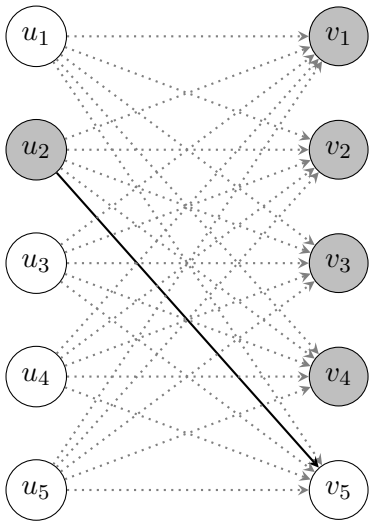


图1：图 G，其中每条边都可以通过一个割隔离

基集，[34] 断言 3.3 指出 $\sum_{e\in E} p_e \le Bn = 20$。通过证明 $p_e = 1$ 对每个$e \in E$均成立，我们便可确立

$$\sum_{e\in E} p_e = |E| = |L| \cdot |R| = 5 \cdot 5 = 25 > 20,$$

这与该断言矛盾。固定一条边 $e = (u, v)$。令 $S := \{u\} \cup (R \setminus \{v\})$，我们得到$f_e(S) = 1$，并且对于每条边 $e' \neq e$，都有 $f_{e'}(S) = 0$。后者是因为 S 要么

- 包含 e' 的右端点，或者
- 不包含 e' 的左端点。

这两种情况互斥。根据 $f_{e'}$ 的定义，两种情况都推出 $f_{e'}(S) = 0$。因此，

$$p_e = \max_{A\subseteq V} \frac{f_e(A)}{F(A)} \ge \frac{f_e(S)}{F(S)} = \frac{f_e(S)}{f_e(S) + \sum_{e'\neq e} f_{e'}(S)} = \frac{1}{1+0} = 1,$$

由此可得 $p_e = 1$，证毕。

B.3 单调次模性

现在，我们具备了建立 $\sum_{i=1}^{N} p_i$上界所需的全部工具，从而使定理3(ii) 中的大小界更简洁、更具表现力。

以下引理在 [34] 中以断言 3.3 的形式出现，证明非常相似。然而，下面引理 15 的证明仅在基多面体极点的坐标均为非负时才成立。我们将更详细地解释各个步骤，并指出具体在哪些地方需要非负性，也就是 [34] 中的证明在哪些地方 [34] 不成立。推论 13 保证了当所有 f_i均为单调时的非负性。对于下面的引理，我们专门考虑 $\mathcal{D} = 2^E$，其中基集 E 的大小为 n。

引理 15. 设给定归一化的单调次模函数 $f_1,\ldots,f_N\colon 2^E \to \mathbb{R}$，其峰值贡献是相对于 $F = f_1 + \cdots + f_N$ 定义的。则 $\sum_{i=1}^N p_i \le Bn$，其中 $B = \max_{1\le i\le N} |\mathsf{EX}(\mathcal{B}(f_i))|$ 表示这些 f_i的基多面体的极点数量的最大值。

证明。 关键在于引理 14。它给出

$$\sum_{i=1}^N p_i = \sum_{i=1}^N \max_{A\subseteq E} \frac{f_i(A)}{F(A)} = \sum_{i=1}^N \max_{A\subseteq E} \frac{\max_{y\in\mathsf{EX}(\mathcal{B}(f_i))}\langle y, \mathbb{1}_A\rangle}{\sum_{j=1}^N \max_{y\in\mathsf{EX}(\mathcal{B}(f_j))}\langle y, \mathbb{1}_A\rangle} \tag{32}$$

利用 p_i 的定义。由于所有 f_i 都是单调的，我们知道，对于每个 $1 \le i \le N$， $y \in \mathbb{R}^E_{\ge 0}$ 对所有 $y \in \mathcal{B}(f_i)$ 都成立，因此上式中涉及的所有项 $\langle y, \mathbb{1}_A\rangle$ 都是非负的。对于非负数 $x_1, \ldots, x_n \in \mathbb{R}_{\ge 0}$, 我们总有

$$\frac{x_1 + \cdots + x_n}{n} \overset{(1)}{\le} \max_{1\le i\le n} x_i \overset{(2)}{\le} x_1 + \cdots + x_n, \tag{33}$$

也就是说，最大值介于平均值与总和之间。这是 Rafiey 和 Yoshida 在证明中使用的一个技巧 [34] ，但对一般次模函数并不成立。在本证明余下的几个步骤中，非负性发挥着重要作用。将（1）应用于式（32）的分母，并将（2）应用于其分子，我们得到

$$\frac{\max_{y\in\mathsf{EX}(\mathcal{B}(f_i))}\langle y, \mathbb{1}_A\rangle}{\sum_{j=1}^N \max_{y\in\mathsf{EX}(\mathcal{B}(f_j))}\langle y, \mathbb{1}_A\rangle} \le \frac{\sum_{y\in\mathsf{EX}(\mathcal{B}(f_i))}\langle y, \mathbb{1}_A\rangle}{\sum_{j=1}^N \frac{1}{|\mathsf{EX}(\mathcal{B}(f_j))|}\sum_{y\in\mathsf{EX}(\mathcal{B}(f_j))}\langle y, \mathbb{1}_A\rangle} \tag{34}$$

$$= \frac{\sum_{e\in A}\sum_{y\in\mathsf{EX}(\mathcal{B}(f_i))} y_e}{\sum_{e\in A}\sum_{j=1}^N \frac{1}{|\mathsf{EX}(\mathcal{B}(f_j))|}\sum_{y\in\mathsf{EX}(\mathcal{B}(f_j))} y_e}, \tag{35}$$

其中等号成立是根据 $\langle y, \mathbb{1}_A\rangle = \sum_{e\in A} y_e$。接下来，注意到 $\sum_{y\in\mathsf{EX}(\mathcal{B}(f_i))} y_e \ge 0$ 且 $\sum_{j=1}^N \frac{1}{|\mathsf{EX}(\mathcal{B}(f_j))|}\sum_{y\in\mathsf{EX}(\mathcal{B}(f_j))} y_e \ge 0$ 对所有 $e \in A$ 都成立，因为对于所有 $y \in \mathcal{B}(f_i)$, $y_e \ge 0$ 都成立，且 $e \in E$，因此我们可以使用不等式

$$\frac{x_1 + \cdots + x_n}{q_1 + \cdots + q_n} \le \max_{1\le i\le n} \frac{x_i}{q_i} \tag{36}$$

适用于所有非负数 $x_1, \ldots, x_n, q_1, \ldots, q_n \in \mathbb{R}_{\ge 0}$，得到

$$\frac{\sum_{e\in A}\sum_{y\in\mathsf{EX}(\mathcal{B}(f_i))} y_e}{\sum_{e\in A}\sum_{j=1}^N \frac{1}{|\mathsf{EX}(\mathcal{B}(f_j))|}\sum_{y\in\mathsf{EX}(\mathcal{B}(f_j))} y_e} \le \max_{e\in A} \frac{\sum_{y\in\mathsf{EX}(\mathcal{B}(f_i))} y_e}{\sum_{j=1}^N \frac{1}{|\mathsf{EX}(\mathcal{B}(f_j))|}\sum_{y\in\mathsf{EX}(\mathcal{B}(f_j))} y_e}. \tag{37}$$

将此结果与方程（34）和方程（35）结合起来，可将方程（32）推广为

$$\sum_{i=1}^N p_i \le \sum_{i=1}^N \max_{A\subseteq E}\max_{e\in A} \frac{\sum_{y\in\mathsf{EX}(\mathcal{B}(f_i))} y_e}{\sum_{j=1}^N \frac{1}{|\mathsf{EX}(\mathcal{B}(f_j))|}\sum_{y\in\mathsf{EX}(\mathcal{B}(f_j))} y_e} \tag{38}$$

$$= \sum_{i=1}^N \max_{e\in E} \frac{\sum_{y\in\mathsf{EX}(\mathcal{B}(f_i))} y_e}{\sum_{j=1}^N \frac{1}{|\mathsf{EX}(\mathcal{B}(f_j))|}\sum_{y\in\mathsf{EX}(\mathcal{B}(f_j))} y_e} \tag{39}$$

$$\le \sum_{i=1}^N \sum_{e\in E} \frac{\sum_{y\in\mathsf{EX}(\mathcal{B}(f_i))} y_e}{\sum_{j=1}^N \frac{1}{|\mathsf{EX}(\mathcal{B}(f_j))|}\sum_{y\in\mathsf{EX}(\mathcal{B}(f_j))} y_e}, \tag{40}$$

其中，方程（39）成立，是因为对所有 $A \subseteq E$ 和 $e \in A$ 取最大值，等同于在此对所有 $e \in E$ 取最大值，因为右端仅依赖于 e。方程（40）则再次由方程（33）得出——所有商均非负，因为其分子和分母均非负。经过几步代数步骤即可得到所需结论。将 $\sum_{i=1}^{N} \dots$ 与 $\sum_{e \in E} \dots$ 互换，并使用 $|\mathsf{EX}(\mathcal{B}(f_j))| \leq \max_{1 \leq \ell \leq n} |\mathsf{EX}(\mathcal{B}(f_\ell))| = B$，可得

$$\sum_{i=1}^{N} \sum_{e \in E} \frac{\sum_{y \in \mathsf{EX}(\mathcal{B}(f_i))} y_e}{\sum_{j=1}^{N} \frac{1}{|\mathsf{EX}(\mathcal{B}(f_j))|} \sum_{y \in \mathsf{EX}(\mathcal{B}(f_j))} y_e} \tag{41}$$

$$= \sum_{e \in E} \sum_{i=1}^{N} \frac{\sum_{y \in \mathsf{EX}(\mathcal{B}(f_i))} y_e}{\sum_{j=1}^{N} \frac{1}{|\mathsf{EX}(\mathcal{B}(f_j))|} \sum_{y \in \mathsf{EX}(\mathcal{B}(f_j))} y_e} \tag{42}$$

$$\leq \sum_{e \in E} \sum_{i=1}^{N} \frac{\sum_{y \in \mathsf{EX}(\mathcal{B}(f_i))} y_e}{\frac{1}{B} \sum_{j=1}^{N} \sum_{y \in \mathsf{EX}(\mathcal{B}(f_j))} y_e} \tag{43}$$

$$= \sum_{e \in E} B = Bn, \tag{44}$$

从而得到所声称的 $\sum_{i=1}^{N} p_i \leq Bn$ 通过等式（38）、等式（39）和等式（40）。 □

这个引理有趣之处在于，界 $\sum_{i=1}^{N} p_i \leq Bn$ 不依赖于N。在稀疏化有意义的情形下，通常有 $N \gg n$。

注释 16. 如果所有 f_i 都是单调的，我们至少可以得出以下结论，这一点 Rafiey 和 Yoshida [34] 已经观察到。存在一种基于椭球法的算法，给定两个单调次模函数 f, g：$2^E \to \mathbb{R}_{\geq 0}$，可在多项式时间内将 $\max_{A \subseteq E} f(A) g(A)$ 近似到因子 $\mathcal{O}(\sqrt{n} \log n)$（参见 [34] 和 [3]中的定理 3.4）。它可用于获得近似值 $\widehat{p}_1, \dots, \widehat{p}_N$，使得对于 $1 \leq i \leq N$，有 $p_i \leq \widehat{p}_i \leq \mathcal{O}(\sqrt{n} \log n) p_i$。根据推论 6，我们可以使用这些 $\widehat{p}_i$ 执行算法 1，并在多项式时间内得到一个期望大小为 $\mathcal{O}\left(\frac{n + \log \frac{1}{\delta}}{\varepsilon^2} \sum_{i=1}^{N} \widehat{p}_i\right)$ 的 ε-稀疏化器。由于这些 f_i 是单调的，引理 15 适用，因此对于固定的 δ，界变为 $\mathcal{O}\left(Bn^{2.5} \log_{\varepsilon^2} n\right)$。

C 有界元数的单调次模函数

次模稀疏化的动机部分源于图割的稀疏化；图割不仅在图中得到了广泛研究，在超图中也是如此 [39, 25, 10]。如果相关的次模函数 ε-稀疏化器所针对的函数 F 可以分解为 $F = f_1 + \dots + f_N$ 个 f_i 元数有界的 组成函数，则可以同时缩减其规模并加快构造速度。这涵盖了广泛的次模函数，例如割函数。作为核心算法的一个简单应用，我们将证明：如果所有组成函数 ε均为单调函数（引理 $\mathcal{O}(n^2/\varepsilon^2)$ 所需，其中 δ 固定）且为次模函数（引理 f_i 所需），则可以找到一个期望规模为 F 的 15）以及引理 18 和引理 19）。

具体而言，考虑集合函数 $f: 2^E \to \mathbb{R}$。我们不再用 $f(S)$ 表示 $S \subseteq E$ 来描述 f 的取值，而是可以将 S 编码为一个具有 $|E|$ 个坐标的布尔向量。任意固定 E 中元素的顺序，例如 $E = \{e_1, \dots, e_n\}$，并将对应于 S 的向量的 i^{th} 坐标设为 1，当且仅当 $e_i \in S$。这种标准的对应关系使我们可以用 $f(x_1, \dots, x_n)$ 表示 $(x_1, \dots, x_n) \in \{0,1\}^n$，以描述 f 的取值。从这个角度来看，

直观地说， f 的元数 a ，是指其取值仅取决于 a 个自变量。值得注意的是，这种表示方式与在约束满足问题中表示 f 的方式一致，后者是赋值约束语言 [26]的一部分。注意， f 元数为 a 意味着存在一个集合 $C \subseteq E$ ，其大小为 $|C| = a$ ，使得对于所有 $f(S)$， $= f(S \cap C)$。我们称 $S \subseteq E$为 C 的有效支撑。 f。

我们从将次模函数 F 分解为元数为 $\leq a$ 的次模函数 $f_1, \ldots, f_N$ （其中 $\leq a$ 为某个常数）的分解 $F = f_1 + \cdots + f_N$ 开始。令 $C_1, \ldots, C_N$ 分别表示 $f_1, \ldots, f_N$ 的有效支撑。为了构造稀疏化器，我们想要应用算法 1。为了高效地做到这一点，我们需要能够计算峰值贡献，即对于每个 $1 \leq i \leq N$，在所有 $A \subseteq E$ 中最大化商式 $\frac{f_i(A)}{F(A)}$。由于分子仅取决于 A 与各个 C_i 的交集情况（这些集合的大小都是常数），我们可以穷举这些交集，并在一个额外约束下最小化分母。

备注 17。要使这一想法奏效，我们必须知道各个 C_i。在某些情况下，虽然已知所有 f_i 的元数都有界，但它们所支撑的元素在计算上无法获取。不过，如果 f 是单调且非负的，就可以通过对每个元素 $e \in E$ 检验 $f(\{e\}) > f(\varnothing)$ 来找到有效支撑。如果检验不通过，我们就知道 e 不会影响 f 的取值，因此不属于其有效支撑。

算法 3：计算有界元数 f_i 的峰值贡献

要求：函数 $f_1, \ldots, f_N$ 的有效支撑 $C_1, \ldots, C_N$ 的大小为 $\leq a$

确保：对每个 $1 \leq i \leq N$， $p_i = \max_{A \subseteq E} \frac{f_i(A)}{F(A)}$

对于 $i = 1, \ldots, N$ **do**

  $p_i \leftarrow 0$

  对于所有 $H \subseteq C_i$ ，执行

    $p_i \leftarrow$ 最大值 $\left\{ p_i, \frac{f_i(H)}{\widehat{F}(A^*)} \right\}$ ，其中 A^* 是 $\widehat{F}_{f_i}$ 的一个最小化解： $2^{E \setminus C_i} \to \mathbb{R}, A \mapsto F(A \cup H)$

  结束循环

结束循环

返回 $p_1, \ldots, p_N$

有关如何计算峰值贡献的正式说明，请参阅算法 3。对于每个 $1 \leq i \leq N$，按如下方式计算量 p_i 。我们首先选择一个集合 $H \subseteq C_i$ ，然后将范围限制在满足 $A \cap C_i = H$ 的集合 $A \subseteq E$ 上。这意味着 $f_i(A) = f_i(A \cap C_i) = f_i(H)$，因为 C_i 是 f_i 的有效支撑。因此，我们只需处理商 $f_i(H)/F(A)$，并在所有满足 $A \cap C_i = H$ 的 $A \subseteq E$ 上将其最大化。由于分子不再依赖于 A，目标就变成在满足 $A \cap C_i = H$ 的条件下最小化 $F(A)$。事实证明，这只是对辅助函数 F 的一个子模最小化问题。下面的结论是该方法正确性的关键。

引理 18. 设 f、 F： $2^E \to \mathbb{R}$ 为次模函数，并设 $C \subseteq E$ 表示 f 的有效支撑。则

$$\max_{A \subseteq E} \frac{f(A)}{F(A)} = \max_{H \subseteq C} \frac{f(H)}{\min_{A \subseteq E \setminus C} \widehat{F}(A)}$$

其中 $\widehat{F}$： $2^{E \setminus C} \to \mathbb{R}$， $A \mapsto F(A \cup H)$ 。

证明。“≤”令 A^* 为左端的最大化元。设 $H := A^* \cap C$，且 $A := A^* \setminus C$。于是我们得到

$$\frac{f(A^*)}{F(A^*)} = \frac{f(A^* \cap C)}{F((A^* \setminus C) \cup (A^* \cap C))} = \frac{f(H)}{F((A^* \setminus C) \cup H)} = \frac{f(H)}{\widehat{F}(A)}$$

具有 $H \subseteq C$ 和 $A \subseteq E \setminus C$，因此这显然就是上面的右端 $\leq$。

“≥” 设 (H^*, A^*) 是右端的一个最大化对。令 $A := A^* \cup H^*$，可得

$$\frac{f(H^*)}{\widehat{F}(A^*)} = \frac{f(A \cap C)}{F(A^* \cup H)} = \frac{f(A)}{F(A)},$$

由于最大化是在所有 A 上进行的，因此这始终是左端 $\leq$。 □

另一个观察结果使我们能够在有界元数的情形下给出更精确的稀疏化器大小界。根据定理3，由算法 ε生成的稀疏化器，其期望大小为 $\mathcal{O}\left(\log|\mathcal{D}|+_{\log \varepsilon^2} \quad ^{\frac{1}{\delta}} \sum_{i=1}^{N} p_i\right)$，其中 $\sum_{i=1}^{N} p_i \leq Bn$ 根据引理 15，若这些 f_i均为单调函数。一般来说，量 B 可能很大。然而，如果这些函数都具有有界元数，我们可以证明 $B = \mathcal{O}(1)$，从而在组成函数为单调函数的情况下，可以得到期望大小为 $\mathcal{O}(n^2/\varepsilon^2)$ 的稀疏化器——假设 δ 为常数。

下述在 [34] 中提出但未证明的结论给出了一个关于 B 的上界，该上界与 n 或 N 无关，只取决于元数。为求完整起见，我们给出证明。

引理 19. 设 $f : 2^E \to \mathbb{R}$ 是元数为 $\leq a$ 的次模函数。那么，基多面体 $\mathcal{B}(f)$ 至多有 2^{a^2} 个极点，即 $|\mathrm{EX}(\mathcal{B}(f))| \leq 2^{a^2}$。

证明. 根据定理 12，点 $x \in \mathcal{B}(f)$ 是极点，当且仅当存在一条极大链 $\varnothing = S_0 \subset S_1 \subset \cdots \subset S_n = E$，满足 $x(e_i) = f(S_i) - f(S_{i-1})$，其中 $1 \leq i \leq n$，且 $\{e_i\} = S_i \setminus S_{i-1}$。令 C 表示 f 的有效支撑，则这些等式变为 $x(e_i) = f(S_i \cap C) - f(S_{i-1} \cap C)$。现在由 $e_i \notin C$ 可得

$$S_i \cap C = (S_{i-1} \cup \{e_i\}) \cap C = S_{i-1} \cap C,$$

因此 $x(e_i) = f(S_i \cap C) - f(S_{i-1} \cap C) = f(S_{i-1} \cap C) - f(S_{i-1} \cap C) = 0$。所以，只有与元素 $e \in C$ 对应的坐标可能非零。对于任意这样的坐标 $e \in C$，我们有

$$x(e) = f(S \cup \{e\}) - f(S) = f((S \cup \{e\}) \cap C) - f(S \cap C)$$

其中， S 是极大链 e 中加入 2^a 的（唯一）集合。现在注意到，右端最多可以取 $S \subseteq C$种可能的值：每个子集 S 对应一种。这是因为该值完全由 C与 的交集决定。

综合以上结果，非零坐标至多有 a个，每个坐标至多有 2^a 个不同的取值。计算所有组合，可得 x的候选项上界为 $(2^a)^a = 2^{a^2}$。因此， $\mathcal{B}(f)$至多有 2^{a^2} 个极点。

□