arXiv:2603.07387v1 [cs.DS] 8 Mar 2026

Ap使用 Sketch 近似张量网络收缩 es

Mike Heddes 加利福尼亚
大学欧文分校mheddes@uci.edu

Igor Nunes 加利福尼亚
大学欧文分校igord@uci.edu

Tony Givargis 加利福尼亚
大学欧文分校givargis@uci.edu

Alex Nicolau 加利福尼亚大学
欧文分校nicolau@ics.uci.edu

摘要

张量网络收缩是一种基础的数学运算，是点积和矩阵乘法的推广。它广泛应用于数据库系统、图论、机器学习、概率论和量子力学等众多领域。一般而言，张量网络收缩的计算开销很大，通常需要指数级的时间和空间。草图方法包含多种降维技术，广泛应用于近似算法的设计。然而，现有的张量网络收缩草图方法仅支持无环张量网络。我们提出了首个能够近似任意张量网络收缩的方法，包括有环张量网络的收缩。此外，我们证明，现有草图方法的计算复杂度会随收缩次数呈指数级增长。我们还针对无环张量网络提出了第二种方法，其空间和时间复杂度仅随收缩次数呈多项式级增长。

1 引言

张量网络收缩（TNC）是一种基础的数学运算，应用于量子物理、机器学习、数据库系统、图论和概率论等多个领域。正如张量将向量和矩阵推广到更高阶一样，张量网络收缩也将点积和矩阵乘法推广开来。张量网络收缩可以理解为对多个张量的张量收缩（见定义 1）与张量积（见定义 2）进行运算。张量收缩由一组共享指标指定，这些指标构成张量之间的连接，从而形成张量网络。

下面的张量网络收缩定义使用艾弗森括号记法 $[S]$：当命题 S 为真时，其值为 1，否则为 0；整数集合 $[k] \equiv \{1, \dots, k\}$ 表示任意非负整数 k 对应的连续整数集合；此外，使用 $\mathbf{i}_K$ 表示多重索引 $\mathbf{i} \equiv (i_1, \dots, i_q)$，其取值范围限定为 $K \subseteq [q]$ 中的指标。关于张量和多重指标记法的更多细节，请参见第 2 节。

定义 1（张量收缩）。给定任意一个阶数为 q 的张量 $\mathcal{X} \in \mathbb{R}^{n_1 \times \cdots \times n_q}$ 和收缩 $E \subset [q] \times [q]$，使得对于每个 $(u, v) \in E$，$u \neq v$ 和 $n_u = n_v$。令 $K = \{u, v : (u, v) \in E\}$，且 $K' = [q] \setminus K$。张量收缩 $\mathrm{tc}(\mathcal{X}, E)$ 是一个阶数为 $|K'|$ 的张量，定义如下：

$$\mathrm{tc}(\mathcal{X}, E)(\mathbf{i}_{K'}) = \sum_{\mathbf{i}_K = \mathbf{1}_K}^{\mathbf{n}_K} \mathcal{X}(\mathbf{i}) \prod_{(u,v) \in E} [i_u = i_v], \qquad \forall \mathbf{1}_{K'} \leq \mathbf{i}_{K'} \leq \mathbf{n}_{K'}.$$

定义 2（张量积）。给定任意两个张量 $\mathcal{A} \in \mathbb{R}^{n_1 \times \cdots \times n_p}$ 和 $\mathcal{B} \in \mathbb{R}^{m_1 \times \cdots \times m_q}$，其阶数分别为 p 和 q，它们的张量积 $\mathcal{A} \times \mathcal{B}$ 定义为 $(\mathcal{A} \times \mathcal{B})(\mathbf{i}, \mathbf{j}) = \mathcal{A}(\mathbf{i})\mathcal{B}(\mathbf{j})$，对所有 $\mathbf{1} \leq \mathbf{i} \leq \mathbf{n}$ 以及 $\mathbf{1} \leq \mathbf{j} \leq \mathbf{m}$ 均成立，其中 $\mathcal{A} \times \mathcal{B} \in \mathbb{R}^{n_1 \times \cdots \times n_p \times m_1 \times \cdots \times m_q}$。

定义 3（张量网络收缩（TNC））。对于任意 $p \in \mathbb{N}^+$、张量 $\mathcal{X}_1$、.. . 、 $\mathcal{X}_p$ 以及收缩 E，张量网络收缩定义为 $\mathrm{tc}(\mathcal{X}_1 \times \cdots \times \mathcal{X}_p, E)$。

下面给出一个张量网络收缩（TNC）的示例，以阐明上述定义。示例 1 中的张量网络在图 1 中以张量图表示法表示，也称为彭罗斯图形表示法。在这种表示法中，张量用带有腿的形状表示，每条腿代表一个模态，连接的腿则表示一次收缩。

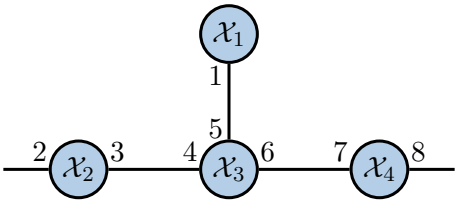


图 1. 一个张量网络示例的张量图

示例 1. 张量网络收缩 $\mathrm{tc}(\mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{X}_3 \times \mathcal{X}_4, E) = \mathcal{Y}$（如图 *1* 所示）由四个张量 $\mathcal{X}_1 \in \mathbb{R}^{n_1}$、$\mathcal{X}_2 \in \mathbb{R}^{n_2 \times n_3}$、$\mathcal{X}_3 \in \mathbb{R}^{n_4 \times n_5 \times n_6}$ 和 $\mathcal{X}_4 \in \mathbb{R}^{n_7 \times n_8}$ 以及三次收缩 $E = \{(1、5)、(3、4)、(6、7)\}$ 组成。根据定义 *1*，我们有 $K = \{1, 3, 4, 5, 6, 7\}$、$K' = \{2, 8\}$，对艾弗森括号进行化简后，得到

$$\mathcal{Y}(i_2, i_8) = \sum_{i_1=1}^{n_1} \sum_{i_3=1}^{n_3} \sum_{i_6=1}^{n_6} \mathcal{X}_1(i_1)\mathcal{X}_2(i_2, i_3)\mathcal{X}_3(i_3, i_1, i_6)\mathcal{X}_4(i_6, i_8), \qquad \forall (1,1) \leq (i_2, i_8) \leq (n_2, n_8).$$

张量网络收缩的一般问题已知是NP难的 [BC17, AL10]。在 [MS08], 中，作者表明精确算法的运行时间随张量网络的树宽呈指数增长。在本文中，我们基于草图化，为无环张量网络的张量网络收缩提出了一种多项式复杂度的(ϵ, δ)-近似。此外，我们提出了第一种能够近似任意张量网络收缩的方法，包括有环张量网络的收缩。据我们所知，这是第一项直接以草图化处理该一般问题的工作，并为该问题的重要实例建立了新的界。

草图化，即降维，因其能够在保持解空间近似的同时迅速缩小问题规模而备受关注。该技术已被应用于数值线性代数和机器学习中的若干基本问题，包括最小二乘回归和低秩近似 [Woo14]。在具有影响力的论文 [AMS96], 引入*AMS* 草图——一种用于估计数据流二阶频率矩的方法——之后，人们对这些方法的兴趣急剧上升。AMS 草图的一项显著改进，称为计数草图 [CCFC02], ，显著降低了草图化的时间复杂度。这两种草图后来作为内积估计器而广受欢迎 [AGMS99, CG05]。

在数据库理论领域，内积估计具有重要意义。它是查询优化的核心环节，尤其是连接大小估计。这一重要性促使数据库领域进一步发展草图方法。其中值得注意的成果包括将内积草图推广到多重连接查询的两种方法：首先是 AMS 草图 [DGGR02], ，随后是计数草图 [HNGN24]。在此基础上，我们的工作

提出了一种新视角，指出先前工作所研究的连接大小估计问题，对应于张量网络收缩的近似问题。

直觉上，将估计从单连接的连接大小扩展到多重连接的连接大小，对应于从近似内积到近似张量网络收缩这一更广泛的推广。基于这一视角，我们工作的动机源于如下观察：（1）上述连接大小估计量仅支持无环查询；（2）其精度会随着连接数量的增加而呈指数级下降，而这些连接对应于收缩。事实上，我们在第4.3节中针对 [HNGN24]所提出的技术给出了一个指数下界。

尽管我们最初的动机源于数据库理论，但我们将所涉问题重新表述为张量网络收缩问题，从而揭示了所提技术更广泛的适用性。在第5节中，我们讨论近似张量网络收缩这一更一般问题在数据库之外所发挥的重要作用。例如，张量网络收缩被用于模拟量子计算机 [BC17]；在概率论中，张量网络收缩对应于图模型中的边缘化 [RS19]；而计算图中三角形的数量可以归约为一个张量网络收缩。

1.1 技术概述

我们提出了两种（ϵ， δ）近似张量网络收缩（ATNC）方法（见定义 4）。ATNC 的误差界是相对于输入张量的范数定义的，而不是相对于张量网络收缩（TNC）的输出定义的。这是不可避免的，因为任何相对于 TNC 输出定义的误差界，都会意味着存在一种次线性空间的单向协议来解决 索引 问题，这与该问题的 $\Omega(n)$ 随机通信下界 [JKS08] 相矛盾。对于近似矩阵乘法，类似的输入相对误差界是标准做法 [Woo14], ；近似矩阵乘法是 TNC 的一个特例[1]。

定义 4 （近似张量网络收缩（ATNC））。 给定 $\epsilon, \delta > 0$ 以及任意正整数 p，随机函数 $\hat{}\mathrm{tc}$ 是一个（ϵ, δ）-近似张量网络收缩；对于任意满足 $\mathcal{X} = \mathcal{X}_1 \times \cdots \times \mathcal{X}_p$ 的张量网络收缩 $\mathrm{tc}(\mathcal{X}, E)$，

$$\Pr\left(\left\|\hat{\mathrm{tc}}(\mathcal{X}, E) - \mathrm{tc}(\mathcal{X}, E)\right\|_{\mathrm{F}} \leq \epsilon \|\mathcal{X}\|_{\mathrm{F}}\right) \geq 1 - \delta.$$

我们的第一种方法能够近似任意 TNC，而此前的草图方法只能近似无环 TNC [Dee25]。当张量网络中包含环时，TNC 称为有环 TNC；否则称为无环 TNC。该结果的关键是补计数草图，这是一种循环反向的计数草图，于第 4.2 节中提出。为每个收缩分别分配一个独立的计数草图及其补草图，使估算过程能够使用草图的循环卷积，而不是循环互相关。我们证明，[HNGN24] 中使用循环互相关，使其无法近似有环张量网络。随后，我们在第 4.3 节中证明了当前最佳方法的一个指数下界，这促使我们寻求一种新方法。

我们的第二种方法消除了无环张量网络收缩中对收缩次数的指数依赖。为实现这一点，我们将张量网络解释为树结构，并在第4.4节中将无环张量网络收缩表述为递归函数，该函数由一系列带有克罗内克积的矩阵乘法组成。随后，利用 [AKK+20]中的递归草图技术对克罗内克积进行草图化。为确保算法在空间和时间上高效，草图从树的叶节点开始，逐步向根节点增量构建。这两种方法得出了我们的主要结果，具体表述见定理 1。我们针对完整张量网络收缩陈述主要结果，即其张量网络收缩

[1]对于矩阵 $A, B \in \mathbb{R}^{n \times n}$，矩阵乘法 $AB = \mathrm{tc}(A \times B, \{(2,3)\})$。

结果为一个标量值。在第4.5节中，我们进一步说明，这些方法同样适用于部分张量网络收缩，其结果为一个非零阶张量。

定理 *1.* 对于任意 $p \in \mathbb{N}^+$、任意阶数为 q_k 的张量 $\mathcal{X}_k$ （ $k \in [p]$），以及任意 ϵ、$\delta > 0$，均存在一个完全张量网络收缩的（ϵ，δ）近似张量网络收缩，可在时间 $O((pm \log m + qN) \log 1/\delta)$ 内计算，并使用 $O(mp \log 1/\delta)$ 空间，其中 *(1)* 在无环情形下 $m = \Omega(t/\epsilon^2)$，*(2)* 一般情形下 $m = \Omega(3^t/\epsilon^2)$；这里 $N = \sum_{k=1}^{p} nnz(\mathcal{X}_k)$、$q = \sum_{k=1}^{p} q_k$ 和 $t = |E|$。

在表 1 中，我们将所提出的两种方法的时间与空间复杂度与现有方法进行了比较。我们用 $N = \sum_{k=1}^{p} \mathrm{nnz}(\mathcal{X}_k)$、$q = \sum_{k=1}^{p} q_k$ 和 $t = |E|$ 对复杂度进行参数化，它们分别表示非零分量的数量、模态数和收缩次数。请注意，现有的草图方法仅支持无环张量网络。在这种情形下，我们的方法使草图维度呈指数级降低，因此时间与空间复杂度也随之大幅改善。此外，我们提出了首个能够近似一般张量网络收缩的方法，包括循环张量网络。

表 1. 近似完全张量网络收缩的复杂度比较

方法	设置	Time	空间	草图大小 (m)
[DGGR02]	无环	$O((pm + mqN) \log 1/\delta)$	$O(mp$对数 $1/\delta)$	$\Omega(3^t/\epsilon^2)$
[HNGN24]	无环	$O((pm$ 对数 $m + qN)$ 对数 $1/\delta)$	$O(mp$对数 $1/\delta)$	$\Omega(3^t/\epsilon^2)$
Ours	无环	$O((pm \log m + qN) \log 1/\delta)$	$O(mp \log 1/\delta)$	$\Omega(t/\epsilon^2)$
Ours	一般	$O((pm \log m + qN) \log 1/\delta)$	$O(mp$对数 $1/\delta)$	$\Omega(3^t/\epsilon^2)$

2 预备知识

本节介绍全文中使用的符号和概念。首先，我们基于 [GVL13] 回顾张量代数中的概念及其符号。

一个阶数为q 的张量 $\mathcal{A} \in \mathbb{R}^{n_1 \times \cdots \times n_q}$ 可以看作一个 q维数组，其中第 k个模态的指标范围为 1 到 n_k。阶数为q 的张量的一个分量 $\mathcal{A}(i_1, \ldots, i_q)$由 q 个指标指定，其中 $1 \le i_k \le n_k$，对所有从 1 到 q的 k均成立。对于多重索引$\mathbf{i} = (i_1, \ldots, i_q)$，张量分量也可以等价地用 $\mathcal{A}(\mathbf{i}) \equiv \mathcal{A}(i_1, \ldots, i_q)$ 编号。长度为 q的全 1 多重索引记作 $\mathbf{1}_q$；当其长度可由上下文明确时，也可简记为$\mathbf{1}$ 。若$\mathbf{i}$ 和$\mathbf{j}$ 是长度相同的多重索引，则$\mathbf{i} \le \mathbf{j}$ 表示对于所有 k， $i_k \le j_k$ 。

$\mathcal{A}$ 的平方弗罗贝尼乌斯范数 由$\|\mathcal{A}\|_{\mathrm{F}}^2 = \sum_{\mathbf{i}=\mathbf{1}}^{\mathbf{n}} \mathcal{A}(\mathbf{i})^2 \equiv \sum_{i_1=1}^{n_1} \cdots \sum_{i_q=1}^{n_q} \mathcal{A}(i_1, \ldots, i_q)^2$,给出， $\mathcal{A}$的非零分量个数记为 $\mathrm{nnz}(\mathcal{A})$，而哈达玛积 是张量之间的逐分量乘法。

定义 5 （哈达玛积）。 给定任意两个张量 $\mathcal{A}$、$\mathcal{B} \in \mathbb{R}^{n_1 \times \cdots \times n_q}$，它们的哈达玛积 $\mathcal{A} \circ \mathcal{B}$ 由 $(\mathcal{A} \circ \mathcal{B})(\mathbf{i}) = \mathcal{A}(\mathbf{i})\mathcal{B}(\mathbf{i})$ 规定，对 $\mathbf{1} \le \mathbf{i} \le \mathbf{n}$，其中 $\mathcal{A} \circ \mathcal{B} \in \mathbb{R}^{n_1 \times \cdots \times n_q}$。

除了对张量的分量进行索引之外，还可以提取子张量。从张量中进行的一阶和二阶提取分别称为纤维 和切片。例如，若$\mathcal{A} \in \mathbb{R}^{2 \times 3 \times 4 \times 5}$ 且 $b = \mathcal{A}(1, :, 2, 4)$，则 b是 $\mathcal{A}$ 的一条纤维，且 $b(i) = \mathcal{A}(1, i, 2, 4)$。此外，若$B = \mathcal{A}(1, :, 2, :)$，则 B是 $\mathcal{A}$ 的一个切片，且 $B(i, j) = \mathcal{A}(1, i, 2, j)$。在此记号中，冒号用于表示对某一模态的提取。

张量可以被 以各种方式排列。张量 $\mathcal{A} \in \mathbb{R}^{n_1 \times \cdots \times n_q}$ 的向量化 ， 记作

$\mathrm{vec}(\mathcal{A}) \in \mathbb{R}^{n_1 \cdots n_q}$，是一个列向量，其分量按字典序排列，取自 $\mathcal{A}$。 $\mathcal{A}$ 的模态-k展开 记作 $\mathrm{mat}_k(\mathcal{A}) \in \mathbb{R}^{n_k \times n_1 \cdots n_q / n_k}$，是一个矩阵，其各列对应于 $\mathcal{A}$ 的模态-k 纤维，并按字典序排列。若省略 k ，如 $\mathrm{mat}(\mathcal{A})$，则默认其值为 1。

我们继续讨论专用于一阶和二阶张量的运算，这些张量通常称为向量和矩阵。首先定义两个矩阵之间的克罗内克积 。

定义 6 (克罗内克积)。 给定任意两个矩阵 $A \in \mathbb{R}^{m \times n}$ 和 $B \in \mathbb{R}^{p \times q}$，它们的克罗内克积 $A \otimes B$ 是一个(mp)乘(nq) 矩阵，其分量由$(A \otimes B)(\alpha, \beta) = A(i,j)B(k,l)$给出，其中 $\alpha = p(i-1)+k$和 $\beta = q(j-1)+l$。

将克罗内克积用于向量时，假定这些向量是列矩阵。这意味着，对于向量 x和 y， $x \otimes y = \mathrm{vec}(x \times y)$。用 $x^{\otimes p} \equiv x \otimes \cdots \otimes x$表示向量 x与自身进行 p次克罗内克积。对集合 S 中的每个元素 u 依次进行克罗内克积，记作 $\bigotimes_{u \in S} X_u$。对于空集，有 $\bigotimes_{u \in \emptyset} X_u = 1$。逐行克罗内克积 是将克罗内克积应用于两个矩阵的各行。

定义 7 （逐行克罗内克积）。 给定任意两个矩阵 $A \in \mathbb{R}^{p \times n}$ 和 $B \in \mathbb{R}^{p \times m}$，它们的逐行克罗内克积 $A \bullet B$ 是一个 $p \times (nm)$ 矩阵，其分量定义为 $(A \bullet B(k, \alpha) = A(k,i)B(k,j)$，其中 $\alpha = m(i-1)+j$。

对于向量 x， $y \in \mathbb{C}^n$，它们的内积记作 $\langle x, y \rangle$； x 的逐分量复共轭记作 $\overline{x}$，而 e_i 是第 i个标准基向量。用 $x * y$表示它们的循环卷积，用 $x \star y$ 表示它们的循环互相关。

定义 8 （循环卷积）。 给定两个向量 $x, y \in \mathbb{C}^n$，它们的循环卷积 $x * y$是一个向量，其分量定义为 $(x*y)_j = \sum_{i=1}^{n} x_i y_{(j-i \bmod n)+1}$。

定义 9 (循环互相关)。 给定两个向量 x、 $y \in \mathbb{C}^n$，它们的循环互相关 $x \star y$ 是一个向量，其分量定义为 $(x \star y)_j = \sum_{i=1}^{n} \overline{x_i} y_{(j+i-2 \bmod n)+1}$。

用 I 表示单位矩阵，而矩阵 $F_n \in \mathbb{C}^{n \times n}$（或当上下文已明确时，简称 F ）表示离散傅里叶变换矩阵。使用快速傅里叶变换，可以在 $O(n \log n)$ 时间内高效计算乘积 $F_n x$ 和 $F_n^{-1} x$ 。下面的引理阐明了循环卷积和循环互相关与哈达玛积之间的关系。

Lem引理 1。 对于任意向量 x、 $y \in \mathbb{C}^n$、 $F(x*y) = Fx \circ Fy$，以及 $F(x \star y) = F\overline{x} \circ Fy$ 。

2.1 计数草图

计数草图（见定义 *11*）最初用于在数据流 [CCFC02]中查找频繁项。此后，它作为一种降维技术得到了更广泛的应用 [Woo14]。计数草图矩阵的每一列在随机行中只有一个随机符号。符号由 4-独立哈希函数确定（见定义 10），行则由 2-独立哈希函数确定。如果向量 x 以其非零分量列表的形式表示，那么可以在 $O(\mathrm{nnz}(x))$ 时间内高效计算该向量的计数草图。

定义 10 (k-独立性)。 如果对于任意 k 个不同的元素 $x_1, \ldots, x_k \in [n]$，其哈希值 $h(x_1) \ldots, h(x_k)$ 彼此独立且在 $[m]$ 中均匀分布，则称哈希函数族 $H = \{h \colon [n] \to [m]\}$ 为 k-独立。

定义 11 (计数草图**)。** 计数草图矩阵是一个随机矩阵 $C \in \mathbb{R}^{m\times n}$，其每个分量定义为 $C(j,i) = s(i)[h(i) = j]$，其中 $s\colon [n] \to \{-1,1\}$ 和 $h\colon [n] \to [m]$ 分别取自 *4*-独立和 *2*-独立哈希函数族。给定任意向量 $x \in \mathbb{R}^n$，通过 Cx 得到 x 的计数草图。

引理 2 ([WDL+09, 引理 2])。 对于任意正整数 m, n，令 $C \in \mathbb{R}^{m\times n}$ 为一个计数草图矩阵。那么，对于任意向量 $x, y \in \mathbb{R}^n$，

$$\mathbb{E}\Big[(Cx)^\mathsf{T}(Cy)\Big] = x^\mathsf{T}y \quad and \quad \mathrm{Var}\Big((Cx)^\mathsf{T}(Cy)\Big) \le \frac{2}{m}\|x\|_2^2\|y\|_2^2.$$

2.2 张量草图

张量草图（参见定义 12）率先引入了一种计数草图，使其能够高效地对张量积进行草图化。其思路源自关于压缩矩阵乘法 [Pag13] 的研究，该方法用于近似多项式核 $\langle x,\ y\rangle^q = \langle x^{\otimes q},\ y^{\otimes q}\rangle$（适用于任意向量 x 和 y），以加速核机器 [PP13]。

定义 12 (张量草图**)。 </style}阶数为 的张量草图矩阵 是一个随机矩阵，定义为 ((**$FC_1) \bullet (FC_2) \bullet \cdots \bullet (FC_q))$，其中 $C_1, \ldots, C_q \in \mathbb{R}^{m\times n}$ 是相互独立的计数草图矩阵。给定任意向量 $x \in \mathbb{R}^{n^q}$，x 的张量草图通过 Tx 得到。

借助张量草图矩阵 $T \in \mathbb{R}^{m\times n^q}$和向量 $x \in \mathbb{R}^{n^q}$，可以在 $O(q\,\mathsf{nnz}(x))$ 时间内计算草图 Tx。然而，给定向量 $y \in \mathbb{R}^n$，利用快速傅里叶变换，可以在 $O(qm\log\ m + q\,\mathsf{nnz}(y))$ 时间内计算草图 $T(y^{\otimes q})$。与先计算 $y^{\otimes q}$ 再应用张量草图相比，这种方法的效率可以显著提高。

引理 3 ([ANW14, 引理 2])。 对于任意正整数 m、n、q，设 $T \in \mathbb{R}^{m\times n^q}$ 为一个阶数为 q 的张量草图矩阵。则对于任意向量 x、$y \in \mathbb{R}^{n^q}$，

$$\mathbb{E}\Big[(Tx)^\mathsf{T}(Ty)\Big] = x^\mathsf{T}y \quad and \quad \mathrm{Var}\Big((Tx)^\mathsf{T}(Ty)\Big) \le \frac{3^q}{m}\|x\|_2^2\|y\|_2^2.$$

2.3 递归草图

递归草图 （见定义 13）改进了张量草图的一项局限性：张量草图的方差对张量的阶数呈指数依赖。其核心思想是，当按照二叉树对张量的模进行递归草图处理时，可以避免这种指数依赖。 [AKK+20] 中介绍了多种可用于执行递归的草图。我们在下面的定义中使用张量草图的递归。对于给定的草图大小，递归草图仅使张量草图的时间复杂度增加一个常数因子，同时还能使近似精度获得指数级提升。

定义 13 (递归草图**)。** 阶数为 q 的递归草图矩阵 $R \in \mathbb{R}^{m\times n^q}$ （其中 q 是二的幂）是一个随机矩阵，定义为 $R = Q_2Q_4\cdots Q_{q/2}Q_qS$，其中 $S = C_1 \otimes \cdots \otimes C_q$ 和 $Q_l = T_{l,\ 1} \otimes T_{l,\ 2} \otimes \cdots \otimes T_{l,\ l/2}$，包含相互独立的计数草图矩阵 C_1、……、$C_q \in \mathbb{R}^{m\times n}$，以及相互独立的张量草图矩阵 $T_{l,\ 1}$、……、$T_{l,\ l/2} \in \mathbb{R}^{m\times m^2}$，其中 $l \in \{2, 4, \ldots, q/2, q\}$。给定任意向量 $x \in \mathbb{R}^{n^q}$，x 的递归草图通过 Rx 得到。

[2]为简化记号，我们允许 q 取任意正整数。实际阶数默认为下一个二的幂 $\tilde{q}$，并按照 [AKK+20]中的描述，通过与 e_1进行克罗内克积，将 $\mathbb{R}^{n^q}$ 中的张量嵌入到 $\mathbb{R}^{n^q}$ 中。重要的是，这样做既能保持方差界，也不会影响渐近复杂度。

引理 4 ([AKK+20，引理 12 和 15])。对于任意正整数 m、n、q，令 $R \in \mathbb{R}^{m \times n^q}$ 为一个阶数为 q 的递归草图矩阵。那么，对于任意向量 x、$y \in \mathbb{R}^{n^q}$，

$$\mathbb{E}\Big[(Rx)^\mathsf{T}(Ry)\Big] = x^\mathsf{T} y \quad and \quad \mathrm{Var}\Big((Rx)^\mathsf{T}(Ry)\Big) \leq \left(\left(1+\frac{8}{m}\right)^{2q} - 1\right)\|x\|_2^2\|y\|_2^2.$$

3 相关工作

In t本节回顾相关工作，并将其与我们的贡献进行比较 s.

基于草图的张量网络收缩。文献 [DGGR02] 提出了一种 AMS 草图的推广方法，用于估计多重连接查询的连接大小。尽管 [DGGR02]，中并未明确提及，连接大小估计问题等价于完全张量网络收缩的近似问题（见第 5.1 节）。他们的方法通过 q 个独立的 AMS 草图矩阵对一个阶数为 q 的张量进行草图化。对所有张量完成草图化后，通过对这些草图求哈达玛积，再对其分量取平均，即可得到近似结果。由于 AMS 草图矩阵是稠密的，计算一个阶数为 q 的张量 $\mathcal{X}$ 的 m 维草图需要 $O(mq\,\mathsf{nnz}(\mathcal{X}))$ 时间。文献 [HNGN24] 提出的方法通过使用稀疏计数草图矩阵对张量进行草图化，将这一开销降低了 m 倍。这两种方法都仅支持无环张量网络，且其空间和时间复杂度会随收缩次数呈指数增长（与我们的方法的比较见表 1）。

边缘化。Robeva 和 Seigal [RS19] 证明了，离散变量无向图模型中的边缘化等价于张量网络收缩（TNC）。因此，经典图模型推理算法可以视为张量网络收缩算法。这包括精确方法，例如无环张量网络上的置信传播，以及用于有环张量网络的连接树算法 [WJ+08]。正如 [WJ+08]和 [MS 08], 所示，此类精确算法的运行时间关于张量网络的树宽呈指数增长。为缓解这一复杂性，实践中通常采用环状置信传播等近似方法 [WJ+08]，但这类方法通常缺乏最坏情况下的近似保证。

张量分解。张量分解方法旨在将给定张量分解为张量网络，具体网络结构由所选分解方式（例如 CP、Tucker 或张量列车）决定。这与本文研究的张量网络收缩问题不同：该问题中张量网络已给定，目标是高效计算其收缩。大量研究探讨了如何利用随机化算法和草图化加速张量分解。例如，[WTSA15] 提出了用于 CP 分解的算法，通过使用草图加速张量幂迭代。在 [MB18] 和 [MS21] 中，研究者提出了用于低秩 Tucker 分解的草图化交替最小二乘法。最近，[MWZ24] 提出了用于张量列车、Tucker 和 CP 分解，以及分解为任意张量网络的算法。虽然这些工作证明了草图化在张量分解中的有效性，但它们并未解决本文研究的问题。

4 方法

我们提出两种方法，它们是 (ϵ, δ)-近似张量网络收缩。我们首先描述针对完全张量网络收缩的方法，随后证明这些方法也能近似部分张量网络收缩，其结果为非零阶张量。

在确立四项简化假设（不失一般性）之后，我们在本节开头描述第一种方法，该方法可以近似任意张量网络收缩。随后，我们通过证明当前最佳方法的方差对收缩次数具有指数依赖，来说明需要一种新的草图技术。受此启发，我们给出无环完全张量网络收缩的递归表达式，以此作为引入第二种方法的基础，该方法对收缩次数具有多项式依赖。

4.1 不失一般性假设

为简化后续的解释和符号表示，我们对张量网络收缩（TNC）作出以下四个不失一般性（WLOG）的假设:

不失一般性 1. 每个模态恰好参与一次收缩。否则，可以为参与多次收缩的模态创建虚拟副本 。考虑以下向量 a，b，c的张量网络收缩：tc($a \times b \times c$，{(1，2)，(2，3)})，其中 b 的唯一模态参与了两次收缩。定义$B(i, j) = b(i)[i = j]$，则可以改写为 tc($a \times B \times c$，{(1，2)，(3，4)})。

不失一般性 2. 每次收缩涉及两个张量。当被收缩的模态属于同一个张量时（例如矩阵的迹），精确收缩的计算时间与该张量的非零分量数量成正比，因此可以直接计算。

不失一般性 3. 任意两个张量之间至多存在一次收缩。多次收缩可通过重塑张量合并为一次更大的收缩。以下矩阵A 与 B的示例：tc($A \times B$, {(1, 3), (2,4)})可表示为 tc(vec(A) $\times$ vec(B) {(1,2)})。

不失一般性 4. 每个模态具有相同数量的分量。本文讨论的方法依赖于张量的非零分量数量；因此，可以向较小的张量追加零，以确保所有模具有相同数量的分量。

容易验证，这些变换不会增加变换后张量的非零分量数量或弗罗贝尼乌斯范数。此外，它们不会引入任何额外的张量或收缩。这些是我们的结果成立的充分条件。

4.2 近似任意张量网络收缩

我们第一种方法的洞见源于对 [HNGN24]中方法的一项观察；即，当张量网络收缩为无环时，使用循环互相关来组合草图会产生一种模式，使得每次收缩都有一个共轭模态。这种模式的产生是因为循环互相关在表示为哈达玛积时会对其第一个参数取共轭（见引理1），从而形成一层层复共轭，并在其数量为偶数时相互抵消。因此，收缩树中每隔一层的张量，其模被取共轭，从而确保每次收缩都有一个模态被共轭。然而，这依赖于张量网络具有树结构，因此对有环张量网络并不适用。

下面给出的两个简单示例分别展示了无环和有环张量网络收缩（见图2）：在无环张量网络收缩（例2）中，每次收缩都有一个共轭模态（两者都属于张量 Y）；而在有环张量网络收缩（例3）中，收缩 (1，6) 没有共轭模态。我们强调这一区别，是因为我们针对一般情形的证明（见附录A.2）依赖于每次收缩都有一个共轭模态。

例2. 采用 *[HNGN24]*, 中的方法，具有 x、$z \in \mathbb{R}^n$ 和 $Y \in \mathbb{R}^{n \times n}$ 的无环张量网络收缩 $\mathrm{tc}(x \times Y \times z, \{(1, 2), (3, 4)\})$（如图*2*左侧所示）可近似如下。首先，收缩

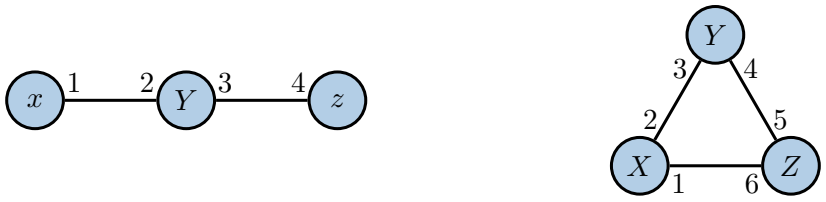


图2. 简单无环（左）和循环（右）张量网络的张量图

(1, 2) 和 (3,4) 分别被分配独立的计数草图矩阵 C_1 和 C_2。随后计算计数草图 C_1x 和 C_2z 以及张量草图 $F^{-1}(FC_1 \bullet FC_2)\,\mathrm{vec}(Y)$，并通过循环互相关将它们组合起来。最后，选取结果向量的第一个分量，得到以下表达式：

$$e_1^\mathsf{T}(C_1x \star F^{-1}(FC_1 \bullet FC_2)\,\mathrm{vec}(Y) \star C_2z) = e_1^\mathsf{T}F^{-1}(FC_1x \circ \overline{(FC_1 \bullet FC_2)\,\mathrm{vec}(Y)} \circ FC_2z).$$

例*3.* 使用 *[HNGN24],* 中的方法，循环张量网络收缩 *tc(*$X \times Y \times Z$*, {(2, 3), (4, 5), (1, 6)})*（如图*2*右侧所示），其中 $X, Y, Z \in \mathbb{R}^{n\times n}$，可按如下方式近似。首先，为收缩 *(2, 3)*、*(4, 5)* 和 *(1, 6)* 分别分配独立的计数草图矩阵 C_1、C_2 和 C_3。然后计算张量草图 F^{-1}*(*$FC_1 \bullet FC_2$*) vec(X)*、F^{-1}*(*$FC_2 \bullet FC_3$*) vec(Y)* 和 F^{-1}*(*$FC_3 \bullet FC_1$*) vec(Z)*，并通过循环互相关将它们合并。最后，选取所得向量的第一个分量，得到以下表达式：

$$\begin{aligned} &e_1^\mathsf{T}(F^{-1}(FC_1 \bullet FC_2)\,\mathrm{vec}(X) \star F^{-1}(FC_2 \bullet FC_3)\,\mathrm{vec}(Y) \star F^{-1}(FC_3 \bullet FC_1)\,\mathrm{vec}(Z)) \\ &= e_1^\mathsf{T}F^{-1}((FC_1 \bullet FC_2)\,\mathrm{vec}(X) \circ \overline{(FC_2 \bullet FC_3)\,\mathrm{vec}(Y)} \circ (FC_3 \bullet FC_1)\,\mathrm{vec}(Z)). \end{aligned}$$

为介绍我们估算循环张量网络收缩的方法，我们首先引入补计数草图，这是我们方法的关键。

定义14 (补计数草图)。 给定一个计数草图矩阵 $C \in \mathbb{R}^{m\times n}$ 及哈希函数 $s: [n] \to \{-1, 1\}$ 和 $h: [n] \to [m]$, C 的补计数草图矩阵 C' 定义为 $C'(j, i) = s(i)[-h(i) + 2 \equiv j$ (模 $m)]$。

注释 1。 计数草图矩阵的补矩阵本身也是一个计数草图矩阵，而补矩阵的补矩阵则会还原为原矩阵，即 $(C')' = C$。

补计数草图是计数草图的循环反转形式，这意味着对于任意 $x \in \mathbb{R}^n$，都有 $FC'x = FCx$，如引理 5 所示（证明见附录A.1）。无需依靠循环互相关来对每次收缩中的一个模态进行共轭，使用补计数草图矩阵即可显式控制每次收缩中哪个模态进行共轭。

引理 5。 对于任意正整数 m 和 n，设 $C \in \mathbb{R}^{m\times n}$ 为一个计数草图矩阵，C' 为其补矩阵。那么，对于任意 $x \in \mathbb{R}^n$、$FC'x = FCx$。

我们的方法首先为每个收缩的一个模态独立采样计数草图矩阵，然后将其补计数草图矩阵分配给另一个模态。具体而言，对于每个 $(u, v) \in E$, C_u 是一个独立的计数草图矩阵，且 $C_v := C'_u$。因此，任意完全张量网络收缩的近似可表示为

$$\hat{\mathrm{tc}}(\mathcal{X}_1 \times \cdots \times \mathcal{X}_p, E) = e_1^\mathsf{T}F^{-1}(x_1 \circ \cdots \circ x_p), \qquad x_k = \left(\bullet_{u\in V(k)} FC_u\right)\mathrm{vec}(\mathcal{X}_k), \tag{1}$$

其中， $V(k)$ 是 $\mathcal{X}_k$ 的模态，而 $\bullet_{u\in S}$ 表示 S 中每个成员 u 对应的逐行克罗内克积。下面的例子表明，采用我们的方法后，例3 中所用简单循环张量网络收缩的每次收缩现在都有一个共轭模态。引理 6 中估计量的期望和方差界的证明（见附录 A.2）依赖于这一性质。

例4. 使用我们的第一种方法，循环张量网络收缩$\mathrm{tc}(X \times Y \times Z, \{(2,3), (4, 5),(1, 6)\})$ （如图*2*右侧所示）在 X、 Y、 $Z \in \mathbb{R}^{n\times n}$ 下的近似如下。首先，分别为模*2*、*4*和*1*分配独立的计数草图矩阵 C_2、C_4和 C_1。然后，分别为模*3*、*5*和*6*分配补计数草图矩阵 C_2'、 C_4'和 C_1'。随后计算张量草图 $F^{-1}($ $FC_1 \bullet FC_2) \operatorname{vec}(X)$、 $F^{-1}(FC_2' \bullet FC_4) \operatorname{vec}(Y)$和 $F^{-1}(FC_4' \bullet FC_1') \operatorname{vec}(Z)$，并使用循环卷积将它们组合。最后，选取所得向量的第一个分量，得到如下表达式:

$$
\begin{aligned}
&e_1^{\mathsf{T}}(F^{-1}(FC_1 \bullet FC_2) \operatorname{vec}(X) * F^{-1}(FC_2' \bullet FC_4) \operatorname{vec}(Y) * F^{-1}(FC_4' \bullet FC_1') \operatorname{vec}(Z)) \\
&= e_1^{\mathsf{T}} F^{-1}((FC_1 \bullet FC_2) \operatorname{vec}(X) \circ (\overline{FC_2} \bullet FC_4) \operatorname{vec}(Y) \circ (\overline{FC_4} \bullet \overline{FC_1}) \operatorname{vec}(Z)).
\end{aligned}
$$

我们现在证明，对于任意张量网络收缩，我们的第一种方法都是无偏估计量，并通过将其归约到引理 6 中的张量草图来界定其方差（证明见附录 A.2）。据我们所知，虽然我们的第一种方法与[HNGN24], 中的方法具有相同的方差界和渐近复杂度，但我们的方法是首个能够估计循环张量网络收缩的草图方法。

引理 6. 对于每个 m， $p \in \mathbb{N}^+$ 以及任意完全张量网络收缩，令 $\hat{}\mathrm{tc}$ 按照式 (1) 定义，草图大小为m。则，当 $y = \hat{}\mathrm{tc}(\mathcal{X}_1 \times \cdots \times \mathcal{X}_p, E)$ 且 $t = |E|$时，

$$
\mathbb{E}[y] = \mathrm{tc}(\mathcal{X}_1 \times \cdots \times \mathcal{X}_p, E), \qquad \textit{and} \qquad \operatorname{Var}(y) \le \frac{3^t}{m} \prod_{k=1}^{p} \|\mathcal{X}_k\|_{\mathrm{F}}^2.
$$

我们第一种方法的伪代码见算法 1。我们的估计量支持最一般的流式数据设置——一般转门模型，在该模型中，更新流可以任意改变任何分量。具体而言，我们存储所有 $k \in [p]$ 对应的草图 x_k（见算法 1 第 8 行），每次更新的处理时间为 $O(q_k)$。

算法 1 近似一般的完整张量网络收缩。

输入：对于 $k \in [p]$, 度数为 q_k 的张量 $\mathcal{X}_k \in \mathbb{R}^{n\times\cdots\times n}$，收缩次数 E，草图大小 m输出：近似张量网络收缩 $y \in \mathbb{R}$1：对于 (u, v), 执行 $\in E$ 2：令 $C_u \in \mathbb{R}^{m\times n}$ 为一个独立的计数草图矩阵。3：$C_v \leftarrow C_u'$4：结束循环 5：$z \in \{1\}^m$6：对于 $k \in [p]$ 执行 7：$T_k \leftarrow F^{-1}(\bullet_{u\in V(k)} FC_u)$ ▷ 由计数草图矩阵组成一个张量草图矩阵 8：$x_k \leftarrow T_k \mathrm{vec}(\mathcal{X}_k)$ 9：$z \leftarrow z \circ Fx_k$10：结束循环 11：$y \leftarrow \frac{1}{m}\sum_{i=1}^{m} z_i$

—

4.3 指数下界

使用草图方法进行近似完全张量网络收缩的已知最佳界限，包括 [DGGR02], [HNGN24], 中提出的方法以及我们在第 4.2 节中的第一种方法，都对收缩次数呈指数依赖，如表 1 所示。为了说明为何需要一种新方法，我们在引理 7 中（证明见附录 A.3）声称， [HNGN24] 中的方法需要对收缩次数呈指数依赖。具体而言，我们基于 [AKK+20 给出了近似一系列矩阵乘法时方差的一个下界；引理 43], 为使用张量草图进行范数近似提供了方差下界。

引理 7。 对于任意正整数 n、m、q，其中 q 为偶数，设 C_1、...、$C_q \in \mathbb{R}^{m\times n}$ 是相互独立的计数草图矩阵，且 $T_i = F^{-1}((FC_{i-1}) \bullet (FC_i))$。那么，对于全 *1* 张量 $\mathcal{X}_1$、$\mathcal{X}_{q+1} \in \{1\}^n$ 和 $\mathcal{X}_i \in \{1\}^{n\times n}$ （其中 $2 \le i \le q$），

$$\mathrm{Var}\Big(e_1^\mathsf{T}(C_1\mathcal{X}_1 \star T_2\,\mathrm{vec}(\mathcal{X}_2) \star \cdots \star T_q\,\mathrm{vec}(\mathcal{X}_q) \star C_q\mathcal{X}_{q+1})\Big) \ge \left(\frac{3^q}{2m^2} - 1\right)\prod_{i=1}^{q+1}\|\mathcal{X}_i\|_\mathrm{F}^2.$$

这一结果也适用于我们的第一种方法，表明仅靠更精细的分析无法消除以往最佳方法中的指数依赖。因此，我们得出结论：要为近似张量网络收缩问题获得更优的界，必须采用新方法。本节余下部分将介绍我们的第二种近似方法，并说明这种新方法如何消除对收缩次数的指数依赖。

4.4 无环张量网络收缩的改进近似

对于第二种方法，除了假设进行完全张量网络收缩外，我们还假设张量网络是无环的。无环张量网络常见于若干重要场景。特别是，许多关系数据库查询都具有无环连接图，而估计这些查询的连接大小（这对应于张量网络收缩）对构建高效的数据库系统至关重要（见第 5.1 节）。尽管已有针对无环张量网络收缩的易处理算法，中间张量仍可能大到难以处理，因此有必要采用近似张量网络收缩。

为便于解释我们提出的解决方案，我们首先介绍一些额外的符号和概念。具体而言，我们引入根张量 $\mathcal{X}_o$的概念，其中 $o \in [p]$从参与收缩的张量中任意选取一个，以便从张量网络构成有根树结构。随后，我们将 $\Gamma(k)$ 定义为张量 $\mathcal{X}_k$ 的出邻居[3]，即与 $\mathcal{X}_k$ 进行收缩且距离根张量更远的张量集合。注意，有根树的所有叶节点都是向量。基于此，我们可以递归地表示无环完全张量网络收缩，如引理 8 所示（证明见附录 A.4）。

Le引理 8。 给定任意 $o \in [p]$，任何无环完全张量网络收缩都可以递归地表示为 as

$$\mathrm{tc}(\mathcal{X}_1 \times \cdots \times \mathcal{X}_p, E) = \mathrm{vec}(\mathcal{X}_o)^\mathsf{T} g(o), \qquad g(k) = \bigotimes_{l\in\Gamma(k)} \mathrm{mat}(\mathcal{X}_l)g(l).$$

我们用于近似张量网络收缩的第二种方法基于以下观察：张量网络收缩的递归表述由一系列涉及克罗内克积的矩阵乘法组成。

[3]为简化符号，假设邻居函数 Γ 按照与相应张量的模相匹配的顺序返回邻居，并假设非根张量的第一个模与其在树中的父节点进行收缩。通过简单地置换张量的指标，始终可以实现这种安排。

利用这一观察，我们使用递归草图矩阵对克罗内克积进行草图化。这便得到如下张量网络收缩近似：

$$\hat{\mathrm{tc}}(\mathcal{X}_1 \times \cdots \times \mathcal{X}_p, E) = \mathrm{vec}(\mathcal{X}_o)^\mathsf{T} R_o^\mathsf{T} R_o \hat{g}(o), \qquad \hat{g}(k) = \bigotimes_{l \in \Gamma(k)} \mathrm{mat}(\mathcal{X}_l) R_l^\mathsf{T} R_l \hat{g}(l), \tag{2}$$

其中，每个 R_k 将向量化的 $q_k - 1$ 阶张量映射为大小为 m 的向量，但 R_o 除外，它将向量化的 q_o 阶张量映射为大小为 m 的向量。

在引理 9 中，我们证明第二种方法是无环完全张量网络收缩的无偏估计量，并给出其方差界。证明见附录 A.5：沿有根树遍历时，将引理 4 中递归草图方差界的 $(1+8/m)^p$ 项累加起来，即可得到该方差界。期望结果则由递归草图的期望得出。

引理 9. 对于每个 m、$p \in \mathbb{N}^+$ 以及任意无环完全张量网络收缩，设 $\hat{\mathrm{tc}}$ 按照式 (2) 定义，草图大小为 m。那么，对于 $y = \hat{\mathrm{tc}}(\mathcal{X}_1 \times \cdots \times \mathcal{X}_p, E)$ 和 $t = |E|$，

$$\mathbb{E}[y] = \mathrm{tc}(\mathcal{X}_1 \times \cdots \times \mathcal{X}_p, E) \qquad \textit{and} \qquad \mathrm{Var}(y) \le \left(\left(1 + \frac{8}{m}\right)^{2t} - 1\right) \prod_{k=1}^{p} \|\mathcal{X}_k\|_\mathrm{F}^2.$$

高效地计算估计值并非易事，尤其是因为每个 $\mathrm{mat}(\mathcal{X}_l) R_l^\mathsf{T}$ 都会产生一个（可能很大的） $n \times m$ 矩阵。为了解决这一问题，我们对前一个递归步骤中的递归草图矩阵进行分解，并将其初始降维直接整合到递归的下一步中。由于 $R_k = Q_k S_k$，其中 $S_k = \bigotimes_{l \in \Gamma(k)} C_{k,l}$ 是相互独立的计数草图矩阵的克罗内克积，因此我们可以利用混合乘积恒等式 $(A \otimes B)(x \otimes y) = (Ax) \otimes (By)$ 得到以下因式分解：

$$\mathrm{vec}(\mathcal{X}_o)^\mathsf{T} R_o^\mathsf{T} Q_o S_o \left(\bigotimes_{l \in \Gamma(o)} \mathrm{mat}(\mathcal{X}_l) R_l^\mathsf{T} R_l \hat{g}(l) \right) = \mathrm{vec}(\mathcal{X}_o)^\mathsf{T} R_o^\mathsf{T} Q_o \left(\bigotimes_{l \in \Gamma(o)} C_{o,l} \, \mathrm{mat}(\mathcal{X}_l) R_l^\mathsf{T} R_l \hat{g}(l) \right),$$

该方法可应用于每个递归步骤。由此得到 $C_{k,l}\mathrm{mat}(\mathcal{X}_l) R_l^\mathsf{T}$，它们是规模易于处理的 m-by-m 矩阵，可在 $O(q_l \mathrm{nnz}(\mathcal{X}_l))$ 时间内高效计算。因此，我们将一个涉及任意阶张量的问题转化为一个易于处理得多、只涉及二阶张量的问题。

为彻底避免存储中间矩阵，我们提出一种渐进式计算方案，按照从叶节点到根节点的顺序遍历张量网络的有根树。该方法在生成每个后续草图的过程中高效计算矩阵-向量乘积，从而在每个递归步骤中得到紧凑的草图向量。

具体而言，在叶节点处，我们将一阶张量的计数草图计算为 $x_l = C_{k,l}\mathrm{vec}(\mathcal{X}_l)$，其耗时为 $O(\mathrm{nnz}(\mathcal{X}_l))$。利用这些叶节点层级的草图向量，我们接着沿树向上一层。在这一层，我们首先计算 $r_k = Q_k(\bigotimes_{l \in \Gamma(k)} x_l)$，它表示对叶节点层级草图向量的递归草图。然后我们计算 $x_k = C_{l,k}\,\mathrm{mat}(\mathcal{X}_k) R_k^\mathsf{T} r_k$，将草图矩阵 $C_{l,k}\mathrm{mat}(\mathcal{X}_k) R_k^\mathsf{T}$ 的计算与涉及 r_k 的矩阵-向量乘积结合起来。联合执行这些操作需要 $O(q_k \mathrm{nnz}(\mathcal{X}_k))$ 的时间；若分别执行，则需要 $O(q_k\,\mathrm{nnz}(\mathcal{X}_k) + m^2)$ 的时间。该过程（称为 Sketched-MatrixVectorProduct）的伪代码见附录 B 中的算法 5，而我们第二种方法的伪代码见算法 2。例 5 展示了将第二种方法应用于图 3 所示无环张量网络的情形。

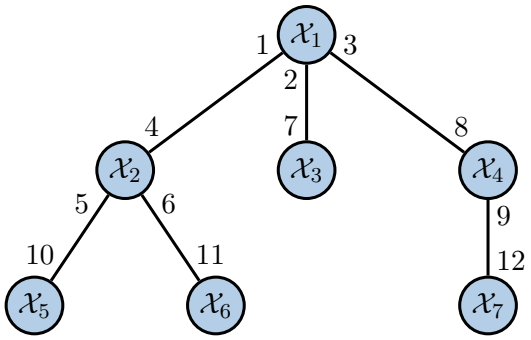


图 3. 一个无环张量网络示例的张量图

例 5. 使用我们的第二种方法，图 *3* 所示的张量网络收缩，

$$\mathrm{tc}(\mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{X}_3 \times \mathcal{X}_4 \times \mathcal{X}_5 \times \mathcal{X}_6 \times \mathcal{X}_7, \{(1,4),(2,7),(3,8),(5,10),(6,11),(9,12)\})$$

对于 $\mathcal{X}_1$、$\mathcal{X}_2 \in \mathbb{R}^{n\times n\times n}$、$\mathcal{X}_4 \in \mathbb{R}^{n\times n}$ 和 $\mathcal{X}_3$、$\mathcal{X}_5$、$\mathcal{X}_6$、$\mathcal{X}_7 \in \mathbb{R}^{n}$，其近似计算方法如下。首先，选取 $\mathcal{X}_1$ 作为任意根张量，从而可以将该张量网络视为一棵有根树。然后，为每个张量分配一个独立的递归草图矩阵，其阶数与该张量在有根树中的子节点数相匹配。因此，我们有 $R_1 \in \mathbb{R}^{m\times n^3}$、$R_2 \in \mathbb{R}^{m\times n^2}$ 和 $R_4 \in \mathbb{R}^{m\times n}$。回顾一下，每个递归草图矩阵都可以分解为 $R_k = Q_k(\bigotimes_{l\in\Gamma(k)} C_{k,l})$，其中 C_k、l 是独立的计数草图矩阵。现在，从树的最低层开始，计算计数草图 $x_5 = C_{2,5}\,\mathrm{vec}(\mathcal{X}_5)$、$x_6 = C_{2,6}\,\mathrm{vec}(\mathcal{X}_6)$、$x_7 = C_{4,7}\,\mathrm{vec}(\mathcal{X}_7)$。接着向上移动一层，先计算递归草图 $r_2 = Q_2(x_5 \otimes x_6)$ 和 $r_4 = Q_4 x_7$，然后使用算法 *5* 计算 $x_2 = C_{1,2}\,\mathrm{mat}(\mathcal{X}_2)R_2^\mathsf{T} r_2$，$x_3 = C_{1,3}\,\mathrm{vec}(\mathcal{X}_3)$，以及 $x_4 = C_{1,4}\,\mathrm{mat}(\mathcal{X}_4)R_4^\mathsf{T} r_4$。最后，在树的根节点处，计算 $x_1 = R_1\,\mathrm{vec}(\mathcal{X}_1)$、$r_1 = Q_1(x_2 \otimes x_3 \otimes x_4)$ 和 $y = x_1^\mathsf{T} r_1$。

算法 2 近似无环全张量网络收缩。

输入： 度数为 q_k 的张量 $\mathcal{X}_k \in \mathbb{R}^{n\times\cdots\times n}$，适用于 $k \in [p]$，收缩 E，草图大小 m

输出： 近似张量网络收缩 $y \in \mathbb{R}$

1: 令 o 为 $[p]$的任意成员。2: 令 $R_o \in \mathbb{R}^{m\times n^{q_o}}$ 为一个独立的递归草图。3: 对于 $k \in [p]\setminus\{o\}$ 执行 4: 令 $R_k = Q_k(\bigotimes_{l\in\Gamma(k)} C_{k,l}) \in \mathbb{R}^{m\times n^{q_k-1}}$ 为一个独立的递归草图。5: 结束循环 6: 函数 NeighborsSketch(k) 7: 如果 $|\Gamma(k)| = 0$ 则返回 1 8: 对于 $l \in \Gamma(k)$执行 9: $r_l \leftarrow$ NeighborsSketch(l) 10: $x_l \leftarrow$ SketchedMatrixVectorProduct($C_{k,l}, \mathcal{X}_l, R_l, r_l$)▷ $C_{k,l}\,\mathrm{mat}(\mathcal{X}_l)R_l^\mathsf{T} r_l$ 11: 结束循环 12: 返回 $Q_k(\bigotimes_{l\in\Gamma(k)} x_l)$▷ 应用递归草图的剩余部分 13: 结束函数 14: $x_o \leftarrow R_o\mathrm{vec}(\mathcal{X}_o)$ 15: $r_o \leftarrow$ NeighborsSketch(o) 16: $y \leftarrow x_o^\mathsf{T} r_o$

4.5 主要结果

在本节中，我们结合前几节的结果，得到定理 1 中所述的主要结果。该定理表明，对于任意完全张量网络收缩，我们的第一种方法是一个$(\epsilon,\ \delta)$-近似张量网络收缩；对于无环完全张量网络收缩，我们的第二种方法是一个$(\epsilon,\ \delta)$-近似张量网络收缩。证明（见附录 A.6）使用切比雪夫不等式，以及引理 6 和引理 9 中的期望值和方差上界，从而获得具有常数成功概率的估计量。通过选取 $O(\log 1/\delta)$ 个相互独立的估计值的中位数（每个估计值都具有常数成功概率），我们可以使用切尔诺夫界，得到一个以 $1-\delta$ 的概率成功的估计量。这种方法有时被称为中位数技巧。它通常用于将具有常数成功概率的估计量提升为具有高成功概率的估计量。

定理 *1.* 对于每一个 $p \in \mathbb{N}^+$、任意 q_k 阶张量 $\mathcal{X}_k$ （针对 $k \in [p]$），以及每一个 ϵ、$\delta > 0$，都存在完全张量网络收缩的一个 (ϵ, δ)-近似张量网络收缩，可在时间 $O((pm\log m+qN)\log 1/\delta)$ 内、使用 $O(mp\log 1/\delta)$ 空间计算得到，其中 *(1)* 在无环情形下 $m=\Omega(t/\epsilon^2)$，*(2)* 在一般情况下 $m=\Omega(3^t/\epsilon^2)$，这里 $N=\sum_{k=1}^{p} nnz(\mathcal{X}_k)$、$q=\sum_{k=1}^{p} q_k$ 和 $t=|E|$。

最后，在推论 1（证明见附录 A.7）中，我们表明我们的两种方法也可应用于部分张量网络收缩，即那些求值为非零阶张量的收缩。这一推广的时间复杂度随所得张量中的分量数目线性增长。有趣的是，我们方法的计算复杂度仅取决于输入张量与输出张量的大小。

推论 *1.* 对于任意 p、$n \in \mathbb{N}^+$，任意阶数为 q_k 的张量 $\mathcal{X}_k \in \mathbb{R}^{n\times\cdots\times n}$ （对应 $k \in [p]$），以及任意 ϵ、$\delta > 0$，都存在一个部分张量网络收缩的 (ϵ, δ)-近似张量网络收缩，其结果为一个阶数为 $\bar{q}$ 的张量，可在 $O((n^{\bar{q}}pm\log m+qN)\log 1/\delta)$ 时间内计算，并使用 $O(n^{\bar{q}}mp\log 1/\delta)$ 空间，其中 *(1)* 在无环情形下 $m=\Omega(t/\epsilon^2)$，*(2)* 在一般情形下 $m=\Omega(3^t/\epsilon^2)$，其中 $N=\sum_{k=1}^{p} nnz(\mathcal{X}_k)$、$q=\sum_{k=1}^{p} q_k$ 和 $t=|E|$。

至此，我们已经证明，第一种方法在一般情形下是近似张量网络收缩，而第二种方法在无环情形下是近似张量网络收缩。在本文余下部分，我们将重点介绍允许近似解的张量网络收缩应用。

5 张量网络收缩的应用

张量网络收缩凭借其一般性，在许多学科中都很重要。从对物理现象的数学建模，到解决计算机科学中的理论和实践问题，皆是如此。在量子力学中，张量用于建模多体量子波函数，而近似张量网络收缩则用于模拟量子计算机 [BC17]。在机器学习中，对高效张量表示进行收缩，旨在降低训练大型模型的计算成本 [Ose11, NPOV15]。与此同时，在概率论中，离散随机变量的联合概率分布可自然地表示为张量，而图模型中的边缘化已被证明等价于张量网络收缩 [RS19]。在图论中，一个图的 q-着色数量可以归约为张量网络收缩 [BC17]。本节余下部分将重点介绍张量网络收缩的两项应用：数据库系统中的连接大小估计和图中的三角形计数。

5.1 连接大小估计

近似张量网络收缩的一项尤为相关的应用位于数据库系统中，即连接大小估计。在该设定下，连接大小是由一次连接所产生的元组数量。

一般而言，查询的执行时间对连接执行顺序 [LGM+15]高度敏感。由于确定最优连接顺序已知是NP难的 [IK84], ，查询优化器通常依赖中间连接大小估计[4] 作为其主要输入，以评估特定连接顺序的代价。

```
SELECT COUNT(*)
FROM R1, R2, R3, R4
WHERE R1.1 =R2.2
AND R3.4 =R2.2 AND
R4.5 =R2.3
```

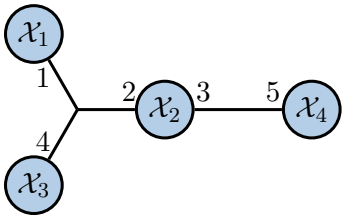


图4。 SQL 查询示例（左）及其对应的张量网络图（右）。对于所有$k \in [4]$，每个分量 $\mathcal{X}_k$**(i)** 表示关系 R_k中**i** 的频数。

包含等值连接的查询的连接大小可以表示为张量网络收缩，其中每个关系由一个张量表示，而连接则指定了收缩 [AGMS99, DGGR02, HNGN24]。因此，连接大小估计可以表述为一个近似张量网络收缩问题。由此，我们的结果为无环查询情况下的连接大小估计提供了更优界，并将先前结果推广到有环查询，如推论2所述。图4展示了一个 SQL 查询示例及其对应的张量网络图。

推论2。 对于每个 $p \in \mathbb{N}^+$，任意由关系 R_k 构成、满足 $k \in [p]$的等值连接查询，以及每个$\epsilon, \delta > 0$，都存在一个(ϵ, δ)-近似的连接大小（如定义 *4* 所示），其计算时间为 $O((pm \log m + qN) \log 1/\delta)$，空间为 $O(mp \log 1/\delta)$；当查询为无环时，复杂度为**(1)** $m = \Omega(t/\epsilon^2))$，一般情况下为**(2)** $m = \Omega(3^t/\epsilon^2)$，其中 $N = \sum_{k=1}^{p} |R_k|$, $q = \sum_{k=1}^{p} q_k$和 $t = |E|$。

证明。 令 $\mathcal{X}_k$ 为关系 R_k的频数张量，使得对于所有 $k \in [p]$，每个分量$\mathcal{X}_k$**(i)** 表示**i** 在关系 R_k中的频数，即 $\mathcal{X}_k$**(i)** $= \sum_{\mathbf{j} \in R_k} [\mathbf{i} = \mathbf{j}]$；并令 E 为等值连接的集合。由于草图具有线性，每个频数张量的草图都可以通过对关系进行单遍扫描来计算。结果可直接由定理 1 得出。 □

5.2 三角形计数

即使对于已知易处理算法的组合问题，近似张量网络收缩也可能很有用。例如，可以考虑图中的三角形计数任务。随着涉及大规模图的应用不断涌现，即使是多项式时间算法也可能变得成本高昂。由于三角形数量可以表示为 tr(AAA)，其中邻接矩阵为 A，因此也可以使用近似张量网络收缩对其进行近似。此前已有研究探索用于三角形计数的高效基于草图的算法，例如 [JG05]。事实上，AMS 草图启发了若干用于图问题的此类高效算法 [HRR98, FKM+05b, FKM+04, FKM+05a]。推论 3 中给出的三角形计数结果达到 [JG05], 中第二个估计量的空间界，同时改进了其 $O(mN \log 1/\delta)$ 时间复杂度。此外， [JG05] 中的第二个估计量需要 12-独立哈希函数，而我们的方法至多依赖 4-独立哈希函数，在实践中效率显著更高。

推论 3. 对于任意有向图 G ，其包含 n 个节点、 N 条边，邻接矩阵为 A，且任意给定ϵ、 $\delta > 0$，可以在单次遍历边的过程中计算出 G 中的三角形数量，绝对误差至多为 ϵ

[4]中间查询结果本身也可以表示为一个查询（即待优化查询的子查询）。

，使用算法 *1* 的时间为 $O((m \log m + N) \log 1/\delta)$，占用 $O(m \log 1/\delta)$ 空间，且概率至少为 $1 - \delta$，其中 $m = \Omega(N^3/\epsilon^2)$。

证明。令 $\mathcal{X}_1$、$\mathcal{X}_2$、$\mathcal{X}_3 = A$ 和 $E = \{$(1、 6)、 (2、3)、 (4、 5)$\}$。则结论直接由定理 1 得出。请注意，为了从定义 4 中移除相对项 $\|\mathcal{X}\|_{\mathrm{F}} = \|A\|_{\mathrm{F}}^3 = N^{3/2}$，对 m 的界中额外包含了 N^3 项。 □

致谢

感谢 Ioannis Panageas 提供的宝贵反馈。

参考文献

[AGMS99] Noga Alon、Phillip B Gibbons、Yossi Matias 和 Mario Szegedy。有限存储下连接与自连接大小的跟踪。载于第十八届 *ACM SIGMOD-SIGACT-SIGART* 数据库系统原理研讨会论文集，第 10–20 页，1999 年。[AKK+20] Thomas D Ahle、Michael Kapralov、Jakob BT Knudsen、Rasmus Pagh、Ameya Velingker、David P Woodruff 和 Amir Zandieh。高次多项式核的无偏草图化。载于*ACM-SIAM* 离散算法研讨会（*SODA*），第 141–160 页。SIAM，2020 年。另见扩展版 arXiv 预印本 arXiv:1909.01410。[AL10] Itai Arad 和 Zeph Landau。量子计算与张量网络求值。《*SIAM* 计算杂志》，39(7):3089–3121，2010 年。[AMS96] Noga Alon、Yossi Matias 和 Mario Szegedy。近似频率矩的空间复杂度。载于*ACM* 计算理论研讨会（*STOC*），第 20–29 页，1996 年。[ANW14] Haim Avron、Huy Nguyen 和 David Woodruff。多项式核的子空间嵌入。神经信息处理系统进展（*NeurIPS*），27，2014 年。[BC17] Jacob C Bridgeman 和 Christopher T Chubb。浅谈与诠释性舞蹈：张量网络入门课程。《物理学杂志 *A*：数学与理论》，50(22):223001，2017 年。[CCFC02] Moses Charikar、Kevin Chen 和 Martin Farach-Colton。查找数据流中的频繁项。载于国际自动机、语言与编程研讨会（*ICALP*），第 693–703 页。Springer，2002 年。[CG05] Graham Cormode 和 Minos Garofalakis。通过网络对数据流进行草图化：分布式近似查询跟踪。载于超大型数据库国际会议（*VLDB*），第 13–24 页，2005 年。[Dee25] Kyle Deeds。数据感知复杂度分析与程序优化。博士论文，华盛顿大学，2025 年。[DGGR02] Alin Dobra、Minos Garofalakis、Johannes Gehrke 和 Rajeev Rastogi。处理数据流上的复杂聚合查询。载于*2002* 年 *ACM SIGMOD* 数据管理国际会议论文集，第 61–72 页，2002 年。

[FKM+04] Joan Feigenbaum、Sampath Kannan、Andrew McGregor、Siddharth Suri 和 Jian Zhang。半流式模型中的图问题。载于国际自动机、语言与编程研讨会（ICALP），第 531–543 页。Springer，2004 年。[FKM+05a] Joan Feigenbaum、Sampath Kannan、Andrew McGregor、Siddharth Suri 和 Jian Zhang。流式模型中的图距离：空间的价值。载于 ACM-SIAM 离散算法研讨会（SODA），第 5 卷，第 745–754 页，2005 年。[FKM+05b] Joan Feigenbaum、Sampath Kannan、Andrew McGregor、Siddharth Suri 和 Jian Zhang。半流式模型中的图问题。《理论计算机科学》，348(2-3):207–216，2005 年。[GVL13] Gene H Golub 和 Charles F Van Loan。矩阵计算。约翰斯·霍普金斯大学出版社，2013 年。[HNGN24] Mike Heddes、Igor Nunes、Tony Givargis 和 Alex Nicolau。计数草图的卷积与互相关可实现多重连接查询的快速基数估计。载于 2024 年 ACM SIGMOD 数据管理国际会议论文集，2024 年。[HRR98] Monika Rauch Henzinger、Prabhakar Raghavan 和 Sridhar Rajagopalan。数据流计算。外存算法，50:107–118，1998 年。[IK84] Toshihide Ibaraki 和 Tiko Kameda。计算 n 元关系连接的最优嵌套顺序。《ACM数据库系统汇刊》（TODS），9(3):482–502，1984 年。[JG05] Hossein Jowhari 和 Mohammad Ghodsi。用于图中三角形计数的新型流式算法。载于《计算与组合数学：第十一届年度国际会议，COCOON 2005》，中国昆明，2005 年 8 月 16–19 日，论文集 11，第 710–716 页。Springer，2005 年。[JKS08] T. S. Jayram、Ravi Kumar 和 D. Sivakumar。汉明距离的单向通信复杂度。《计算理论》，4(6):129–135，2008 年。[LGM+15] Viktor Leis、Andrey Gubichev、Atanas Mirchev、Peter Boncz、Alfons Kemper 和 Thomas Neumann。查询优化器究竟有多好？VLDB基金会会刊，9(3):204–215，2015 年。[MB18] Osman Asif Malik 和 Stephen Becker。利用 TensorSketch 对大型张量进行低秩 Tucker 分解。载于 S. Bengio、H. Wallach、H. Larochelle、K. Grauman、N. Cesa-Bianchi 和 R. Garnett 编，《神经信息处理系统进展》，第 31 卷。Curran Associates, Inc.，2018 年。[MS08] Igor L Markov 和 Yaoyun Shi。通过收缩张量网络来模拟量子计算。《SIAM 计算杂志》，38(3):963–981，2008 年。[MS21] Linjian Ma 和 Edgar Solomonik。用于低秩张量分解的快速、准确随机化算法。载于 M. Ranzato、A. Beygelzimer、Y. Dauphin、P.S. Liang 和 J. Wortman Vaughan 编，《神经信息处理系统进展》，第 34 卷，第 24299–24312 页。Curran Associates, Inc.，2021 年。[MWZ24] Arvind V. Mahankali、David P. Woodruff 和 Ziyu Zhang。张量分解的近线性时间与固定参数可解算法。载于 Venkatesan Guruswami 编，第 15 届理论计算机科学创新会议（ITCS

2024)，载于莱布尼茨国际信息学论文集 *(LIPIcs)*第 287 卷，第 79:1–79:23 页，达格施图尔，德国，2024 年。达格施图尔城堡——莱布尼茨信息学中心。[NPOV15]Alexander Novikov、Dmitrii Podoprikhin、Anton Osokin 和 Dmitry P Vetrov。神经网络张量化。神经信息处理系统进展 *(NeurIPS)*，第 28 卷，2015 年。[Ose11] Ivan V Oseledets。张量列车分解。*SIAM*科学计算杂志，33(5):2295–2317，2011 年。[Pag13] Rasmus Pagh。压缩矩阵乘法。*ACM*计算理论汇刊 *(TOCT)*，5(3):1–17，2013 年。[PP13] Ninh Pham 和 Rasmus Pagh。通过显式特征映射实现快速且可扩展的多项式核。载于*ACM SIGKDD*知识发现与数据挖掘国际会议，第 239–247 页，2013 年。[RS19] Elina Robeva 和 Anna Seigal。图模型与张量网络的对偶性。*Information and Inference: A Journal of the IMA*，8(2):273–288，2019 年。[WDL+09] Kilian Weinberger、Anirban Dasgupta、John Langford、Alex Smola 和 Josh Attenberg。面向大规模多任务学习的特征哈希。载于国际机器学习会议 *(ICML)*，第 1113–1120 页，2009 年。[WJ+08] Martin J Wainwright、Michael I Jordan 等。图模型、指数族与变分推断。基础与趋势® 机器学习，1(1–2):1–305，2008 年。[Woo14] David P Woodruff。草图化：数值线性代数的工具。基础与趋势® 理论计算机科学，10(1–2):1–157，2014 年。[WTSA15] Yining Wang、Hsiao-Yu Tung、Alexander J Smola 和 Anima Anandkumar。通过草图化实现快速且有保证的张量分解。神经信息处理系统进展 *(NeurIPS)*，第 28 卷，2015 年。

A 证明

本附录包含论文中所述结果的证明。

A.1 引理 5 的证明

证明。 我们首先利用离散傅里叶变换和计数草图的定义，得到

$$\begin{aligned}(\overline{FCx})_k &= \overline{\sum_{j=1}^{m} e^{-i2\pi\frac{k-1}{m}(j-1)} \sum_{i=1}^{n} x_i s(i)[h(i)=j]} \\ &= \sum_{j=1}^{m} e^{i2\pi\frac{k-1}{m}(j-1)} \sum_{i=1}^{n} x_i s(i)[h(i)=j].\end{aligned}$$

这里用到了复共轭的以下性质： $x+y=\overline{x}+\overline{y}$、 $xy=\overline{x}\,\overline{y}$、$\exp(\overline{i})=\exp(-i)$，以及当 $\overline{x}=x$ 时的 $x\in\mathbb{R}$。根据补计数草图的定义，我们有

$$(FC'x)_k = \sum_{l=1}^{m} e^{-i2\pi\frac{k-1}{m}(l-1)} \sum_{i=1}^{n} x_i s(i)[-h(i)+2 \equiv l \pmod m].$$

通过变量替换，令 $l=-j+2$，我们得到

$$\begin{aligned}(FC'x)_k &= \sum_{j=1}^{m} e^{-i2\pi\frac{k-1}{m}(-j+2-1)} \sum_{i=1}^{n} x_i s(i)[-h(i)+2 \equiv -j+2 \pmod m] \\ &= \sum_{j=1}^{m} e^{i2\pi\frac{k-1}{m}(j-1)} \sum_{i=1}^{n} x_i s(i)[h(i)=j] = (\overline{FCx})_k,\end{aligned}$$

由此证明得证。 □

A.2 引理 6 的证明

证明。 首先注意到，每个张量都可以写成标准基向量张量积的线性和：$\mathcal{X}_k = \sum_{\mathbf{i}_k=\mathbf{1}}^{\mathbf{n}_k} \mathcal{X}_k(\mathbf{i}_k) \times_{u\in V(k)} e_{i_u}$，其中 $\times_{u\in S}$ 表示 S 的每个成员 u 对应的张量积。由于 $x_k = \left(\bullet_{u\in V(k)} FC_u\right)\mathrm{vec}(\mathcal{X}_k)$，根据线性性和混合乘积恒等式 $(A\bullet B)(x\otimes y) = (Ax)\circ(By)$，

$$x_k = \left(\bullet_{u\in V(k)} FC_u\right)\left(\sum_{\mathbf{i}_k=\mathbf{1}}^{\mathbf{n}_k} \mathcal{X}_k(\mathbf{i}_k) \bigotimes_{u\in V(k)} e_{i_u}\right) = \sum_{\mathbf{i}_k=\mathbf{1}}^{\mathbf{n}_k} \mathcal{X}_k(\mathbf{i}_k)\left(\circ_{u\in V(k)} FC_u e_{i_u}\right),$$

其中，$\circ_{u\in S}$ 是每个 u 属于 S的成员对应的哈达玛积。这将张量$\mathcal{X}_k$的草图 x_k 表示为其各元素之和。注意，$\bullet_{u\in V(k)} FC_u$ 是张量草图矩阵的傅里叶变换，而 $\mathrm{vec}(\mathcal{X}_k) = \sum_{\mathbf{i}_k=\mathbf{1}}^{\mathbf{n}_k} \mathcal{X}_k(\mathbf{i}_k)\bigotimes_{u\in V(k)} e_{i_u}$。随后通过组合以下草图与哈达玛积，得到估计值：

如下将草图与哈达玛积组合：

$$\hat{\mathrm{tc}}(\mathcal{X}_1 \times \cdots \times \mathcal{X}_p, E) = e_1^\mathsf{T} F^{-1}(x_1 \circ \cdots \circ x_p) \tag{3}$$

$$= \sum_{\mathbf{i}=\mathbf{1}}^{\mathbf{n}} \mathcal{X}_1(\mathbf{i}_1) \cdots \mathcal{X}_p(\mathbf{i}_p) e_1^\mathsf{T} F^{-1} \left(\mathop{\circ}_{u \in [q]} FC_u e_{i_u} \right) \tag{4}$$

$$= \sum_{\mathbf{i}=\mathbf{1}}^{\mathbf{n}} \mathcal{X}_1(\mathbf{i}_1) \cdots \mathcal{X}_p(\mathbf{i}_p) e_1^\mathsf{T} F^{-1} \left(\left(\mathop{\circ}_{(u,v) \in E} FC_u e_{i_u} \right) \circ \left(\mathop{\circ}_{(u,v) \in E} FC_v e_{i_v} \right) \right) \tag{5}$$

$$= \sum_{\mathbf{i}=\mathbf{1}}^{\mathbf{n}} \mathcal{X}_1(\mathbf{i}_1) \cdots \mathcal{X}_p(\mathbf{i}_p) e_1^\mathsf{T} F^{-1} \left(\left(\mathop{\circ}_{(u,v) \in E} FC_u e_{i_u} \right) \circ \left(\mathop{\circ}_{(u,v) \in E} FC'_u e_{i_v} \right) \right) \tag{6}$$

$$= \sum_{\mathbf{i}=\mathbf{1}}^{\mathbf{n}} \mathcal{X}_1(\mathbf{i}_1) \cdots \mathcal{X}_p(\mathbf{i}_p) e_1^\mathsf{T} F^{-1} \left(\left(\mathop{\circ}_{(u,v) \in E} FC_u e_{i_u} \right) \circ \overline{\left(\mathop{\circ}_{(u,v) \in E} FC_u e_{i_v} \right)} \right) \tag{7}$$

$$= \frac{1}{m} \sum_{\mathbf{i}=\mathbf{1}}^{\mathbf{n}} \mathcal{X}_1(\mathbf{i}_1) \cdots \mathcal{X}_p(\mathbf{i}_p) \left\langle \mathop{\circ}_{(u,v) \in E} FC_u e_{i_u}, \mathop{\circ}_{(u,v) \in E} FC_u e_{i_v} \right\rangle \tag{8}$$

$$= \frac{1}{m} \sum_{\mathbf{i}=\mathbf{1}}^{\mathbf{n}} \mathcal{X}_1(\mathbf{i}_1) \cdots \mathcal{X}_p(\mathbf{i}_p) \left\langle FT \bigotimes_{(u,v) \in E} e_{i_u}, FT \bigotimes_{(u,v) \in E} e_{i_v} \right\rangle \tag{9}$$

$$= \sum_{\mathbf{i}=\mathbf{1}}^{\mathbf{n}} \mathcal{X}_1(\mathbf{i}_1) \cdots \mathcal{X}_p(\mathbf{i}_p) \left\langle T \bigotimes_{(u,\ v) \in E} e_{i_u}, T \bigotimes_{(u,v) \in E} e_{i_v} \right\rangle. \tag{10}$$

式 (5) 将计数草图和补计数草图分成两组。请注意，对于每次收缩，一个计数草图被分配给一个模态，其补项被分配给另一个模态，因此两者数量相等。式 (7) 使用引理 5，式 (8) 使用 $e_1^T F^{-1}(x \circ y) = \frac{1}{m}\langle x, y \rangle$。证明的这一步要求每次收缩都具有一个共轭模态，这通过使用补计数草图实现。式 (9) 利用了式 (8) 包含张量草图的傅里叶变换这一事实。式 (10) 利用了内积在酉变换下的不变性，因为 F 是一个酉矩阵乘以 $\sqrt{m}$。估计量的期望为

$$\mathbb{E}[\hat{\mathrm{tc}}(\mathcal{X}_1 \times \cdots \times \mathcal{X}_p, E)] = \sum_{\mathbf{i}=\mathbf{1}}^{\mathbf{n}} \mathcal{X}_1(\mathbf{i}_1) \cdots \mathcal{X}_p(\mathbf{i}_p) \mathbb{E} \left\langle T \bigotimes_{(u,v) \in E} e_{i_u}, T \bigotimes_{(u,v) \in E} e_{i_v} \right\rangle \tag{11}$$

$$\mathbb{E} \left\langle T \bigotimes_{(u,v) \in E} e_{i_u}, T \bigotimes_{(u,v) \in E} e_{i_v} \right\rangle = \left\langle \bigotimes_{(u,v) \in E} e_{i_u}, \bigotimes_{(u,v) \in E} e_{i_v} \right\rangle = \prod_{(u,v) \in E} \langle e_{i_u}, e_{i_v} \rangle = \prod_{(u,v) \in E} [i_u = i_v]. \tag{12}$$

式 12 中的期望值和方差界可直接由引理 3 得出。方差界中每个张量的弗罗贝尼乌斯范数平方来自张量分量之和。请注意，在我们的证明中，我们不对张量网络的结构作任何假设。

□

A.3 引理 7 的证明

证明。 由于 $\mathrm{vec}(\mathcal{X}_i) = \mathbf{1}_n \otimes \mathbf{1}_n$ 对于 $2 \le i \le q$，我们使用卷积定理 $F(x \star y) = Fx \circ Fy$ 以及混合乘积恒等式 $(A \bullet B)(x \otimes y) = (Ax) \circ (By)$ 得到

$$y = e_1^\mathsf{T} F^{-1}((FC_1\mathbf{1}) \circ \overline{(FC_1\mathbf{1}) \circ (FC_2\mathbf{1})} \circ \cdots \circ \overline{(FC_{q-1}\mathbf{1}) \circ (FC_q\mathbf{1})} \circ (FC_q\mathbf{1})),$$

其中复共轭作用于每隔一对计数草图。我们可以利用哈达玛积的结合律重新排列该表达式，使所有共轭项位于同一侧。注意 $e_1^\top F^{-1}(x \circ y) = \frac{1}{m}\mathbf{1}^\top(x \circ y) = \frac{1}{m}\langle x, y\rangle$，以及 $\|x\|_2^2 = \langle x, x\rangle$，从而得到

$$\frac{1}{m}\langle (FC_1\mathbf{1}) \circ \cdots \circ (FC_q\mathbf{1}), (FC_1\mathbf{1}) \circ \cdots \circ (FC_q\mathbf{1})\rangle = \frac{1}{m}\|(FC_1\mathbf{1}) \circ \cdots \circ (FC_q\mathbf{1})\|_2^2.$$

我们现在应用 [AKK+20，引理 43]，，因为 $\|T(x^{\otimes q})\|_2^2 = \frac{1}{m}\|(FC_1\mathbf{1}) \circ \cdots \circ (FC_q\mathbf{1})\|_2^2$，由于 F^{-1}是酉矩阵乘以 $\frac{1}{\sqrt{m}}$。最后，注意到$\|x^{\otimes q}\|_2^4 = \prod_{i=1}^{q+1}\|\mathcal{X}_i\|_{\mathrm{F}}^2$，从而完成证明。 □

A.4 引理 8 的证明

证明。 我们可以假设张量网络收缩满足第 4.1 节中的不失一般性假设。然后，通过展开 $g(o)$以及内积的定义，我们有

$$\mathrm{vec}(\mathcal{X}_o)^\top g(o) = \sum_{\mathbf{i}_o=\mathbf{1}}^{\mathbf{n}} \mathcal{X}_o(\mathbf{i}_o)\left(\bigotimes_{l_1\in\Gamma(o)} \mathrm{mat}(\mathcal{X}_{l_1})g(l_1)\right)(\mathbf{i}_o),$$

其中$\mathbf{i}_k$ 包含张量 $\mathcal{X}_k$的所有指标。令$\mathbf{o}_k$ 包含与张量 $\mathcal{X}_k$的出边相关联的所有指标，并令$\mathbf{l}_i$ 包含距离根节点 i处所有张量的所有指标，也称为树的第 i层。注意到$\mathbf{i}_o = \mathbf{l}_1$ ，因为根张量的指标就是第一层中的所有指标。令 $E(l)$表示属于树的前 l层的收缩子集。于是我们得到如下结果：

$$\mathrm{vec}(\mathcal{X}_o)^\top g(o) = \sum_{\mathbf{l}_1=\mathbf{1}}^{\mathbf{n}}\sum_{\mathbf{l}_2=\mathbf{1}}^{\mathbf{n}} \mathcal{X}_o(\mathbf{i}_o)\left(\prod_{l_1\in\Gamma(o)} \mathcal{X}_{l_1}(\mathbf{i}_{l_1})g(l_1)(\mathbf{o}_{l_1})\right)\prod_{(u,v)\in E(1)}[i_u = i_v],$$

其中，引入$\mathbf{i}_o$ 后，指标会与$\mathbf{i}_{l_1}$成对出现。添加的艾弗森括号确保这两个指标相等。随后，我们可以对树的所有 w层重复展开递归，得到

$$\begin{aligned}\mathrm{vec}(\mathcal{X}_o)^\top g(o) &= \sum_{\mathbf{l}_1=\mathbf{1}}^{\mathbf{n}}\sum_{\mathbf{l}_2=\mathbf{1}}^{\mathbf{n}}\cdots\sum_{\mathbf{l}_w=\mathbf{1}}^{\mathbf{n}} \mathcal{X}_o(\mathbf{i}_o)\left(\prod_{l_1\in\Gamma(o)} \mathcal{X}_{l_1}(\mathbf{i}_{l_1})\left(\prod_{l_2\in\Gamma(l_1)} \mathcal{X}_{l_2}(\mathbf{i}_{l_2})\dots\right)\right)\prod_{(u,v)\in E}[i_u = i_v]\\ &= \sum_{\mathbf{i}=\mathbf{1}}^{\mathbf{n}} \mathcal{X}_1(\mathbf{i}_1)\cdots\mathcal{X}_p(\mathbf{i}_p)\prod_{(u,v)\in E}[i_u = i_v],\end{aligned}$$

其中$\mathbf{i}$ 包含所有指标，因为这是一次完全张量网络收缩，由此证明得证 f. □

A.5 引理 9 的证明

证明。证明 接下来先求期望，再给出方差界 .

期望。 注意，所有递归草图矩阵 $R_1, \dots, R_p$彼此独立。于是，根据全期望定律和引理 4，并令 $x_o = \mathrm{vec}(\mathcal{X}_o)$，

$$\mathbb{E}[y] = \mathbb{E}\left[x_o^\top R_o^\top R_o\hat{g}(o)\right] = \mathbb{E}\left[\mathbb{E}\left[x_o^\top R_o^\top R_o\hat{g}(o)\middle|\hat{g}(o)\right]\right] = x_o^\top\left(\bigotimes_{l\in\Gamma(o)} \mathbb{E}\left[\mathrm{mat}(\mathcal{X}_l)R_l^\top R_l\hat{g}(l)\right]\right).$$

因此， $\mathbb{E}[x_o^\top R_o^\top R_o\hat{g}(o)] = x_o^\top g(o)$，因为对任意$k$， $\mathbb{E}\left[\mathrm{mat}(\mathcal{X}_k)R_k^\top R_k\hat{g}(k)\middle|\hat{g}(k)\right] = \mathrm{mat}(\mathcal{X}_k)\hat{g}(k)$，这与引理 8 中的结果一致。

方差界。 根据全方差公式、引理 4，以及 $x_o = \mathsf{vec}(\mathcal{X}_o)$，

$$\begin{aligned}\mathrm{Var}\big(\hat{\mathrm{tc}}(\mathcal{X}_1 \times \cdots \times \mathcal{X}_p, E)\big) &= \mathbb{E}\Big[\mathrm{Var}\Big(x_o^\mathsf{T} R_o^\mathsf{T} R_o \hat{g}(o)\Big|\hat{g}(o)\Big)\Big] + \mathrm{Var}\Big(\mathbb{E}\Big[x_o^\mathsf{T} R_o^\mathsf{T} R_o \hat{g}(o)\Big|\hat{g}(o)\Big]\Big) \\ &= \left(\left(1+\frac{8}{m}\right)^{2q_o} - 1\right)\|\mathcal{X}_o\|_\mathrm{F}^2\,\mathbb{E}\big[\|\hat{g}(o)\|_2^2\big] + \mathrm{Var}\Big(x_o^\mathsf{T}\hat{g}(o)\Big).\end{aligned}$$

对于 $\hat{}g$ 的范数，取 $M_l = \mathsf{mat}(\mathcal{X}_l)$，我们有

$$\begin{aligned}\mathbb{E}\big[\|\hat{g}(k)\|_2^2\big|\hat{g}(l)\ \forall l \in \Gamma(k)\big] &= \prod_{l\in\Gamma(k)} \mathbb{E}\Big[\|\mathrm{mat}(\mathcal{X}_l)R_l^\mathsf{T} R_l \hat{g}(l)\|_2^2\Big|\hat{g}(l)\Big] \\ &= \prod_{l\in\Gamma(k)}\sum_{i=1}^{n}\mathbb{E}\Big[\Big(M_l(i,:)R_l^\mathsf{T} R_l\hat{g}(l)\Big)^2\Big|\hat{g}(l)\Big] \\ &= \prod_{l\in\Gamma(k)}\sum_{i=1}^{n}\mathrm{Var}\Big(M_l(i,:)R_l^\mathsf{T} R_l\hat{g}(l)\Big|\hat{g}(l)\Big) + \mathbb{E}\Big[M_l(i,:)R_l^\mathsf{T} R_l\hat{g}(l)\Big|\hat{g}(l)\Big]^2.\end{aligned}$$

利用引理 4 和柯西-施瓦茨不等式，我们得到以下界：

$$\begin{aligned}\mathbb{E}\big[\|\hat{g}(k)\|_2^2\big|\hat{g}(l)\ \forall l \in \Gamma(k)\big] &\le \prod_{l\in\Gamma(k)}\left(\left(1+\frac{8}{m}\right)^{2q_l-2} - 1\right)\|\mathcal{X}_l\|_\mathrm{F}^2\,\mathbb{E}\big[\|\hat{g}(l)\|_2^2\big] + \|\mathcal{X}_l\|_\mathrm{F}^2\,\mathbb{E}\big[\|\hat{g}(l)\|_2^2\big] \\ &= \prod_{l\in\Gamma(k)}\left(1+\frac{8}{m}\right)^{2q_l-2}\|\mathcal{X}_l\|_\mathrm{F}^2\,\mathbb{E}\big[\|\hat{g}(l)\|_2^2\big].\end{aligned}$$

回到剩余的方差项，我们有

$$\begin{aligned}\mathrm{Var}\Big(x_o^\mathsf{T}\hat{g}(o)\Big) &= \mathbb{E}\left[\left(x_o^\mathsf{T}\left(\textstyle\bigotimes_{l\in\Gamma(o)}\mathrm{mat}(\mathcal{X}_l)R_l^\mathsf{T} R_l\hat{g}(l)\right) - x_o^\mathsf{T}\left(\textstyle\bigotimes_{l\in\Gamma(o)}\mathrm{mat}(\mathcal{X}_l)g(l)\right)\right)^2\right] \\ &= \mathbb{E}\left[\left(x_o^\mathsf{T}\left(\textstyle\bigotimes_{l\in\Gamma(o)}\mathrm{mat}(\mathcal{X}_l)R_l^\mathsf{T} R_l\hat{g}(l) - \mathrm{mat}(\mathcal{X}_l)g(l)\right)\right)^2\right] \\ &\le \|\mathcal{X}_o\|_\mathrm{F}^2\prod_{l\in\Gamma(o)}\mathbb{E}\Big[\|\mathrm{mat}(\mathcal{X}_l)R_l^\mathsf{T} R_l\hat{g}(l) - \mathrm{mat}(\mathcal{X}_l)g(l)\|_2^2\Big] \qquad \text{(Cauchy-Schwarz)} \\ &= \|\mathcal{X}_o\|_\mathrm{F}^2\prod_{l\in\Gamma(o)}\sum_{i=1}^{n}\mathbb{E}\Big[\Big(M_l(i,:)R_l^\mathsf{T} R_l\hat{g}(l) - M_l(i,:)g(l)\Big)^2\Big] \\ &= \|\mathcal{X}_o\|_\mathrm{F}^2\prod_{l\in\Gamma(o)}\sum_{i=1}^{n}\mathrm{Var}\Big(M_l(i,:)R_l^\mathsf{T} R_l\hat{g}(l)\Big)\end{aligned}$$

然后，我们应用全方差公式，得到：

$$\begin{aligned}\mathrm{Var}\Big(x_o^\mathsf{T}\hat{g}(o)\Big) &\le \|\mathcal{X}_o\|_\mathrm{F}^2\prod_{l\in\Gamma(o)}\sum_{i=1}^{n}\mathbb{E}\Big[\mathrm{Var}\Big(M_l(i,:)R_l^\mathsf{T} R_l\hat{g}(l)\Big|\hat{g}(l)\Big)\Big] + \mathrm{Var}\Big(\mathbb{E}\Big[M_l(i,:)R_l^\mathsf{T} R_l\hat{g}(l)\Big|\hat{g}(l)\Big]\Big) \\ &\le \|\mathcal{X}_o\|_\mathrm{F}^2\prod_{l\in\Gamma(o)}\left(\left(1+\frac{8}{m}\right)^{2q_l-2} - 1\right)\|\mathcal{X}_l\|_\mathrm{F}^2\,\mathbb{E}\big[\|\hat{g}(l)\|_2^2\big] + \sum_{i=1}^{n}\mathrm{Var}(M_l(i,:)\hat{g}(l)).\end{aligned}$$

综合目前为止的结果，我们得到

$$\begin{aligned}\mathrm{Var}(y) \le &\left(\left(1+\frac{8}{m}\right)^{2q_o}-1\right)\|\mathcal{X}_o\|_{\mathrm{F}}^2 \prod_{l\in\Gamma(o)}\left(1+\frac{8}{m}\right)^{2q_l-2}\|\mathcal{X}_l\|_{\mathrm{F}}^2\,\mathbb{E}\big[\|\hat{g}(l)\|_2^2\big] \\ &+\|\mathcal{X}_o\|_{\mathrm{F}}^2 \prod_{l\in\Gamma(o)}\left(\left(1+\frac{8}{m}\right)^{2q_l-2}-1\right)\|\mathcal{X}_l\|_{\mathrm{F}}^2\,\mathbb{E}\big[\|\hat{g}(l)\|_2^2\big]+\sum_{i=1}^{n}\mathrm{Var}(M_l(i,:)\hat{g}(l)),\end{aligned}$$

第一项通过从根节点到叶节点遍历张量图，递归地累加所有张量范数和 $(1+8/m)$ 系数。总共有 $q_o+\sum_{k=1:\ k\neq o}^{p} q_k-1=q-p+1=t$ 个这样的 $(1+8/m)$ 系数。注意，对于树的每个叶节点，$\hat{g}(l)=1$，其方差为零。若设 $t'=t-q_o$，则得到如下结果：

$$\begin{aligned}\mathrm{Var}(y) \le &\left(1+\frac{8}{m}\right)^{2t}\prod_{k=1}^{p}\|\mathcal{X}_k\|_{\mathrm{F}}^2-\left(1+\frac{8}{m}\right)^{2t'}\prod_{k=1}^{p}\|\mathcal{X}_k\|_{\mathrm{F}}^2 \\ &+\|\mathcal{X}_o\|_{\mathrm{F}}^2 \prod_{l\in\Gamma(o)}\left(\left(1+\frac{8}{m}\right)^{2q_l-2}-1\right)\|\mathcal{X}_l\|_{\mathrm{F}}^2\,\mathbb{E}\big[\|\hat{g}(l)\|_2^2\big]+\sum_{i=1}^{n}\mathrm{Var}(M_l(i,:)\hat{g}(l)),\end{aligned}$$

最后一个不等式是根据以下观察得出的：对叶张量范数的重复计数会使最后两项相消。不过应注意，我们至少多计了范数的乘积，因此最终得到

$$\mathrm{Var}(y) \le \left(1+\frac{8}{m}\right)^{2t}\prod_{k=1}^{p}\|\mathcal{X}_k\|_{\mathrm{F}}^2-\left(1+\frac{8}{m}\right)^{2t'}\prod_{k=1}^{p}\|\mathcal{X}_k\|_{\mathrm{F}}^2+\left(1+\frac{8}{m}\right)^{2t'}\prod_{k=1}^{p}\|\mathcal{X}_k\|_{\mathrm{F}}^2-\prod_{k=1}^{p}\|\mathcal{X}_k\|_{\mathrm{F}}^2,$$

其等于 $\left(\left(1+\frac{8}{m}\right)^{2t}-1\right)\prod_{k=1}^{p}\|\mathcal{X}_k\|_{\mathrm{F}}^2$，证明完毕。

□

A.6 定理 1 的证明

证明。令 $y=\hat{\mathrm{tc}}(\mathcal{X}_1\times\cdots\times\mathcal{X}_p,E)$，那么，由于这是完全张量网络收缩，

$$\mathbb{E}\Big[\|y-\mathrm{tc}(\mathcal{X}_1\times\cdots\times\mathcal{X}_p,E)\|_{\mathrm{F}}^2\Big]=\mathbb{E}\Big[(y-\mathrm{tc}(\mathcal{X}_1\times\cdots\times\mathcal{X}_p,E))^2\Big]=\mathrm{Var}(y).$$

注意，当 $m\ge 16t$ 时，$(1+8/m)^{2t}-1\le\exp(16t/m)-1\le 32t/m$。因此，根据(1)引理 9，令 $m=\Omega(t/\epsilon^2)$；根据(2)引理 6，令 $m=\Omega(3^t/\epsilon^2)$，从而利用切比雪夫不等式，以常数成功概率得到完全近似张量网络收缩的估计量（见定义 4）。注意$\prod_{k=1}^{p}\|\mathcal{X}_k\|_{\mathrm{F}}=\|\mathcal{X}_1\times\cdots\times\mathcal{X}_p\|_{\mathrm{F}}$。通过切尔诺夫界，对 $O(\log 1/\delta)$ 个相互独立且同分布、具有常数成功概率的估计值取中位数，即可获得高成功概率。

□

A.7 推论 1 的证明

证明。将误差的弗罗贝尼乌斯范数平方展开，其中 $\mathcal{Y}$ 和 $\mathcal{Y}$ 分别表示近似张量网络收缩和精确张量网络收缩，可得

$$\mathbb{E}\Big[\|\hat{\mathrm{tc}}(\mathcal{X}_1\times\cdots\times\mathcal{X}_p,E)-\mathrm{tc}(\mathcal{X}_1\times\cdots\times\mathcal{X}_p,E)\|_{\mathrm{F}}^2\Big]=\sum_{\mathbf{r}=\mathbf{1}}^{\mathbf{n}}\mathbb{E}\Big[\Big(\hat{\mathcal{Y}}(\mathbf{r})-\mathcal{Y}(\mathbf{r})\Big)^2\Big]=\sum_{\mathbf{r}=\mathbf{1}}^{\mathbf{n}}\mathrm{Var}\Big(\hat{\mathcal{Y}}(\mathbf{r})\Big).$$

ˆ令 $\mathcal{X}_k(\mathbf{s}_k)$ 表示用于计算条目 $\mathcal{Y}(\mathbf{r})$ 的 $\mathcal{X}_k$ 的子张量，其中

$$\mathbf{s}_k = (\gamma(u) : u \in V(k)), \qquad \text{and} \qquad \gamma(u) = \begin{cases} : & \text{if } u \in K, \\ i_u & \text{otherwise,} \end{cases}$$

其中 $V(k)$ 是 $\mathcal{X}_k$ 的模， K 是收缩模态。然后，根据（**1**）引理 9，令$m \geq 32t/(\epsilon^2\delta)$；并根据（**2**）引理 6，令 $m \geq 3^t/(\epsilon^2\delta)$,

$$\sum_{\mathbf{r}=\mathbf{1}}^{\mathbf{n}} \mathrm{Var}\Big(\hat{\mathcal{Y}}(\mathbf{r})\Big) \leq \sum_{\mathbf{r}=\mathbf{1}}^{\mathbf{n}} \epsilon^2\delta \prod_{k=1}^{p} \|\mathcal{X}_k(\mathbf{s}_k)\|_{\mathrm{F}}^2 = \epsilon^2\delta \prod_{k=1}^{p} \|\mathcal{X}_k\|_{\mathrm{F}}^2.$$

根据切比雪夫不等式和定义 4，证明至此完成。 □

B 补充算法

本附录包含论文中主要算法所用函数的伪代码。我们首先介绍算法 3，该算法针对给定的列（对应于一个张量索引）计算张量草图矩阵的符号和行。算法 4 使用此函数来计算给定列对应的递归草图矩阵的符号和行。随后，算法 5 使用该函数同时计算张量的草图及其矩阵-向量乘积。

阶数为 q 的张量草图矩阵由 q 个计数草图矩阵组成，而每个计数草图矩阵又由 q 个符号与行哈希函数构成。张量草图矩阵 $T \in \mathbb{R}^{m \times n^q}$ 的另一种定义将每个分量指定为 $T(j,\ i) = s_1(i_1) \cdots s_q(i_q)$ $[(h_1(i_1) + \cdots + h_q(i_q) - q + 1) \equiv j \text{(模 } m)]$,其中 $i = 1 + \sum_{k=1}^{q}(i_k - 1)n^{q-k}$ (张量索引 $\mathbf{i}$ 的向量化)。基于此公式，算法 3 针对给定列计算张量草图矩阵的符号和行。具体而言，对每个模态的计数草图哈希函数所得的符号相乘，并将行相加（模草图大小）。

算法 3 按列计算张量草图矩阵的符号和行。

输入：索引 $\mathbf{i} = (i_1, \ldots, i_q)$ 和张量草图矩阵 $T \in \mathbb{R}^{m \times n^q}$，该矩阵由相互独立的计数草图矩阵 $C_1, \ldots, C_q \in \mathbb{R}^{m \times n}$ 组成；这些矩阵对所有 $k \in [q]$ 均包含符号哈希函数 s_k 和行哈希函数 h_k 输出：$\sigma \in \{-1, 1\}$ 和 $b \in [m]$1：函数 TensorSketchHash(**i** T) **2：** $\sigma \leftarrow 1$**3：** $b \leftarrow 0$**4：对于** $k \in [q]$ **执行 5：** $\sigma \leftarrow \sigma \cdot s_k(i_k)$ **6：** $b \leftarrow b + h_k(i_k)$ **7：结束循环 8：** $b \leftarrow ((b - q)$ **模** $m) + 1$**9：返回** σ, b**10：结束函数**

算法 4 计算给定列在递归草图矩阵中的符号和行。这通过使用算法 3，按照二叉树的方式归约一个行指标列表来实现；这些行指标由各模态的计数草图矩阵的行哈希函数得到。在每一步中，符号通过乘法累积。

算法 5 在计算阶数为 q 的张量 $\mathcal{X}$ 的大小为 m 的草图的同时，将所得草图矩阵与一个向量相乘。这是我们第二种方法中的一项重要功能，因为它显著降低了计算复杂度：若先计算草图，再进行矩阵-向量乘积，则需要 $O(q\,\mathrm{nnz}(\mathcal{X}) + m^2)$ 时间和 $O(m^2)$ 空间；而该算法仅需 $O(q\,\mathrm{nnz}(\mathcal{X}))$ 时间和 $O(m)$ 空间。

算法 4 递归草图矩阵符号与行列。

输入： 下标 $\mathbf{i} = (i_1, \ldots, i_q)$ 以及递归草图矩阵 $R \in \mathbb{R}^{m \times n^q}$，其由 independent 计数草图矩阵 $C_1, \ldots, C_q \in \mathbb{R}^{m \times n}$ 组成，这些矩阵由符号与行哈希函数构成 s_k 与 h_k（对所有 $k \in [q]$），以及独立的张量草图矩阵 $T_{l,1}, \ldots, T_{l,l/2} \in \mathbb{R}^{m \times m}$ [2] for $l \in \{2, 4, \ldots, q/2, q\}$

输出： $\sigma \in \{-1, 1\}$ 与 $b \in [m]$

函数 RecursiveSketchHash($\mathbf{i}$ R)
  $\sigma \leftarrow 1$
  $b \in \{0\}^q$
  对于 $k \in [q]$，**执行**
    $\sigma \leftarrow \sigma \cdot s_k(i_k)$
    $b_k \leftarrow h_k(i_k)$
  循环结束
  对于 $l \in \{q, q/2, \ldots, 4, 2\}$，**执行**
    对于 $k \in [l/2]$，**执行**
      $s,\ h \leftarrow$ TensorSketchHash$(\ (b_{2k-1},\ b_{2k})\ T_{l,\ k})$
      $\sigma \leftarrow \sigma \cdot s$
      $b_k \leftarrow h$
    结束 for
  结束 for
  返回 σ, b_1
结束函数

算法 5 组合张量草图与矩阵-向量乘积。

输入： 阶 q 张量 $\mathcal{X} \in \mathbb{R}^{n \times \cdots \times n}$，计数草图矩阵 $C \in \mathbb{R}^{m \times n}$ 由符号与行哈希函数 s 和 h组成，递归草图矩阵 $R \in \mathbb{R}^{m \times n^{q-1}}$，以及任意向量 $z \in \mathbb{R}^m$

输出： $y = C\,\mathrm{mat}(\mathcal{X})R^\mathsf{T} z$

函数 SketchedMatrixVectorProduct($C, \mathcal{X}, R, z$)
  $y \in \{0\}^m$
  对于 每个分量 $\mathcal{X}$**(i) 执行**
    $j \leftarrow h(i_1)$
    $\sigma, b =$ RecursiveSketchHash$((i_2, i_3, \ldots, i_q)\ R)$
    $y_j \leftarrow y_j + \mathcal{X}(\mathbf{i}) \cdot z_b \cdot \sigma \cdot s(i_1)$
  结束循环
  返回 y
结束函数