Electronic Colloquium on Computational Complexity, Report No. 25 (2026)

Beyond Bilinear Complexity: What Works and What Breaks with Many Modes?

Cornelius Brand* Radu Curticapean† Petteri Kaski‡ Baitian Li§ Ian Orzel¶
Tim Seppelt‖ Jiaheng Wang**

Abstract

The complexity of bilinear maps (equivalently, of 3-mode tensors) has been studied extensively, most notably in the context of matrix multiplication. While circuit complexity and tensor rank coincide asymptotically for 3-mode tensors, this correspondence breaks down for $d \geq 4$ modes. As a result, the complexity of d-mode tensors for larger fixed d remains poorly understood, despite its relevance, e.g., in fine-grained complexity. Our paper explores this intermediate regime.

First, we give a "graph-theoretic" proof of Strassen's $2\omega/3$ bound on the asymptotic rank exponent of 3-mode tensors. Our proof directly generalizes to an upper bound of $(d-1)\omega/3$ for d-mode tensors. Using refined techniques available only for $d \geq 4$ modes, we improve this bound beyond the current state of the art for ω. We also obtain a bound of $d/2+1$ on the asymptotic exponent of *circuit complexity* of generic d-mode tensors and optimized bounds for $d \in \{4,5\}$.

To the best of our knowledge, asymptotic circuit complexity (rather than rank) of tensors has not been studied before. To obtain a robust theory, we first ask whether low complexity of T and U imply low complexity of their Kronecker product $T \otimes U$. While this crucially holds for rank (and thus for circuit complexity in 3 modes), we show that assumptions from fine-grained complexity rule out such a *submultiplicativity* for the circuit complexity of tensors with many modes. In particular, assuming the Hyperclique Conjecture, this failure occurs already for $d = 8$ modes. Nevertheless, we can salvage a restricted notion of submultiplicativity.

From a technical perspective, our proofs heavily make use of the *graph tensors* T_H, as employed by Christandl and Zuiddam (*Comput. Complexity* 28 (2019) 27–56) and Christandl, Vrana and Zuiddam (*Comput. Complexity* 28 (2019) 57–111), whose modes correspond to the vertices of undirected graphs H. We make the simple but conceptually crucial observation that Kronecker products $T_G \otimes T_H$ are isomorphic to T_{G+H}, and that G and H may also be *fractional* graphs. By asymptotically converting generic tensors to specific graph tensors, we can use nontrivial results from algorithmic graph theory to study the rank and complexity of d-mode tensors for fixed d.

*University of Regensburg. Funded by ERC project CountHom.

†University of Regensburg and IT University of Copenhagen. Funded by the European Union (ERC, CountHom, 101077083). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency.

‡Aalto University.

§Columbia University. Funded by a Columbia SEAS Presidential Fellowship.

¶University of Copenhagen. Funded by the European Research Council (ERC) under grant agreement no. 101125652 (ALBA).

‖IT University of Copenhagen. Funded by ERC project CountHom.

**University of Regensburg and University of Helsinki. Funded by ERC project CountHom.

ISSN 1433-8092

Contents

1 Introduction

In several fundamental computational problems, the input data is represented by two vectors u, v and the output is a vector $f(u, v)$ that captures a meaningful combination of the input data. Prominent examples are *matrix multiplication* (when the input matrices are flattened to vectors), *polynomial multiplication* (when the input polynomials are given as coefficient vectors), and more general *convolution* problems such as the *subset convolution* in parameterized and exact algorithms, which is essentially multiplication in the algebra of square-free multivariate polynomials.

Bilinear Maps and Rank. All of the above problems ask us to evaluate maps $f : \mathbb{F}^n \times \mathbb{F}^n \to \mathbb{F}^n$ that are *bilinear* over the field $\mathbb{F}$; they satisfy $f(\alpha x+\beta x', y) = \alpha f(x, y)+\beta f(x', y)$, with an analogous rule for the second argument. A bilinear map can also be viewed as a polynomial map that outputs the evaluations of n set-multilinear polynomials in variable sets $x_1, \ldots, x_n$ and $y_1, \ldots, y_n$. (In this setting, a polynomial is set-multilinear if each monomial is multilinear and contains at most one variable among $x_1, \ldots, x_n$ and at most one among $y_1, \ldots, y_n$.)

As laid out in the remainder of the introduction, bilinear maps enjoy a strong connection between computational complexity and *rank*, an algebraic complexity measure. The rank $\mathrm{R}(f)$ is

the minimum number of terms required to express f as a linear combination of rank-1 maps, where a rank-1 map is a product of linear functions in disjoint variable sets.

Rank and Kronecker Products. Several extremely well-studied bilinear maps, such as matrix multiplication and subset convolution, have a very useful structure: They can be expressed as *Kronecker products* of lower-dimensional bilinear maps. To make this precise, given vectors $u, u' \in \mathbb{F}^n$, define $u \otimes u' \in \mathbb{F}^{n^2}$ by setting $(u \otimes u')_{i,j} = u_i u'_j$ and viewing the index set $[n]^2$ as $[n^2]$ under some bijection. For bilinear maps $f, f' : \mathbb{F}^n \times \mathbb{F}^n \to \mathbb{F}^n$, the *Kronecker product* $f \otimes f' : \mathbb{F}^{n^2} \times \mathbb{F}^{n^2} \to \mathbb{F}^{n^2}$ is defined by mapping

$$(u \otimes u', v \otimes v') \mapsto f(u, v) \otimes f'(u', v'),$$

and taking the unique bilinear extension of this map. For example, the Kronecker product of the bilinear map g_n for multiplication of two $n \times n$ matrices and the map g_m for multiplication of two $m \times m$ matrices is the map $g_n \otimes g_m \equiv g_{nm}$ for multiplication of two $nm \times nm$ matrices. Likewise, subset convolution over a universe of size n (i.e., square-free multiplication of n-variate multilinear polynomials) is the Kronecker power of subset convolution over a universe of size 1.

Rank and Kronecker products interact very favorably: If f_1 and f_2 have rank $r_1 = \mathrm{R}(f_1)$ and $r_2 = \mathrm{R}(f_2)$, then $\mathrm{R}(f_1 \otimes f_2) \leq r_1 r_2$, since the rank decompositions for f_1 and f_2 can be combined to a rank decomposition for $f_1 \otimes f_2$. Unlike for linear maps, this combined rank decomposition of bilinear maps may however not be optimal, and we may have $\mathrm{R}(f_1 \otimes f_2) < r_1 r_2$. Consequently, we say that rank is *submultiplicative* under Kronecker products.

Asymptotic Rank and Circuit Complexity. The rank of bilinear maps like matrix multiplication and subset convolution of increasing dimensions can be understood by fixing a constant-dimension bilinear map $g : \mathbb{F}^n \times \mathbb{F}^n \to \mathbb{F}^n$ of rank $r = \mathrm{R}(g)$ and considering its *Kronecker power* $g^{\otimes k} = g \otimes \ldots \otimes g$, a bilinear map between spaces of dimension n^k. As discussed above, we then have $\mathrm{R}(g^{\otimes k}) \leq r^k$, but this bound may be exponentially loose: One of the simplest examples of this phenomenon is the subset convolution map $g : \mathbb{F}^2 \times \mathbb{F}^2 \to \mathbb{F}^2$ on 1-element universes, or what is essentially the same, the three-qubit W-state in quantum information theory. We have $\mathrm{R}(g) = 3$ and thus $\mathrm{R}(g^{\otimes k}) \leq 3^k$, while it is known that $\mathrm{R}(g^{\otimes k}) \leq O(2^k k)$. Combined with algorithmic ideas dating back to Yates [51], this nontrivial rank bound for $g^{\otimes k}$ translates into nontrivial $2^{k+o(k)}$ size circuits for subset convolution on k-element universes. Similar observations are crucial in the study of matrix multiplication.

This algorithmic connection between rank and algorithmic complexity of Kronecker powers motivates defining the *asymptotic rank exponent* $\omega(g)$ of a bilinear map $g : \mathbb{F}^n \times \mathbb{F}^n \to \mathbb{F}^n$ as the infimum over all $\beta > 0$ such that $\mathrm{R}(g^{\otimes k}) \leq O(n^{\beta k})$. Likewise, we define the *asymptotic circuit complexity exponent* $\eta(g)$ as the infimum over all $\beta > 0$ such that $g^{\otimes k}$ has arithmetic circuits of size $O(n^{\beta k})$. For bilinear maps, these two exponents are known to be *equal* (cf. the discussion following Theorem 8), so asymptotic circuit complexity can be understood entirely in terms of asymptotic rank, which in turn is amenable to techniques from pure mathematics.

Trivially, the asymptotic rank exponent of every bilinear map is bounded by 2, because $\mathrm{R}(g) \leq n^2$. Surprisingly however, Strassen [46] showed a bound of $2\omega/3 < 2$, where $\omega < 3$ is the asymptotic rank exponent of matrix multiplication. This translates directly to improved algorithms, e.g., for convolution problems [6]. The *asymptotic rank conjecture* [47] postulates that Strassen's upper bound can be pushed further down to 1, i.e., that every bilinear map g among n-dimensional vector spaces satisfies $\mathrm{R}(g^{\otimes k}) \leq O(n^k)$. This would lead to near-linear time algorithms for evaluating any

map of the form $g^{\otimes k}$, which in turn would directly imply that $n \times n$ matrices can be multiplied with $n^{2+o(1)}$ operations. Additionally, it was recently shown that the asymptotic rank conjecture would imply breakthrough exponential-time algorithms, e.g., for the permanent [5], for the chromatic number of graphs [3], and for the set cover problem [4, 41].

More Modes. So far, we described *bi*-linear maps, which take *two* input vectors and produce an output vector. More generally, a *multi*-linear map takes some number d of input vectors, called *modes*, and produces an output vector, and is linear in each argument. This is a natural generalization of the bilinear case, and is one of the open research directions mentioned, e.g., in the survey by Wigderson and Zuiddam [50, Sect. 13]. Beyond their intrinsic algebraic appeal, such maps also capture multi-party states in quantum information theory and important problems in fine-grained complexity.

A prominent example for a multilinear map with d modes is the *iterated matrix multiplication*, which takes as input d matrices $A_1, \ldots, A_d$ (flattened to vectors) and outputs their product $A_1 \ldots A_d$. Other examples include the *determinant* and *permanent* of an $n \times n$ matrix A, which are multilinear *forms* (rather than *maps*, i.e., they output scalars) with n modes corresponding to the columns of A. In fine-grained complexity, the k-hyperclique problem admits a natural formulation as a multilinear form with k modes [36]. The complexity of this problem is interesting even for small, fixed values of k, starting with $k = 4$, and has attracted significant attention in recent years, especially as a source of conditional hardness [33, 7, 25].

Remark. We will only speak of multilinear *forms* in the following, because maps with $d-1$ modes canonically correspond to forms with d modes by taking duals in the involved (finite-dimensional) spaces. After fixing coordinates, multilinear forms g with d modes can be viewed as set-multilinear polynomials: Denoting the entries of the input vector $\mathbf{x^{(i)}}$ at mode $i \in [d]$ by $x_1^{(i)}, \ldots, x_n^{(i)}$, we have

$$g(\mathbf{x^{(1)}}, \ldots, \mathbf{x^{(d)}}) = \sum_{i_1,\ldots,i_d} t_{i_1,\ldots,i_d} \, x_{i_1}^{(1)} \ldots x_{i_d}^{(d)}$$

with coefficients $t_{i_1,\ldots,i_d} \in \mathbb{F}$ that fully specify g. Arranging these coefficients in an array with d modes, we call this array the *tensor* of g, and overloading terminology further, we call g *itself* a tensor with d modes. After tacitly performing all required identifications, we can view bilinear maps as 3-mode tensors.

Rank versus Circuit Complexity for More Modes. Rank and circuit complexity can be generalized directly to tensors with an arbitrary number of modes. For $d \geq 4$ modes however, the direct correspondence between asymptotic circuit complexity and asymptotic rank breaks down. As a simple, folklore example, consider the tensor P_n with four n-dimensional modes defined by

$$P_n(\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{w}) = \left(\sum_{j=1}^{n} x_j y_j \right) \left(\sum_{j=1}^{n} w_j z_j \right). \tag{1}$$

In algebraic complexity, P_n is known as a product of inner products (PIP). We observe that $P_n \otimes P_m \equiv P_{nm}$, so the asymptotic rank exponent of every single P_n equals the exponent of the monotonically increasing sequence $\mathrm{R}(P_n)$ as $n \to \infty$. The defining formula for P_n of size $O(n)$

implies an asymptotic circuit complexity exponent of 1, while the asymptotic rank exponent of P_n is 2: Multiplying out (1) gives an upper bound of n^2, matched by a simple flattening lower bound.

Thus, unlike for $d = 3$ modes, asymptotic rank and asymptotic circuit complexity of 4-mode tensors for $d \geq 4$ need not agree. The separation can be exacerbated by generalizing the PIP tensor to a fixed number of $2d$ modes, which increases the asymptotic rank exponent to d, while the asymptotic circuit complexity exponent stays 1. As the example shows, the separation holds even for $\Pi\Sigma\Pi$-formula complexity, which morally still resembles rank in that rank decompositions can be viewed as $\Sigma\Pi\Sigma$-formulæ.

1.1 Our Results: Phenomena of Tensors with More Modes

Starting from the observation that rank and circuit complexity need not agree for relevant tensors with $d \geq 4$ modes, we are led to systematically study asymptotic rank and asymptotic circuit complexity for d-mode tensors when $d \geq 3$ is fixed and moderately large or may tend to infinity. As part of this study, we obtain better upper bounds on asymptotic rank and asymptotic circuit complexity for generic d-mode tensors. Moreover, we investigate to which extent desirable properties like submultiplicativity fail for the circuit complexity of d-mode tensors with $d \geq 4$. The foundation to our results are *graph tensors* in the form used by Christandl and Zuiddam [14] and Christandl, Vrana and Zuiddam [13], which allow us to import techniques from graph theory.

We elaborate on these tensors in Section 1.2; for now, it suffices to note that the modes of the graph tensor T_H correspond to the vertices of a graph H. We also shortly remark that graph tensors can be viewed as a special class of tensors that admit representation as a *tensor network*. Tensor networks have received extensive attention across a range of disciplines, each pursuing its own distinctive terminology and notation, from pure mathematics and physics [34, 38, 40] to the study of probabilistic graphical models in artificial intelligence and machine learning [31, 32, 39], and to the study of restricted circuit models in algebraic computation, where tensor networks give rise to a class of set-multilinear circuits for evaluating multilinear maps [1].

Generic Asymptotic Rank. To our surprise and best of knowledge, an analogue of Strassen's $2\omega/3$ bound on the asymptotic exponent of 3-mode tensors for d-tensors with $d \geq 4$ is not mentioned in the literature. We give a simple proof of Strassen's bound in the framework of graph tensors that yields an upper bound of $(d-1)\omega/3$ on the asymptotic exponent of all d-mode tensors for $d \geq 3$.

Our proof admits some freedom in the choice of low-rank tensor L to reduce to: While Strassen's original proof requires L to be the matrix multiplication tensor, which will turn out to be the graph tensor of a triangle K_3, our proof more generally allows L to be the graph tensor of the complete graph K_s for $s \leq d$. In the case $d = 3$, we recover the aforementioned bound of $(d-1)\omega/3$, but for $d \geq 4$, better upper bounds on the exponent of the graph tensor of K_d are known and directly yield better upper bounds. Moreover, with a nontrivial application of Strassen's *laser method*, we improve the best known exponent on the graph tensor for K_4 and obtain:

Theorem 1 (Upper bound on asymptotic rank)**.** *For every d-mode tensor with $d \geq 4$, the asymptotic rank exponent of T is at most* $0.772318(d-1)$.

For comparison, the PIP tensor presented above (1) establishes a lower bound of $\lfloor d/2 \rfloor$ on the rank exponent of d-mode tensors. Moreover, $\omega = 2$ would imply an upper bound of $2/3 \cdot (d-1)$. This is elaborated more formally in Remark 21.

Submultiplicativity. While the PIP tensor rules out a functional dependence between rank and circuit complexity for $d \geq 4$ modes, we ask whether circuit complexity is at least *submultiplicative*: From *rank* decompositions for tensors T and U, we can explicitly construct a *rank* decomposition for $T \otimes U$ of rank $\mathrm{R}(T) \cdot \mathrm{R}(U)$. If T and U are given by circuits of small size, can we analogously expect a circuit of small size for $T \otimes U$? Restricted circuit models are known to have submultiplicativity under Kronecker powering; in particular, a set-multilinear circuit arising from a tensor network and its contraction tree for a tensor T gives rise to a tensor network and its contraction tree for the Kronecker power $T^{\otimes k}$, with submultiplicativity of amortized cost [1]. But does *unrestricted* circuit complexity have submultiplicativity?

Our results strongly suggest a negative answer, but we can obtain such statements only *conditionally*, since an unconditional result would imply strong circuit lower bounds. In the following, we write $\mathrm{C}(T)$ for the minimum size of an arithmetic circuit that computes the tensor T.

Theorem 2 (Submultiplicativity of C implies $\mathsf{VP} = \mathsf{VNP}$)**.** *There are explicit tensors $T_1, T_2, \ldots$ and $U_1, U_2, \ldots$ with $T_d, U_d \in (\mathbb{C}^2)^{\otimes d}$ for all $d \in \mathbb{N}$ such that the following holds: If $\mathrm{C}(T_d \otimes U_d) \leq \mathrm{poly}(d, \mathrm{C}(T_d), \mathrm{C}(U_d))$ for all $d \in \mathbb{N}$, then $\mathsf{VP} = \mathsf{VNP}$.*

Thus, the assumption $\mathsf{VP} \neq \mathsf{VNP}$ rules out submultiplicativity on d-mode tensors for all $d \in \mathbb{N}$. Note that $\mathsf{VP} \neq \mathsf{VNP}$ implies super-polynomial circuit lower bounds for permanents [48]. If we assume even further that permanents require exponential-size circuits, then we can rule out submultiplicativity on d-mode tensors for concrete values of d.

Theorem 3 (Submultiplicativity of C implies faster permanents)**.** *For all $0 < c < 1$, if circuit complexity is submultiplicative on $\lceil 4/c \rceil^2$-mode tensors, then permanents have circuits of size $O(2^{cn})$.*

In particular, unless the permanent has circuits of size $2^{o(n)}$, there is a constant $D \in \mathbb{N}$ such that C is not submultiplicative for D-mode tensors. This assumption on the permanent is implied by a suitable non-uniform variant of the *exponential-time hypothesis*. For a more fine-grained approach, we can consider concrete values of $c > 0$ in Theorem 3. For example, if the permanent does not have circuits of size $2^{0.1n}$, then submultiplicativity fails for tensors with $d = \lceil 4/0.1 \rceil^2 = 40^2 = 1600$ modes. If even circuits of size $2^{0.8n}$ can be ruled out for the permanent, then submultiplicativity fails already on $d = 25$ modes; this is the smallest number d of modes for which Theorem 3 still yields a statement. As for the plausibility of this assumption, Knuth asks in *The Art of Computer Programming* [30, Volume 2, §4.6.4, Problem 11] whether there is any way to evaluate the permanent of a general $n \times n$ matrix using fewer than 2^n arithmetic operations. This is still an open problem. We remark that there is no asymptotic notation around 2^n.[1]

Using another conjecture from fine-grained complexity theory, we come closer to $d = 4$, the smallest number of modes where rank and circuit complexity differ: The (h, k)-*hyperclique conjecture* [36] rules out $O(n^{k-\varepsilon})$-time algorithms for detecting k-hypercliques in h-uniform hypergraphs, and it is easy to formulate a non-uniform algebraic variant of it (see Section 4 for details).

Theorem 4 (Submultiplicativity of C implies faster hypercliques)**.** *For every even $d \geq 8$, the non-uniform $(\frac{d}{2} - 1, \frac{d}{2})$-hyperclique conjecture rules out submultiplicativity of circuit complexity on d-mode tensors.*

[1] Knuth's question is currently known to admit an affirmative answer only under the asymptotic rank conjecture [5]; more specifically, the asymptotic rank conjecture implies uniform arithmetic circuits of size 2^{cn} for the permanent where $c \approx 0.9183$.

While our results suggest that circuit complexity is not submultiplicative for d-mode tensors, we can construct small circuits for $T \otimes U$ from a low-rank decomposition of T and a small circuit for U. This proves useful for us in obtaining better upper bounds on the asymptotic circuit complexity.

Theorem 5 (Mixed asymptotic rank and circuit complexity)**.** *Given d-mode tensors T and U, where T has asymptotic rank exponent $r \geq 1$ and U has asymptotic circuit complexity exponent $s \geq 1$, the asymptotic circuit complexity exponent of $T \otimes U$ is at most $r + s$.*

Asymptotic Circuit Complexity. With tools like Theorem 5 and insights from graph theory, we adapt the graph-theoretic proof of our asymptotic rank bound in Theorem 1 to obtain stronger upper bounds for asymptotic circuit complexity rather than rank.

First, using a known nontrivial upper bound on the treewidth of line graphs of complete graphs [26], we obtain an upper bound on the circuit complexity of generic d-mode tensors:

Theorem 6 (Upper bound on asymptotic circuit complexity)**.** *The asymptotic circuit complexity exponent of every d-mode tensor T is less than $d/2 + 1$.*

While a lower bound of $d/2$ on the asymptotic *rank* exponent of d-mode tensors is immediate from (1), an unbounded *circuit complexity* exponent in d would constitute a strong circuit lower bound and hence seems exceedingly unlikely to be obtainable.

Next, we observe that Theorem 6 does not yield optimal upper bounds for small fixed d. For example, the theorem gives a circuit complexity exponent bound of 3 for $d = 4$, but 3 is even above the asymptotic *rank* upper bound of 2.317. To obtain meaningful upper bounds for small fixed d, we combine Theorem 5 with fractional decompositions of specific small graphs into low-rank and low-treewidth parts. These graph decompositions are found through computer-aided optimization. Our technique is applicable for arbitrary fixed d, albeit with higher computational load, and we obtain exemplary results for $d = 4, 5$:

Theorem 7 (Upper bound on asymptotic circuit complexity for small orders)**.** *For every 4-mode tensor T, the asymptotic circuit complexity exponent of T is at most 2.2967. For every 5-mode tensor, it is at most 2.8774.*

For comparison, the upper bounds on the asymptotic *rank* exponents from Theorem 1 for $d = 4, 5$ are about 2.317 and 3.0893, respectively. Table 1 provides a comprehensive overview of the kinds of results we obtain.

1.2 Our Techniques: New Insights into Graph Tensors

The protagonists in our proofs are so-called *graph tensors*, i.e., tensors with many modes that are composed from 2-mode tensors by using graphs as "composition blueprints". Such tensors appear in quantum information theory as *matrix-product states* and *projected entangled pairs* [15], in quantum machine learning as (restricted) *tensor networks* [28], and in counting complexity (under the right abstraction) as *Holant problems* [49, 11, 12]. For us, graph tensors act as a bridge between graph theory and multilinear algebra that allows us to import nontrivial graph-theoretic concepts (e.g., graph decompositions and treewidth) and results (e.g., bounds on treewidth for specific graph classes) to the study of tensors with many modes. Moreover, graph tensors are closed under Kronecker products, which makes them particularly well-suited for studying the asymptotic behaviour of rank and circuit complexity. We remark that graph tensors can be generalized further to tensors

Method	$d=3$	$d=4$	$d=5$	$d=6$	$d=10$
$\underset{\sim}{\mathrm{R}}(T)$, asymptotic submult. (Thm. 1)	1.59	2.32	3.09	3.87	6.96
$\underset{\sim}{\mathrm{C}}(T)$, treewidth-based (Thm. 31)	2.00	2.50	3.20	3.67	5.80
$\underset{\sim}{\mathrm{C}}(T)$, specialized decompositions (Thm. 40, 41)	–	2.30	2.88	–	–
Flattening lower bound on $\underset{\sim}{\mathrm{R}}(T)$ from (1)	1.00	2.00	2.00	3.00	5.00

Table 1: Bounds on asymptotic rank and circuit complexity exponents of generic d-mode tensors. The first row states rank upper bounds obtained via Theorem 1. The next two rows state upper bounds on asymptotic circuit exponents. For comparison, we include the known case $d = 3$ and the strongest known lower bounds.

admitting representation by more general tensor networks or factor-graph models (e.g. [1, 32]), but graph tensors will be sufficient for our present purposes.

As a simple example, the outer tensor product of k tensors on b modes gives a tensor T with kb modes. In the extreme case of $b = 1$, the resulting tensor T has rank 1, but already for $b = 2$, examples of large rank such as the PIP tensor (1) emerge. Graph tensors are obtained by additionally allowing identification (with flattening) of modes. Such tensors admit natural interpretations both as quantum states and as generic instances of Holant problems; we discuss both in the following.

Graph Tensors in Terms of Quantum States

In quantum information theory, d-mode tensors are the pure states of d-party systems. In the following, let $e_1^{(i)}, \ldots, e_r^{(i)}$ be a basis of the r-dimensional local state space of system $i \in [d]$. A canonical example, the *Greenberger–Horne–Zeilinger* state (equivalently, the *unit tensor* on d modes and dimension r per mode), is defined by

$$U_{d,r} \coloneqq \sum_{s=1}^{r} e_s^{(1)} \otimes \ldots \otimes e_s^{(d)}.$$

It represents d parties of local dimension r in genuine multipartite entanglement, and every d-mode tensor of rank r can be obtained as a projection from $U_{d,r}$. The two-party case $U_{2,r}$ is known as a generalized *Bell state*.

In important applications, the global d-party state decomposes into states on collections of b parties for $b < d$; these collections may overlap nontrivially. We focus on the case $b = 2$, i.e., on states that decompose into local Bell states. These so-called *graph tensors* are fully determined by entanglements between *pairs* of parties and can be described by an undirected graph H with d vertices, where an edge uv is present if parties u and v share a Bell state. Formally, we have

$$T_{H,n} \coloneqq \bigotimes_{uv \in E(H)} \left(\sum_{s=1}^{r} e_s^{(u)} \otimes e_s^{(v)} \right). \tag{2}$$

As a concrete example, assume the d parties correspond to vertices of a path P with entanglement between adjacent parties i and $i+1$ for $1 \leq i < d$. This gives rise to the global state

$$T_{P,n} = \left(\sum_{s=1}^{r} e_s^{(1)} \otimes e_s^{(2)} \right) \otimes \left(\sum_{s=1}^{r} e_s^{(2)} \otimes e_s^{(3)} \right) \otimes \ldots \otimes \left(\sum_{s=1}^{r} e_s^{(d-1)} \otimes e_s^{(d)} \right). \tag{3}$$

A more "computational perspective" on graph tensors is also possible, and it will be more useful when studying their complexity. Namely, by distributing the Kronecker product over the $d-1$ pairs of parentheses in (3), we can also interpret $T_{P,n}$ as a sum over r-ary assignments $s\colon [d-1] \to [r]$ to the $d-1$ edges between adjacent parties:

$$T_{P,n} = \sum_{s\colon [d-1]\to[r]} e^{(1)}_{s(1)} \otimes \left(e^{(2)}_{s(1)} \otimes e^{(2)}_{s(2)}\right) \otimes \ldots \otimes \left(e^{(d-1)}_{s(d-1)} \otimes e^{(d-1)}_{s(d)}\right) \otimes e^{(d)}_{s(d)}. \tag{4}$$

From (4), we see that the vector space at modes 1 and d is isomorphic to $\mathbb{C}^r$, while it is isomorphic to $(\mathbb{C}^r)^{\otimes 2}$ at all other modes. In other words, if $I(v)$ denotes the edges incident with mode v, then the space at mode v admits a basis of vectors $e^{(v)}_a$ that are indexed by *local* assignments of the form $a\colon I(v) \to [r]$. Applying the same reasoning to general graph tensors in (2), we obtain:

$$T_{H,n} = \sum_{s\colon E(H)\to[r]} \bigotimes_{v\in V(H)} e^{(v)}_{s|_{I(v)}}, \tag{5}$$

where $s|_{I(v)}$ is the restriction of s to the edges $I(v)$ incident with vertex $v \in V(H)$.

Graph Tensors in Terms of Holant Problems

To analyze the complexity of tensors $T_{H,n}$ for graphs H, we observe that they can be interpreted as *Holant problems*, which are very well-studied in counting complexity [49, 11, 9, 10, 45]: On input a graph $G = (V, E)$, a Holant problem asks to compute a weighted count of edge-assignments $s\colon E \to [r]$, with weights determined locally at vertices; the vast majority of the literature focuses on the case $r = 2$. As a concrete example, the number of *perfect matchings* in G is the number of edge-assignments $s\colon E(H) \to \{0, 1\}$ such that every vertex $v \in V$ has exactly one 1-labeled edge in the local restriction $s|_{I(v)}$ of the global assignment s.

In general Holant problems, the weights are determined by *signatures* $f_v\colon [r]^{I(v)} \to \mathbb{C}$ at the vertices $v \in V$. On input G and signatures $\{f_v\}_{v\in V}$, we then wish to determine

$$\text{Holant}(G) = \sum_{s\colon E\to[r]} \prod_{v\in V} f_v(s|_{I(v)}). \tag{6}$$

While the analogy between (6) and $T_{G,r}$ from (5) is immediate, Holant problems have been studied from a very different perspective in the literature: Usually, a set of possible signatures $\mathcal{F}$ is fixed, and the input is a graph G with signatures $f_v \in \mathcal{F}$. We consider the converse setting: The graph G is fixed, but the signatures can vary freely in that they are provided as input vectors to the $|V(G)|$ modes of the multilinear form $T_{G,n}$.

As an example of particular relevance for us, fix G as the 4-regular toroidal grid on $t \times t$ vertices. Then $T_{G,2}$ is a t^2-mode tensor with a copy of $(\mathbb{C}^2)^{\otimes 4}$ at each mode. By specifying a vertex signature $f_v : \{0, 1\}^4 \to \mathbb{C}$ for each grid vertex and inputting it as a vector $u_v \in \mathbb{C}^{16}$ into the multilinear form $T_{G,2}$, we can count perfect matchings, Eulerian subgraphs, or evaluate any Holant problem definable on the fixed grid G. In particular, for growing grid sizes, this includes VNP-hard problems: By appropriately engineering the signatures, we express the permanent of arbitrary $n \times n$ matrices as a particular Holant problem on the fixed $n \times n$ *grid*, and we obtain hardness of the grid graph tensor. On the algorithmic side, known fixed-parameter tractable algorithms over low-treewidth graphs establish a nontrivial complexity upper bound of $n^{O(t)}$ on the evaluation of $T_{G,n}$, where our generic results would only yield an $n^{O(t^2)}$ bound. Such upper and lower complexity bounds informed by (parameterized) algorithmic graph theory are ubiquitous in our arguments.

2 Preliminaries

For a nonnegative integer n we write $[n] = \{1, 2, \ldots, n\}$. We write $\mathbb{N} = \{1, 2, \ldots\}$. The Kronecker delta δ_{ij} is 1 if $i = j$ and 0 else.

2.1 Algebraic Complexity

Tensors. It will be convenient to work with tensors in coordinates and view tensors as set-multilinear polynomials with an underlying tuple of modes. Let $\mathbb{F}$ be a field. We write $\mathbb{F}[X]$ for the ring of polynomials over $\mathbb{F}$ in a tuple X of one or more indeterminates. We assume the indeterminates X are partitioned into a tuple of d pairwise disjoint nonempty sets $X_1, X_2, \ldots, X_d$ called *modes*. A monomial is *set-multilinear* if each mode has exactly one indeterminate that has degree one in the monomial, and all other indeterminates have degree zero. A polynomial is *set-multilinear* if all of its monomials are set-multilinear. A *d-mode tensor* or briefly a *d-tensor* T is a set-multilinear polynomial together with its d-tuple $(X_1, X_2, \ldots, X_d)$ of modes. The *shape* of T is $|X_1| \times |X_2| \times \cdots \times |X_d|$. We also call d the *degree* or *order* of the tensor, in line with our polynomial interpretation. We suppress d and the modes whenever they are clear from the context. For brevity, we will even suppress the variable names themselves and write $(\mathbb{F}^n)^{\otimes d}$ for the set of d-tensors over $\mathbb{F}$ with each mode of dimension n.

Addition, subtraction, and scalar multiplication of tensors with identical tuples of modes are inherited from polynomial arithmetic. The *Kronecker product* of two d-tensors $S \in \mathbb{F}[X]$ and $T \in \mathbb{F}[Y]$ with X, Y disjoint is the d-tensor $S \otimes T$ obtained from the polynomial product ST by viewing the Cartesian products $X_i \times Y_i$ for $i \in [d]$ as the d modes. When the sets X, Y are not disjoint, such as when taking a Kronecker power of a tensor, we tacitly assume a disjoint copy of one of the sets of indeterminates is formed before taking the Kronecker product.

Projection, Equivalence, and Restriction. A *simple substitution* for sets of indeterminates X, Y is a map $\sigma : Y \to \mathbb{F} \cup X$. For tensors $S \in \mathbb{F}[X]$ and $T \in \mathbb{F}[Y]$ we say that S is a *projection* of T and write $S \leq_{\mathrm{s}} T$ if there exists a simple substitution σ such that the polynomial identity $S = T^\sigma$ holds and the image of each mode Y_j of T under σ intersects at most one mode X_i of S. A simple substitution is *nonscalar* if it does not assume values in $\mathbb{F}$. We say that S and T are *equivalent* and write $S \equiv T$ if the polynomial identity $S = T^\sigma$ holds for a bijective nonscalar σ that is bijective on modes. For d-tensors S and T we say that T *restricts to* S and write $S \leq T$ if the polynomial identity $S = T^\mu$ holds for a substitution $\mu : Y \to \mathbb{F}[X]$ such that for all for all modes Y_j and all indeterminates $y \in Y_j$ the polynomial $\mu(y)$ is a linear polynomial in the indeterminates X_j.

Tensor Rank. A nonzero d-tensor T has *rank one* if there exist linear polynomials $\ell_j \in \mathbb{F}[X_j]$ for $j \in [d]$ such that the polynomial identity $T = \ell_1 \ell_2 \cdots \ell_d$ holds. The *rank* $\mathrm{R}(T)$ of a d-tensor T is the minimum number of rank one tensors whose sum is T. (Note that $d = 2$ recovers matrix rank.) For two d-tensors S and T, we have $\mathrm{R}(S \otimes T) \leq \mathrm{R}(S) \cdot \mathrm{R}(T)$, with equality for $d = 2$ but strict inequality may hold for $d \geq 3$. This property is commonly referred to as the *submultiplicativity* of tensor rank. Moreover, the rank of any tensor $T \in (\mathbb{F}^n)^{\otimes d}$ is obviously bounded by n^d, the number of set-multilinear monomials (and the stronger bound n^{d-1} can easily be shown). Consequently, $T^{\otimes k}$ has rank at most $n^{(d-1)k}$. The *asymptotic rank* of $T \in (\mathbb{F}^n)^{\otimes d}$ is defined as the limit $\underset{\sim}{\mathrm{R}}(T) = \lim_{k \to \infty} \mathrm{R}(T^{\otimes k})^{1/k}$, which exists by submultiplicativity of rank and Fekete's Lemma (the result is folklore, but see [50, Lemma 2.10] and the references given there). Like rank, also the

asymptotic rank satisfies submultiplicativity with respect to Kronecker products. We say that a d-tensor is *concise* if all of its d flattenings to a matrix with one mode forming the rows and all the other modes forming the columns have full rank. A concise tensor $T \in (\mathbb{F}^n)^{\otimes d}$ in particular satisfies both $\mathrm{R}(T) \geq n$ and $\underset{\sim}{\mathrm{R}}(T) \geq n$. The *exponent* of a concise tensor $T \in (\mathbb{F}^n)^{\otimes d}$ is defined by

$$\omega(T) = \inf \{\beta \geq 0 : \mathrm{R}(T^{\otimes k}) = O(n^{\beta k})\}$$

and we have

$$\underset{\sim}{\mathrm{R}}(T) = n^{\omega(T)}\,.$$

For example, the tensor $\mathrm{MM}_m \in \mathbb{F}^{m^2} \otimes \mathbb{F}^{m^2} \otimes \mathbb{F}^{m^2}$ that represents the bilinear map that multiplies two $m \times m$ matrices for any constant $m \geq 2$ is concise and recovers the exponent ω of square matrix multiplication either via the asymptotic rank identity $\underset{\sim}{\mathrm{R}}(\mathrm{MM}_m) = m^\omega$ or, equivalently, via the tensor exponent identity $\omega(\mathrm{MM}_m) = \omega/2$.

Arithmetic Circuits. Tensor rank fails to capture the arithmetic complexity of a tensor for $d \geq 4$, see Example 11 below. In contrast, circuits provide a more faithful model of arithmetic complexity. An *arithmetic circuit over* $\mathbb{F}[X]$ is a directed acyclic graph (DAG) with sinks called *outputs*, sources called *inputs*, vertices called *gates*, and directed edges called *wires.* Non-input gates are labeled with $+$ and $\times$, while inputs are labeled with elements from $\mathbb{F} \cup X$, and wires are labeled with elements of $\mathbb{F}$. A gate *computes* a polynomial in the obvious inductive manner, and the *size* of a circuit is the number of its wires. For a tensor $T \in (\mathbb{F}^n)^{\otimes d}$, we write $\mathrm{C}(T)$ for the size of a smallest circuit computing T, called the *arithmetic circuit complexity* of T.

A rank-r decomposition of $T \in (\mathbb{F}^n)^{\otimes d}$ can be converted to an arithmetic circuit for T of size $O(d \cdot n \cdot r)$. Hence, the trivial bound $\mathrm{C}(T^{\otimes k}) \leq O(d \cdot n^{(d+1)k})$ allows us to define, in analogy with the tensor case, the *circuit exponent* $\eta(T)$ and the *asymptotic circuit complexity* $\underset{\sim}{\mathrm{C}}(T)$ of T via

$$\eta(T) = \inf\{\beta \geq 0 : \mathrm{C}(T^{\otimes k}) \leq O(n^{k\beta})\}, \quad \underset{\sim}{\mathrm{C}}(T) = n^{\eta(T)}\,.$$

We note here that, again, $\underset{\sim}{\mathrm{C}}(T) = \limsup_{k\to\infty} \mathrm{C}(T^{\otimes k})^{1/k}$ can be shown, but it is not clear whether the limit itself exists, as it does for rank.

The crude upper bound of $\mathrm{C}(T) = O(d \cdot n \cdot \mathrm{R}(T))$ only proves $\eta(T) \leq \omega(T) + 1$. However, a classic result of Yates [51] (see also Knuth [30, §4.6.4]) shows that for $d \geq 2$ and concise $T \in (\mathbb{F}^n)^{\otimes d}$, we have that for all $k \geq 1$,

$$\mathrm{C}(T^{\otimes k}) = O(dk \cdot \mathrm{R}(T)^{k+1}). \tag{7}$$

All treatments in the literature we are aware of focus on the cases $d = 2, 3$ of Kronecker powers of matrices and 3-tensors, but the result easily generalizes to $d \geq 4$; for completeness we include a short exposition in Appendix B and record the asymptotic conclusion in the following theorem.

Theorem 8 (Yates's algorithm)**.** *Let $d \geq 2$ and let $T \in (\mathbb{F}^n)^{\otimes d}$ be concise. Then,*

$$\eta(T) \leq \omega(T), \text{ and therefore } \underset{\sim}{\mathrm{C}}(T) \leq \underset{\sim}{\mathrm{R}}(T).$$

Strikingly, for $d = 3$, it is true that $\underset{\sim}{\mathrm{C}}(T) \geq \underset{\sim}{\mathrm{R}}(T)$ holds as well, providing a justification for considering tensor rank as a measure of arithmetic complexity. As mentioned, this behavior does not extend to $d \geq 4$.

Closure Properties. Since restrictions of rank-one tensors are themselves of rank one (or identically zero), $\mathrm{R}(S) \leq \mathrm{R}(T)$ whenever $S \leq T$, and hence $\mathrm{R}(S) = \mathrm{R}(T)$ if $S \equiv T$. The same statements hold for asymptotic rank. For circuits, it is only true that $\mathrm{C}(S) \leq \mathrm{C}(T)$ when $S \leq_{\mathrm{s}} T$ for $S, T \in (\mathbb{F}^n)^{\otimes d}$ while $S \leq T$ only implies $\mathrm{C}(S) \leq \mathrm{C}(T) + n^{d+1}$. Still, asymptotic circuit complexity again admits $\underset{\sim}{\mathrm{C}}(S) \leq \underset{\sim}{\mathrm{C}}(T)$ if $S \leq T$ as a consequence of Yates's algorithm (7), see Fact 45 in the appendix. Moreover, we do have that $\mathrm{C}(U) \leq \mathrm{C}(U \otimes V)$ provided $V \neq 0$, see Fact 46. The questions of whether and when $\underset{\sim}{\mathrm{C}}$ and C are robust with respect to Kronecker products will be a main theme of the article.

A very general description of these properties in relation to tensor rank (and far beyond) is laid out in the work of Wigderson and Zuiddam [50]. For further background on algebraic complexity we refer to the book of Bürgisser, Clausen and Shokrollahi [8].

2.2 Graph Theory

Graphs. A *graph* G consists of a finite set $V(G)$ of *vertices* and a finite set $E(G)$ of *edges* such that each edge is associated with a set of two vertices called the *end-vertices* or *ends* of the edge. We stress that multiple edges may have the same set of ends; such edges are called *parallel* edges. For a vertex $v \in V(G)$ we write $I_G(v) \subseteq E(G)$ for the set of all edges that have v as an end, and call $|I_G(v)|$ the *degree* of v. The maximum degree of a vertex in G is denoted by $\Delta(G)$.

For two graphs G and H with disjoint edge sets, the *sum* $G + H$ is the graph defined by $V(G + H) = V(G) \cup V(H)$ and $E(G + H) = E(G) \cup E(H)$. When the edge sets are not disjoint, such as when taking the sum $G + G$, we tacitly assume a disjoint copy of one of the edge sets is formed before taking the sum.

We say that two graphs G and H are *isomorphic* and write $G \equiv H$ if there exist bijections $\phi\colon V(G) \to V(H)$ and $\psi\colon E(G) \to E(H)$ such that each edge $e \in E(G)$ has the end-vertices $\{v_1, v_2\} \subseteq V(G)$ if and only if the edge $\psi(e)$ has the end-vertices $\{\phi(v_1), \phi(v_2)\} \subseteq V(H)$. Equivalently, $\psi(I_G(v)) = I_H(\phi(v))$ should hold.

Subgraphs. A graph H is a *subgraph* of a graph G if $V(H) \subseteq V(G)$, $E(H) \subseteq E(G)$, and common edges of H and G have identical end-vertices. A *subdivision* of a graph G is a graph obtained by applying the following operation zero or more times: select an edge e with end u and v, insert a new vertex w, delete the edge e, insert a new edge with ends u and w, and insert a new edge with ends v and w. A graph H is a *topological subgraph* of a graph G if there exists a subdivision of H that is isomorphic to a subgraph of G.

Treewidth. A *tree decomposition* of a graph G is a pair (T, β) where T is a tree and β is a map that associates each vertex $a \in V(T)$ with a *bag* $\beta(a) \subseteq V(G)$ such that (i) the union of all bags is $V(G)$; (ii) both ends of each edge of G are contained in at least one bag; and (iii) for all $a, c \in V(T)$ it holds that $\beta(a) \cap \beta(c) \subseteq \beta(b)$ for all $b \in V(T)$ on the path joining a and c in T. The *width* of (T, β) is $\max_{a \in V(T)} |\beta(a)| - 1$. The *treewidth* $\mathrm{tw}(G)$ is the minimum width of a tree decomposition of G. More background on algorithmic aspects on graph theory, in particular algorithms exploiting tree decompositions of small width, can be found in the textbook of Cygan et al. [19].

3 Graph Tensors and the Asymptotic Rank

This section develops our conventions for graph tensors together with their basic algebraic properties in relation to the underlying graphs, and then proceeds to prove our main theorem (Theorem 1) on bounding the asymptotic rank for d-mode tensors for $d \geq 4$, extending Strassen's result for $d = 3$.

3.1 Graph Tensors

Our graphs are undirected, loopless, and may have parallel edges. Let G be a graph and let $n \in \mathbb{N}$. For a mapping $f\colon E(G) \to [n]$ and a vertex $v \in V(G)$, let us write $f|_{I(v)}$ for the restriction of f to the set $I(v)$ of edges incident to v in G. For each vertex $v \in V(G)$, introduce a mode $X^{(v)} = \{x_g^{(v)} \mid g\colon I(v) \to [n]\}$.

Definition 9 (Graph tensor)**.** The *graph tensor* $T_{G,n}$ of the graph G with *length parameter* n has the modes $(X^{(v)} : v \in V(G))$ and is defined by the polynomial identity

$$T_{G,n} = \sum_{f\colon E(G)\to[n]} \prod_{v\in V} x^{(v)}_{f|_{I(v)}}. \tag{8}$$

Note that $T_{G,n}$ has $|V(G)|$ modes and the *length* of the mode $X^{(v)}$ for $v \in V(G)$ is $|X^{(v)}| = n^{|I(v)|}$. Graph tensors are concise.

Remark 10 (The Holant view)**.** As laid out in the introduction, graph tensors and Holant problems are intimately connected. Namely, the tensor $T_{G,n}$ essentially captures all Holant problems definable on the fixed graph G by choosing concrete signatures $s_v : [n]^{I(v)} \to \mathbb{C}$ for the vertices $v \in V(G)$ and substituting $x_a^{(v)} \leftarrow s_v(a)$ for all $a \in [n]^{I(v)}$.

Example 11 (The k-matching tensor)**.** Let us write H_k for the *k-matching graph* consisting of vertices u_i and v_i joined by an edge for $i \in [k]$ and let $n \in \mathbb{N}$. The *k-matching tensor* $T_{H_k,n}$ is known as the *product of inner products* in algebraic complexity theory, and in a quantum-theoretic interpretation it corresponds to n disjoint pairs of parties such that each pair holds a Bell state, and no further entanglement is present. For our purposes, the tensor $T_{H_k,n}$ for $k \geq 2$ provides an example of the separation of algebraic complexity and tensor rank for tensors with at least four modes. Namely, we observe the polynomial identity

$$T_{H_k,n} \equiv \sum_{f\colon [k]\to[n]} \prod_{i\in[k]} x^{(u_i)}_{i\mapsto f(i)} x^{(v_i)}_{i\mapsto f(i)} = \prod_{i\in[k]} \sum_{j\in[n]} x^{(u_i)}_{i\mapsto j} x^{(v_i)}_{i\mapsto j}\,,$$

where the last formula establishes $\mathrm{C}(T_{H_k,n}) \leq 2kn$, yet we have $\mathrm{R}(T_{H_k,n}) = n^k$. To see that the tensor rank is at most n^k, observe the first formula above; to see that the rank is at least n^k, recall that the tensor rank of a tensor is bounded from below by the matrix rank of any flattening of the tensor into a matrix, and study the $n^k \times n^k$ matrix flattening of $T_{H_k,n}$ defined by the modes $u_1, \ldots, u_k$ and $v_1, \ldots, v_k$. In particular observe that the flattening is, up to permutation of rows and columns, the $n^k \times n^k$ identity matrix.

Several useful connections between algebraic properties of $T_{G,n}$ and graph-theoretic properties of G can be shown. Crucially for us, the Kronecker product of graph tensors of graphs G and H is the graph tensor of the *sum graph* $G + H$, as defined in the preliminaries. (Recall that the sum may introduce multiedges.) The proof is elementary.

Lemma 12 (Product of graph tensors corresponds to graph sum)**.** *Let G and H be graphs and let $n \in \mathbb{N}$. Then,*

$$T_{G,n} \otimes T_{H,n} \equiv T_{G+H,n}\,.$$

Proof. Let us set $I_H(v) = \varnothing$ for all $v \in V(G) \setminus V(H)$, and, symmetrically, $I_G(v) = \varnothing$ for all $v \in V(H) \setminus V(G)$. By bilinearity of the Kronecker product and the definition of the sum $G + H$, we have

$$\begin{aligned}
T_{G,n} \otimes T_{H,n} &= \left(\sum_{f\colon E(G)\to[n]} \prod_{v\in V(G)} x^{(v)}_{f|_{I_G(v)}} \right) \otimes \left(\sum_{f'\colon E(H)\to[n]} \prod_{v\in V(H)} x^{(v)}_{f'|_{I_H(v)}} \right) \\
&= \sum_{\substack{f\colon E(G)\to[n] \\ f'\colon E(H)\to[n]}} \left(\prod_{v\in V(G)} x^{(v)}_{f|_{I_G(v)}} \right) \otimes \left(\prod_{v\in V(H)} x^{(v)}_{f'|_{I_H(v)}} \right) \\
&\equiv \sum_{\substack{f\colon E(G)\to[n] \\ f'\colon E(H)\to[n]}} \prod_{v\in V(G)\cup V(H)} x^{(v)}_{f|_{I_G(v)},f'|_{I_H(v)}} \\
&\equiv \sum_{f\colon E(G+H)\to[n]} \prod_{v\in V(G+H)} x^{(v)}_{f|_{I_{G+H}(v)}} \\
&= T_{G+H,n}.
\end{aligned}$$

□

Remark 13 (Single-edge decomposition)**.** Christandl and Zuiddam [14] used a special case of Lemma 12 to *define* $T_{G,n}$. Specifically, let $G = (V, E)$ with $E = \{e_1, \ldots, e_m\}$, and let $F_i = (V, \{e_i\})$ for $i \in [m]$ be the graph on vertex set V that contains only e_i as edge. Then $G = F_1 + \ldots + F_m$ and

$$T_{G,n} \equiv T_{F_1,n} \otimes \ldots \otimes T_{F_m,n}. \tag{9}$$

Lemma 12, which was not stated in their paper, then follows by commutativity of $\otimes$ up to relabeling of modes. Their definition in terms of sums of single-edge graphs is arguably more elegant, but our definition in terms of the explicit monomial expansion of $T_{G,n}$ in (8) enables a direct connection to algebraic complexity, as already witnessed for the k-matching tensor in Example 11.

Remark 14 (Length rule and sum rule)**.** With the same proof as in Lemma 12, for $n_1, n_2 \in \mathbb{N}$ we have the length rule

$$T_{G,n_1} \otimes T_{G,n_2} \equiv T_{G,n_1 \cdot n_2}\,.$$

Indeed, the set of pairs of mappings (f_1, f_2) with $f_i : E(G) \to [n_i]$ is in bijective correspondence with the set of mappings $E(G) \to [n_1 \cdot n_2]$. Consequently, writing $2 \cdot G = G + G$ for any graph G, and more generally $k \cdot G = G + \cdots + G$ for the k-fold sum for $k > 2$, we have the sum rule

$$T_{k\cdot G,n} \equiv T_{G,n^k}\,.$$

The following lemma is similarly elementary, and we postpone the proof to Appendix A.

Lemma 15 (Projection under topological subgraphs)**.** *Let G and H be graphs and let $n \in \mathbb{N}$. If H is isomorphic to a topological subgraph of G, then $T_{H,n} \leq_{\mathrm{s}} T_{G,n}$.*

Moreover, we will use *contractions* of graphs: Given a set $U \subseteq V(G)$, the graph G/U is obtained by replacing U by a single vertex w incident with all edges *leaving* U; multiedges can be created in this process. Contractions can be executed cheaply if the set of involved edges is not too large:

Lemma 16 (Complexity of contractions)**.** *For every G and $n \in \mathbb{N}$ and $U \subseteq V(G)$, it holds that $\mathrm{C}(T_{G,n}) \leq \mathrm{C}(T_{G/U,n}) + |U| \cdot n^{a(U)}$, where $a(U) \coloneqq \sum_{u \in U} \deg_G(u) - |E(G[U])|$ is the number of edges incident with vertices in U.*

A proof is provided in Appendix A. We remark that such contractions for graph tensors have been extensively studied in the context of probabilistic inference in factor graphs and Bayesian networks, cf. [2, 21, 32, 35, 39, 42, 43, 44, 53], and even a contraction-based model of computation for multilinear forms was studied [1].

3.2 Asymptotic Rank

We now prove the main result of this section, Theorem 1. It will be convenient to work with exponents of graphs, in line with Christandl, Vrana and Zuiddam [13]. Define the *exponent* of a graph G as

$$\omega(G) \coloneqq \inf \{\beta \mid \mathrm{R}(T_{G,n}) = O(n^{\beta})\}.$$

A related quantity is the *exponent per edge* of G, defined as

$$\tau(G) \coloneqq \omega(G)/|E(G)|.$$

The following useful property can be shown in a standard manner. The proof is deferred to Appendix A.

Lemma 17 (Sum rule for exponents)**.** *Let G be a graph and let $k \in \mathbb{N}$. Then,*

$$\omega(k \cdot G) = k \cdot \omega(G). \tag{10}$$

To motivate the use of graph tensors in our next result, let us recall Strassen's upper bound on the maximal asymptotic rank of bilinear maps (i.e., 3-mode tensors) together with a high-level intuition of its proof. This property is often referred to as asymptotic submultiplicativity.

Theorem 18 (Strassen [46]; Upper bound on asymptotic rank for 3-tensors)**.** *Let $T \in (\mathbb{F}^n)^{\otimes 3}$ be an arbitrary 3-mode tensor. Then,*

$$\underset{\sim}{\mathrm{R}}(T) \leq n^{2\omega/3}.$$

Strassen's proof of Theorem 18 works with the universal tensor corresponding to the canonical bilinear map $U_3 : \mathbb{F}^n \times \mathbb{F}^n \to \mathbb{F}^n \otimes \mathbb{F}^n,\ (u, v) \mapsto u \otimes v$, which can be written down in coordinates as a 3-tensor in our notation as

$$U_3 = \sum_{i,j \in [n]} x_i y_j z_{ij}.$$

In the language of graph tensors, the tensor U_3 with the three modes x, y, z admits immediate representation as a graph tensor, namely we have $U_3 \equiv T_{P_2,n}$, where P_2 is the path graph with two

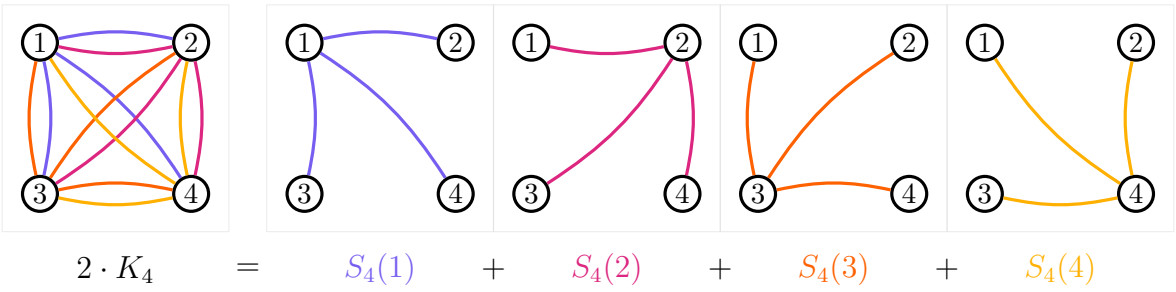


Figure 1: Decomposition of $2 \cdot K_d$ into d stars, (13), for $d = 4$.

edges. We can now generalize this observation on three modes to $d > 3$ modes as follows. The universal $(d-1)$-linear map $U_d : \mathbb{F}^n \times \cdots \times \mathbb{F}^n \to (\mathbb{F}^n)^{\otimes d-1}$ is represented as the d-mode tensor

$$U_d = \sum_{i_1,\ldots,i_d \in [n]} x^{(1)}_{i_1} \cdots x^{(d-1)}_{i_{d-1}} \cdot x^{(d)}_{i_1,\ldots,i_{d-1}} .$$

In the language of graph tensors, we observe that $U_d \equiv T_{S_d,n}$, where S_d is the star graph with $d-1$ edges, and $S_3 \equiv P_2$ in particular. This fact enables us to consider universality under Kronecker powers combinatorially from the perspective of graph sums and graph decompositions (cf. Lemma 12) and arrive at our main theorem. In fact, all our upper bounds are a consequence of this observation.

We now start the work towards our main theorem. Let $d \geq 3$ and write K_d for the complete graph on d vertices. We recall that $\tau(K_d) = \omega(K_d)/\binom{d}{2}$, and write $\tau(d) := \tau(K_d)$ for brevity in what follows.

Remark 19 (The exponents $\tau(d)$ for $d \geq 3$)**.** Recall that we write ω for the exponent of square matrix multiplication. Since $T_{K_3,n}$ is equivalent to the tensor of $n \times n$ matrix multiplication, we have $\tau(3) = \omega/3$. Moreover, it is known [13, Proposition 1.31] that $\tau(d) \leq \tau(e)$ whenever $d \geq e$.

We have the following generalization of Strassen's result in the language of graph tensors and graph exponents.

Lemma 20 (Generalization of Strassen's upper bound to d modes)**.** *Let $T \in (\mathbb{F}^n)^{\otimes d}$ be an arbitrary d-mode tensor with $d \geq 3$. Then,*

$$\underset{\sim}{\mathrm{R}}(T) \leq n^{(d-1)\tau(d)}.$$

Proof. For $u \in [d]$, let us write $S_d(u)$ for the d-vertex star graph with vertex set $[d]$ such that u is the unique *center* vertex of degree $d-1$. An arbitrary d-mode tensor $T \in (\mathbb{F}^n)^{\otimes d}$ admits representation in coordinates as the polynomial

$$T = \sum_{i_1,\ldots,i_d \in [n]} T_{i_1,\ldots,i_d} x^{(1)}_{i_1} \cdots x^{(d)}_{i_d} = \sum_{i_1,\ldots,i_{d-1} \in [n]} x^{(1)}_{i_1} \cdots x^{(d-1)}_{i_{d-1}} \left(\sum_{i_d \in [n]} T_{i_1,\ldots,i_d} x^{(d)}_{i_d} \right). \tag{11}$$

Explicitly, the linear substitution

$$x_{i_1,\ldots,i_{d-1}} \mapsto \sum_{i_d \in [n]} T_{i_1,\ldots,i_d} x^{(d)}_{i_d}$$

into the d-tensor $T_{S_d(d),n}$ shows by (11) and symmetry that

$$T \leq T_{S_d(u),n} \quad \text{for all } u \in [d]\,. \tag{12}$$

We now observe the graph-sum identity (see Figure 1)

$$S_d(1) + \cdots + S_d(d) \equiv 2 \cdot K_d\,, \tag{13}$$

where we recall that K_d is the complete graph on d vertices and $2 \cdot K_d$ is the graph obtained from K_d by taking two copies of each edge. From (12), Lemma 12, (13), and Remark 14 thus

$$T^{\otimes d} \leq T_{2\cdot K_d,n} \equiv T_{K_d,n^2}. \tag{14}$$

By definition of $\tau(d)$ and the properties of $\underset{\sim}{\mathrm{R}}$, it follows from (14) that

$$\underset{\sim}{\mathrm{R}}(T) \leq \underset{\sim}{\mathrm{R}}(T_{K_d,n})^{2/d} \leq n^{2\binom{d}{2}\tau(d)/d} = n^{(d-1)\tau(d)}. \qquad \square$$

For $d = 3$, Lemma 20 replicates Strassen's upper bound from Theorem 18, and generalizes it for $d \geq 4$. Our main result of this section is the following upper bound for $d \geq 4$ that combines Lemma 20 with an improved upper bound for $\tau(d)$.

Theorem 1 (Upper bound on asymptotic rank)**.** *For every d-mode tensor with $d \geq 4$, the asymptotic rank exponent of T is at most* $0.772318(d-1)$.

Proof. Christandl, Vrana, and Zuiddam [13] showed that $\underset{\sim}{\mathrm{R}}(T_{K_d}) \leq n^{\binom{d}{2}\log_7(9/2)}$ for $d \geq 4$, which implies $\tau(d) < 0.7729$. By a careful analysis and modification of their argument, we obtain the sharper bound $\tau(d) < 0.772318$; see Appendix C. The theorem then follows by Lemma 20. $\square$

Remark 21 (Bounds on $\tau(d)$ for $d \geq 4$)**.** To obtain the constant 0.772318 in Theorem 1, an upper bound for $\tau(4)$ suffices since $\tau(d) \leq \tau(4)$ for $d \geq 4$. It is natural to ask whether better upper bounds can be obtained by finding better upper bounds on $\tau(d)$ for $d \geq 4$. The best-known lower bounds on $\tau(d)$ allow for a scenario in which $\tau(d) \leq 1/2 + o(1)$, implying $\underset{\sim}{\mathrm{R}}(T) \leq n^{(d-1)/2+o(d)}$ would follow for d-mode tensors. On the other hand, the current best-known upper bound on ω yields an exponent greater than the exponent $(d-1)\cdot 0.772318$ in Theorem 1, while $\omega = 2$ would imply $\tau(3) = \tau(4) = 2/3$. It is an open problem whether there is a d such that $\tau(d) < 2/3$. Recalling from Example 11 that the k-matching tensor $T_{H_k,n}$ has $d = 2k$ modes and its rank $\mathrm{R}(T_{h_k,n}) = n^k$ equals a matrix-flattening-rank lower bound, this lower bound applies to also asymptotic rank and thus $\underset{\sim}{\mathrm{R}}(T_{H_k,n}) = n^k$, implying $\tau(d) \geq 1/2$ by Theorem 1 for all $d \geq 4$.

4 Asymptotic Circuit Complexity

This section proceeds to study tensors with $d \geq 4$ modes from the standpoint of the asymptotic circuit complexity of a tensor viewed as a set-multilinear polynomial. Recalling Example 11 and our discussion in the introduction, for $d \geq 4$ the asymptotic circuit complexity and the asymptotic tensor rank are no longer tightly connected as is the case for $d = 3$. Yet families tensors of interest from the standpoint of algorithms and complexity (e.g. matrix permanent, hypercliques, general convolutions, iterated matrix multiplication, ...) can be captured as Kronecker powers of individual base tensors with $d \geq 4$ modes, making the asymptotic circuit complexity of these constant-size base tensors worth studying. Graph tensors provide a convenient tool for this study.

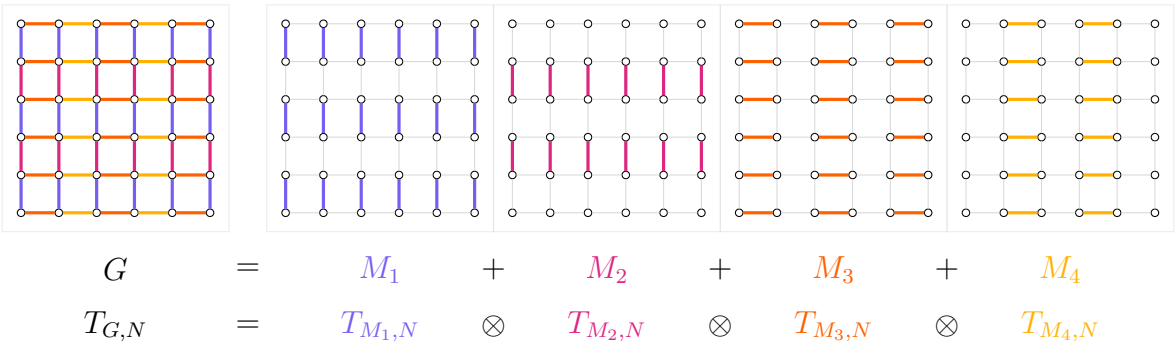


Figure 2: The grid G is the sum of four matchings, so Lemma 12 shows that its graph tensor $T_{G,N}$ is the Kronecker product of four graph tensors of matchings. But since grids are hard by Lemma 25 and matchings are easy by Example 11, we do not expect complexity to be submultiplicative.

4.1 Complexity of Kronecker Powers

We ask whether tensors $U, T \in (\mathbb{F}^n)^{\otimes d}$ of low circuit complexity also have Kronecker products $U \otimes T$ of low circuit complexity. Recalling that tensor rank is submultiplicative for tensors of all orders, a natural first hope promoted by the $d = 3$ connection of circuit complexity to tensor rank would be to establish the *submultiplicativity* property $\mathrm{C}(T \otimes U) \leq \mathrm{C}(T) \cdot \mathrm{C}(U)$ of the complexity measure C for all orders $d \in \mathbb{N}$. However, already at $d = 4$ submultiplicativity would have a breakthrough consequence.

Theorem 22 (Fast matrix multiplication under submultiplicativity)**.** *Submultiplicativity of* C *on* 4*-mode tensors implies* $\omega = 2$.

Proof. The cycle graph C_4 is a sum of two edge-disjoint matchings. By virtue of Lemma 12 and Example 11, submultiplicativity implies $\mathrm{C}(T_{C_4,n}) \leq O(n^2)$. Moreover, $T_{C_3,n}$ is a projection of $T_{C_4,n} \equiv \sum_{i,j,k,\ell} x_{ij} y_{jk} z_{k\ell} w_{\ell i}$ via $z_{k\ell} \mapsto \delta_{k\ell}$, see Lemma 15. Hence, $\mathrm{C}(T_{C_3,n}) \leq O(n^2)$ and $\omega = \omega(C_3) = 2$. □

The remainder of this subsection strengthens the case for the absence of submultiplicativity and illustrates the serendipity of graph tensors as a tool when working with tensors with more modes.

We note that that submultiplicativity for d modes implies the same for $d' < d$, as discussed in Fact 47 in Appendix B.

Reductions from Permanents

We proceed to reduce matrix permanent tensors to graph tensors of square grid graphs G. Since grids have large treewidth, results for related problems [22, 18] allow us to expect their graph tensors to have large algebraic complexity. Indeed, we show in Lemma 25 that the $d \times d$ matrix permanent can be projected from the graph tensors of $(d+2) \times (d+2)$ grids, even for mode dimension 2. On the other hand, as shown in Figure 2, every grid G can be written as the sum

$$G \equiv M_1 + M_2 + M_3 + M_4 \tag{15}$$

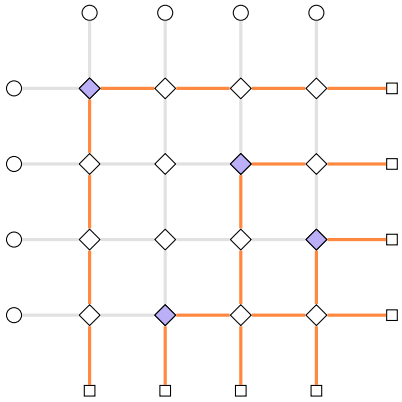


Figure 3: The reduction from the permanent to the graph tensors of grids, as shown in Lemma 25. The **gray** and **orange** edges are assigned 0 and 1, respectively. Each horizontal and each vertical path flips from 0 to 1 exactly once; whenever this happens at some vertex $v_{i,j}$ (a *flip vertex*, shown **purple**), we ensure that both the i-th horizontal and the j-th vertical path flip at $v_{i,j}$. In other words, flip vertices always lie on orange corners. Each flip vertex $v_{i,j}$ contributes $x_{i,j}$ to the total weight of the assignment, otherwise 1. The weight of the assignment shown here is $x_{11}x_{23}x_{34}x_{42}$, which is indeed of the form $\prod_i x_{i,\pi(i)}$ for $\pi = 1342$.

of four edge-disjoint matching graphs M_1, M_2, M_3, M_4. Consequently, for any $N \in \mathbb{N}$ by Lemma 12 we have

$$T_{G,N} = T_{M_1,N} \otimes T_{M_2,N} \otimes T_{M_3,N} \otimes T_{M_4,N}. \tag{16}$$

The graph tensors of matchings have low complexity by Example 11. Thus, if C were submultiplicative, then $T_{G,N}$ would also have low complexity. More generally, this holds for every graph G of small edge-chromatic number, i.e., whenever G can be partitioned into few matchings:

Lemma 23 (Graphs with small edge-chromatic number under submultiplicativity)**.** *Let G be a graph with maximum edge-multiplicity $b \in \mathbb{N}$ such that $E(G)$ can be partitioned into t matchings. Let $N \in \mathbb{N}$. If* C *is submultiplicative on $|V(G)|$-mode tensors, then* $\mathrm{C}(T_{G,N}) \leq 2^t |V(G)|^t \cdot N^{tb}$.

Proof. Let $E(G) = M_1 \cup \ldots \cup M_t$. Example 11 and Remark 14 give $\mathrm{C}(M_i) \leq 2|V(G)| \cdot N^b$. The assumed submultiplicativity of C, together with Lemma 12 gives $\mathrm{C}(T_{G,N}) \leq \prod_{i=1}^{t} \mathrm{C}(T_{M_i,N}) \leq 2^t |V(G)|^t \cdot N^{tb}$, thus proving the claim. □

Using (16), we obtain the following corollary for grids.

Corollary 24. *Let G denote the $n \times m$ grid, for $n, m \in \mathbb{N}$, with maximum edge-multiplicity $b \in \mathbb{N}$. If* C *is submultiplicative on nm-mode tensors, then* $\mathrm{C}(T_{G,N}) \leq 16n^4m^4 \cdot N^{4b}$.

Our results are obtained by invoking this corollary for various choices of grids G and dimensions N. We begin under the comparatively weak assumption $\mathsf{VP} \neq \mathsf{VNP}$ [48]. Towards this end, we first show how the permanent reduces to grid graph tensors.

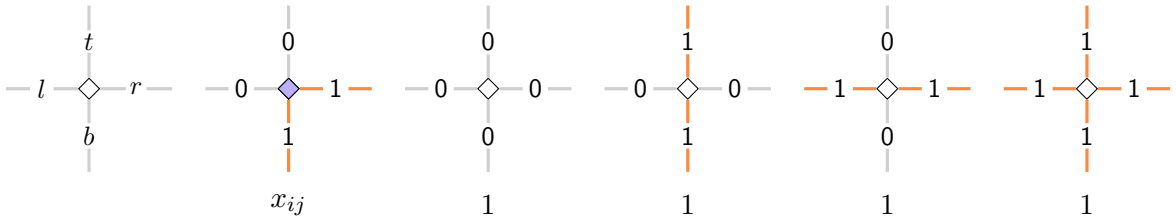


Figure 4: A non-border signature. Below an assignment is the value of the signature. Assignments that are not listed here evaluate to zero.

Lemma 25 (The permanent tensor reduces to a grid-graph tensor)**.** *The $n \times n$ permanent tensor* per_n *is a projection of $T_{G,2}$ for the simple $(n+2) \times (n+2)$ grid graph G.*

On a high level, the reduction in Lemma 25 proceeds as illustrated in Figure 3: A graph G is obtained by attaching pendant vertices of degree 1 to the boundary of an $n \times n$ grid; the pendant vertices are needed to ensure boundary conditions. Each edge of the resulting graph carries a Boolean state, and we substitute values into the indeterminates of $T_{G,2}$ such that the resulting polynomial counts assignments a to the edges of the grid with the following properties:

1. We ensure that each horizontal and each vertical path P starts with a 0-edge and the edge states flip from 0 to 1 at exactly one vertex; we call this the *flip vertex* of P under a.
2. For all $i, j \in [n]$, we ensure that a vertex $v_{i,j}$ is the flip vertex of the i-th horizontal path if and only if the same vertex $v_{i,j}$ also is the flip vertex of the j-th vertical path.
3. Finally, we ensure that a flip vertex contributes a factor $x_{i,j}$ to the weight of the assignment, while non-flip vertices do not contribute (i.e., they contribute 1).

Under these conditions, the flip vertices under an assignment a induce an $n \times n$ permutation matrix P_π, and the weight of a is the product of the indeterminates $x_{i,\pi(i)}$ selected by P_π. Summing over all valid assignments a gives precisely the permanent.

Proof of Lemma 25. The proof implements the idea sketched above, but viewing the pendant vertices as the degree-3 border vertices of the $(n+2) \times (n+2)$ grid. We label the vertices of G as $v_{i,j}$ for $0 \le i, j \le n+1$. For $v_{i,j}$ with $1 \le i, j \le n$, define the signature $f_{i,j} : \{0,1\}^4 \to \{0,1\}$ as follows, where the inputs are read clockwise around $v_{i,j}$, starting from the top edge, as t, r, b, ℓ:

$$f_{i,j}(t,r,b,\ell) = \begin{cases} 1 & t = b \text{ and } \ell = r & \text{(no flip)}\,, \\ x_{ij} & t = \ell = 0 \text{ and } r = b = 1 & \text{(flip)}\,, \\ 0 & \text{otherwise} & \text{(invalid)}\,. \end{cases}$$

Figure 4 depicts all the valid assignments of any non-border signature.

The top border vertices $v_{0,j}$ are given signatures to ensure that the top-most edge of vertical paths is assigned 0:

$$f_{0,j}(r,b,\ell) = \begin{cases} 1 & \text{if } r = b = \ell = 0\,, \\ 0 & \text{otherwise}\,. \end{cases}$$

The bottom border vertices $v_{n+1,j}$ are given signatures to ensure that the bottom-most edge of vertical paths is assigned 1:

$$f_{n+1,j}(t,r,\ell) = \begin{cases} 1 & \text{if } t = r = \ell = 1\,, \\ 0 & \text{otherwise}\,. \end{cases}$$

Analogous signatures are defined for the left and right borders. The four corners of G receive constant all-ones signatures.

Our choice of signatures ensures that, in an assignment $a \in \{0,1\}^{E(G)}$ with $w(a) \coloneqq \prod_{i,j} f_{i,j}(a) \neq 0$, the indices (i,j) with $f_{i,j}(a)$ in state "flip" form a permutation matrix corresponding to some permutation π, and that $w(a) = \prod_i x_{i,\pi(i)}$. It follows that $\operatorname{per}(A) = \sum_\pi \prod_i x_{i,\pi(i)} \leq_{\mathrm{s}} T_{G,2}$. □

Theorem 2 (Submultiplicativity of C implies $\mathsf{VP} = \mathsf{VNP}$)**.** *There are explicit tensors $T_1, T_2, \ldots$ and $U_1, U_2, \ldots$ with $T_d, U_d \in (\mathbb{C}^2)^{\otimes d}$ for all $d \in \mathbb{N}$ such that the following holds: If $\mathrm{C}(T_d \otimes U_d) \leq \operatorname{poly}(d, \mathrm{C}(T_d), \mathrm{C}(U_d))$ for all $d \in \mathbb{N}$, then $\mathsf{VP} = \mathsf{VNP}$.*

Proof. Under the assumption of the theorem, Corollary 24 and Lemma 25 would imply $\mathrm{C}(T_{G,2}) \leq \operatorname{poly}(n)$, contradicting the VNP-hardness of the permanent. □

If we assume that permanents require exponential-size circuits, then we can obtain stronger bounds. Towards this, we first contract large grids into smaller grids with larger edge-multiplicities. (Larger edge-multiplicities essentially correspond to larger mode dimensions: For every graph H and $b, k, n \in \mathbb{N}$ with k dividing n, and writing $B = b^{n/k}$, we have $T_{H,B} \equiv T_{n/k \cdot H, b}$.) Namely, first partition the $n \times n$ grid into $k \times k$ pieces of size $n/k \times n/k$ each, then contract each piece to a single vertex by Lemma 16. The complexity bound in Lemma 16 requires a little care in the choice of contraction sequence to obtain the desired lower bounds later.

Lemma 26 (Grid graph tensor contraction by subgrids)**.** *Let k divide n and let H and G be the $k \times k$ and $n \times n$ grid graphs, respectively. Let $b \in \mathbb{N}$ and $B = b^{n/k}$. Then, $\mathrm{C}(T_{G,b}) \leq (2b^4 + b)n^2 \cdot b^{3(n/k)} + \mathrm{C}(T_{H,B})$.*

Proof. Partition the vertex set of G into $k \times k$ square blocks $A_{i,j}$, each of size $n/k \times n/k$. For each $i, j \in [k]$, we then perform the following:

1. For $t = 1, \ldots, n/k$, contract the vertices in row t of block $A_{i,j}$ by Lemma 16. Each contraction involves $3n/k + 1$ edges and thus incurs an additive cost term of $bn/k \cdot b^{3(n/k)}$ in Lemma 16. Block $A_{i,j}$ now consists of a vertical path of length n/k, with edge-multiplicity n/k between consecutive vertices, and with two additional edges per vertex. The total cost of the contractions in this step is $b(n/k)^2 \cdot b^{3(n/k)}$.
2. For $t = 1, \ldots, n/k - 1$, we contract vertices t and $t+1$ on the path remaining in block $A_{i,j}$. This involves $3n/k + 4$ edges per contraction and incurs a cost term of $2b^4 \cdot b^{3(n/k)}$ in Lemma 16. Block $A_{i,j}$ now consists of a single vertex incident with all edges leaving $A_{i,j}$. The total cost of the contractions in this step is $2b^4(n/k) \cdot b^{3(n/k)}$.

In total over the k^2 cells, we incur cost at most $(2b^4 + b)k^2(n/k)^2 \cdot b^{3(n/k)} = (2b^4 + b)n^2 \cdot b^{3(n/k)}$ for all contractions and obtain the $k \times k$ grid with edge-multiplicity n/k from G, thus proving the lemma. □

A lower bound now follows easily.

Theorem 3 (Submultiplicativity of C implies faster permanents)**.** *For all $0 < c < 1$, if circuit complexity is submultiplicative on $\lceil 4/c \rceil^2$-mode tensors, then permanents have circuits of size $O(2^{cn})$.*

Proof. Choose $k = \lceil 4/c \rceil$ so $4/k \leq c$, and let H be the $k \times k$ grid on k^2 vertices. By Remark 14 and Lemma 26 with $b = 2$, if the k^2-mode tensor $T_{n/k \cdot H, 2}$ admits a circuit of size s, then the n^2-mode tensor $T_{G,2}$ for the $n \times n$ grid G admits a circuit of size $s' \leq n^{O(1)} \cdot 8^{n/k} + s$, and Lemma 25 gives a circuit of size $s' \cdot n^{O(1)}$ for per_n.

If submultiplicativity indeed held, then Corollary 24 would give a circuit of size $s \leq n^{O(1)} \cdot 2^{4n/k}$ for $T_{n/k \cdot H,2}$, so we obtain a circuit of size $s' \leq n^{O(1)} \cdot 2^{4n/k} \leq O(2^{cn})$ for $T_{G,2}$ and the permanent. □

The above strategy for conditional arguments could be refined in different ways to rule out submultiplicativity on fewer modes, e.g., by choosing graph tensors other than grids.

Reductions from Hyperclique Tensors

For $3 \leq h < k$, the (h,k)-*hyperclique conjecture* states that for any $\varepsilon > 0$, there is no $O(n^{k-\varepsilon})$-time algorithm that decides whether an h-uniform hypergraph contains a k-hyperclique. More explicitly, for $N \in \mathbb{N}$, define the following tensor on $\binom{k}{h}$ modes $X^{(S)}$ for $S \in \binom{[k]}{h}$ via

$$H^N_{h,k} = \sum_{f\colon [k] \to [N]} \prod_{S \in \binom{[k]}{h}} x^{(S)}_{f|_S}.$$

Under adequate uniformity assumptions, a circuit family $(C_N)_N$ for the sequence of tensors $(H^N_{h,k})_N$ of size $N^{k-\varepsilon}$ would give an algorithm for the problem that falsifies this conjecture. Even absent such assumptions, we can formulate the hypothesis that there is no such circuit family whatsoever, which we will refer to as the *non-uniform algebraic* (h,k)*-hyperclique conjecture.*

Theorem 4 (Submultiplicativity of C implies faster hypercliques)**.** *For every even $d \geq 8$, the non-uniform $(\frac{d}{2}-1, \frac{d}{2})$-hyperclique conjecture rules out submultiplicativity of circuit complexity on d-mode tensors.*

Proof. We illustrate the case $d = 8$. The corresponding $(3,4)$-hyperclique conjecture rules out algorithms running in time $N^{4-\varepsilon}$ for detecting 4-vertex hypercliques in 3-uniform hypergraphs with N vertices. In this case, we can write

$$H^N_{3,4} \equiv \sum_{i,j,k,\ell=1}^{N} x_{ijk} y_{ij\ell} z_{ik\ell} w_{jk\ell}.$$

Consider the bipartite incidence (non-hyper-)graph $I = I_{3,4}$ associated with the $(3,4)$-hyperclique, which is $K_{4,4}$ minus a perfect matching, as depicted in Fig. 5. The associated graph tensor $T_{I,N}$ can be written down using twelve indices, one for each edge in I, and has 8 modes. For the sake of legibility, we collect the indices into four groups $i,j,k,\ell \in [n]^3$, with three entries each, corresponding to the neighborhoods of the hyperedges in the bottom part of the graph. We recall that the superscript indices indicate the vertex, and the subscript the (images of) incident edges.

$$T_{I,N} \equiv \sum_{i,j,k,\ell \in [n]^3} x^{(1)}_{i_1,i_2,i_3} \cdot x^{(2)}_{j_1,j_2,j_3} \cdot x^{(3)}_{k_1,k_2,k_3} \cdot x^{(4)}_{\ell_1,\ell_2,\ell_3} \cdot x^{(123)}_{i_1,j_1,k_1} \cdot x^{(124)}_{i_2,j_2,\ell_1} \cdot x^{(134)}_{i_3,k_2,\ell_2} \cdot x^{(234)}_{j_3,k_3,\ell_3}.$$

Observe now that $H^N_{3,4}$ is a linear projection of $T_{I,N}$ under the substitution

$$x^{(v)}_{\alpha\beta\gamma} \mapsto \begin{cases} 1, & \text{if } \alpha = \beta = \gamma \\ 0, & \text{otherwise} \end{cases} \text{ for } v \in \{1,2,3,4\}, \alpha,\beta,\gamma \in [n], \text{ and}$$
$$x^{(123)}_{ijk} \mapsto x_{ijk},\ x^{(124)}_{ijk} \mapsto y_{ijk},\ x^{(134)}_{ik\ell} \mapsto z_{ik\ell},\ x^{(234)}_{jk\ell} \mapsto w_{jk\ell} \text{ for } i,j,k,\ell \in [n].$$

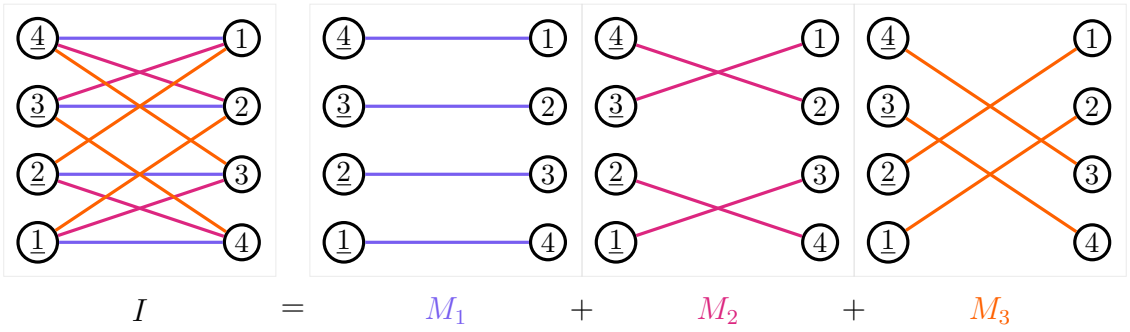


Figure 5: Decomposition of the incidence graph I of the $(3,4)$-hyperclique into matchings M_1, M_2, M_3, where $\underline{1}, \underline{2}, \underline{3}, \underline{4}$ indicate the hyperedges $234, 134, 124, 123$, respectively.

Therefore, $\mathrm{C}(H^N_{3,4}) \leq \mathrm{C}(T_{I,N})$.

On the other hand, we can decompose I as a union of three perfect matchings with four edges each, see again Fig. 5. Via Lemma 23, submultiplicativity of C would imply a circuit of size $O(N^3)$ for $T_{I,N}$, and hence for $H^N_{3,4}$, by Lemma 12. □

Remark 27. In the above proof, it seems cumbersome to transition to the incidence graph, which in particular doubles the degree of the resulting tensors. A more direct approach could involve the hypergraph-tensor of the $(3,4)$-hyperclique itself, which only has four modes. In this case, however, the only possible decompositions into sub-hypergraphs are either into (a) a single hyperedge and a hypergraph with three hyperedges, or into (b) two pairs of hyperedges. Both decompositions would yield circuits of size at least $\Omega(n^4)$, even assuming submultiplicativity, simply because of the numbers of variables involved. Intuitively, passing to the incidence graph allows us to trade higher overall degree (number of modes) for a smaller total number of variables (dimension per mode), which is beneficial when "Kroneckering up".

4.2 Upper Bounds via Line Graph Treewidth

The *line-treewidth* $\mathrm{ltw}(G)$ is the treewidth $\mathrm{tw}(L(G))$ of the *line graph* $L(G)$ of G, where the vertex set of $L(G)$ is $E(G)$ and two vertices corresponding to edges e, e' are adjacent in $L(G)$ if they share a vertex in G. We collect a few facts about the treewidth of a line graph. We recall that $\Delta(H)$ is the maximal degree of any vertex in H.

Fact 28 (Treewidth versus line-treewidth [27, Equations (2) and Proposition 2.3])**.** *For every graph H, we have*

$$\mathrm{tw}(H) - 1 \leq \mathrm{ltw}(H) \leq (\mathrm{tw}(H) + 1)\Delta(H) - 1.$$

Of special interest is the treewidth of the line graph of the complete graph, which has been determined exactly.

Theorem 29 (Line-treewidth of a complete graph [26, Theorem 1])**.** *For $d \geq 1$, we have*

$$\mathrm{ltw}(K_d) = \begin{cases} \left(\frac{d-1}{2}\right)^2 + d - 2, & \text{if } d \text{ is odd,} \\ \left(\frac{d-2}{2}\right)\left(\frac{d}{2}\right) + d - 2, & \text{if } d \text{ is even.} \end{cases}$$

Similar to standard dynamic programming arguments on tree-decompositions in, e.g., [52, Section 5.1], [20, Lemma 5.23], and [23, Example 3.4], we can show:

Lemma 30. *For every H and $n \in \mathbb{N}$, we have $\mathrm{C}(T_{H,n}) \leq |V(H)| \cdot n^{\mathrm{ltw}(H)+1}$.*

Proof. Write $k \coloneqq \mathrm{ltw}(H)$. Let $\beta \colon V(T) \to 2^{E(H)}$ be a tree decomposition for the line graph $L(H)$ of width $\leq k$. We contract H in a bottom-up fashion along the tree decomposition.

Since the set of edges $I(v)$ incident to a vertex $v \in V(H)$ forms a clique in $L(H)$, there exists a node $t_v \in V(T)$ such that $I(v)$ is fully contained in $\beta(t_v)$; there may be multiple such nodes t_v, from which we choose one arbitrarily. Define a map $\chi \colon V(T) \to 2^{V(H)}$ by $\chi(t) \coloneqq \{v \in V(H) \mid t_v = t\}$. By definition, every $t \in V(T)$ satisfies $\bigcup_{v \in \chi(t)} I(v) \subseteq \beta(t)$. In particular, the number of edges incident to some vertex in $\chi(t)$ is

$$\left| \bigcup_{v \in \chi(t)} I(v) \right| \leq k + 1. \tag{17}$$

Let $\ell \in V(T)$ be a leaf of T and write $t \in V(T)$ for its parent. Note that every tree with at least two vertices contains a leaf. We contract the vertices in $\chi(\ell)$ to a fresh vertex v. By (17), the cost of this contraction is bounded by n^{k+1}. The edges incident with the new vertex v are in $\beta(\ell) \cap \beta(t)$. We add the fresh vertex v to $\chi(t)$. This maintains the property (17) of $\gamma(t)$ having at most $k+1$ incident edges. We may now delete ℓ from T and proceed with the next leaf. When the process terminates, a single vertex is left. □

As an immediate consequence the above lemma proves

$$\underset{\sim}{\mathrm{C}}(T_{H,n}) \leq n^{\mathrm{ltw}(H)+1} \tag{18}$$

for every graph H. For general tensors, recall that we write $\eta(T)$ for the exponent of the asymptotic circuit complexity $\underset{\sim}{\mathrm{C}}(T) = n^{\eta(T)}$ of a tensor $T \in (\mathbb{F}^n)^{\otimes d}$. We then have the following more precise version of Theorem 6:

Theorem 31 (Upper bound on the asymptotic circuit complexity of a d-tensor)**.** *Let $T \in (\mathbb{F}^n)^{\otimes d}$ be a d-mode tensor of dimension n. Then, we have*

$$\eta(T) \leq \frac{2}{d}(\mathrm{ltw}(K_d) + 1) = d/2 + 1 - \frac{7 + (-1)^d}{4d}.$$

Proof. Similar to the proof of Lemma 20 and using Lemma 12 as well as Remark 14, we have that

$$T^{\otimes dm} \equiv (T^{\otimes d})^{\otimes m} \leq T_{2 \cdot K_d, n^m} \equiv T_{K_d, n^{2m}} \text{ for all } m.$$

By Fact 46 and more specifically Lemma 43, we have that

$$\mathrm{C}(T^{\otimes k}) \leq \mathrm{C}(T^{\otimes d \cdot \lceil k/d \rceil}) \text{ for all } k,$$

whereas Lemma 30 shows

$$\mathrm{C}(T^{\otimes dm}) \leq d^{O(1)} \cdot n^{2m \cdot (\mathrm{ltw}(K_d)+1)} \text{ for all } m.$$

In combination, for $m = \lceil k/d \rceil$, we find that

$$\mathrm{C}(T^{\otimes k}) \leq d^{O(1)} \cdot n^{2\lceil k/d \rceil (\mathrm{ltw}(K_d)+1)} \leq d^{O(1)} \cdot n^{2(k/d+1)(\mathrm{ltw}(K_d)+1)}.$$

By definition of $\eta(T)$, this shows that $\eta(T) \leq 2/d \cdot (\mathrm{ltw}(K_d) + 1)$. The final bound then follows from Theorem 29. □

Remark 32. It is natural to ask whether the exponent from Theorem 31 is asymptotically tight under standard assumptions from complexity theory. To the best of our knowledge, current techniques would only rule out exponents of the form $o(d/\log d)$.

For instance, known results on the complexity of subgraph isomorphism [37, 29, 17] imply $\eta(T_{H,n}) \in \Omega(d/\log d)$ for a sequence of d-vertex graphs H of maximum degree 3. It is also possible to adapt a recent bound by Pratt [41] to a higher number of modes, which would yield a lower bound under the set cover conjecture. Since all of the resulting lower bounds are fairly loose, we refrain from formally stating and proving them here.

5 Exploratory Upper Bounds in Intermediate Orders

In this final section, we show upper bounds on the asymptotic circuit complexity of generic d-mode tensors for $d = 4$ and $d = 5$. So far, we have presented various techniques to obtain upper bounds on both the asymptotic circuit complexity and rank as d grows. We now combine some of these techniques with computational search to find better upper bounds for 4-mode and 5-mode tensors.

5.1 Restricted Submultiplicativity

The results in Section 4.1 strongly suggest the absence of circuit complexity submultiplicativity for d-mode tensors with $d \geq 4$: For d-mode tensors T and U of low complexity, no useful upper bounds for the circuit complexity $\mathrm{C}(T \otimes U)$ as a function of $\mathrm{C}(T)$ and $\mathrm{C}(U)$ are known. However, we can construct small circuits for $T \otimes U$ from small circuits for T and U, provided that one of the circuits is sufficiently simple.

To this end, we will need the following property of tensor contraction. Let $\ell_1, \dots, \ell_d \colon \mathbb{F}^m \to \mathbb{F}$ and $h_1, \dots, h_d \colon \mathbb{F}^n \to \mathbb{F}$ be linear forms. We consider the Kronecker product of the corresponding rank-one tensors $L = \prod_{i=1}^d \ell_i(x^{(i)})$ and $H = \prod_{i=1}^d h_i(y^{(i)})$, where the i-th mode of $L \otimes H$ has coordinates $z^{(i)}_{jk} \coloneqq x^{(i)}_j \otimes y^{(i)}_k$, with $j \in [m], k \in [n]$. For brevity, we write $z^{(i)}$ for the $m \times n$ matrix with entries $z^{(i)}_{jk}$, and identify h_i with its vector of coefficients in $\mathbb{F}^n$. Then, one can check that

$$L \otimes H = \prod_{i=1}^{d} \ell_i(h_i(z^{(i)}_{11}, \dots, z^{(i)}_{1n}), \dots, h_i(z^{(i)}_{m1}, \dots, z^{(i)}_{mn}))) = L(z^{(1)} \cdot h_1, \dots, z^{(d)} \cdot h_d).$$

By linearity and the fact that every tensor U has a decomposition into rank-one tensors, we obtain that

$$U \otimes H = U(z^{(1)} \cdot h_1, \dots, z^{(d)} \cdot h_d) \tag{19}$$

holds for all tensors U with d modes in coordinates $z^{(i)}_{jk}$. This can be applied in order to obtain the following result.

Lemma 33. *Let $T \in (\mathbb{F}^n)^{\otimes d}$ and $U \in (\mathbb{F}^m)^{\otimes d}$ with T of rank at most r, with $T = \sum_{i=1}^r \prod_{j=1}^d \ell_{ij}(x^{(j)})$. For $1 \leq j \leq d$, let $\Lambda^{(j)}$ be the $n \times r$ matrix of linear forms $(\ell_{ij})_i$ appearing in the j-th mode of the rank-decomposition of T, and we recall that $z^{(i)}$ is an $m \times n$ matrix of indeterminates $z^{(i)}_{j,k}$ for $1 \leq i \leq d$. Then,*

$$\mathrm{C}(U \otimes T) \leq r(\mathrm{C}(U) + 1) + \sum_{i=1}^{d} \mathrm{C}(z^{(i)} \cdot \Lambda^{(i)}) + r \cdot m \cdot d.$$

Proof. Distributing (19) linearly across the terms of the assumed rank-r decomposition of T, we see that

$$U \otimes T = \sum_{i=1}^{r} \left(U \otimes \prod_{j=1}^{d} \ell_{ij}(x^{(j)}) \right) = \sum_{i=1}^{r} U(z^{(1)} \cdot \ell_{i,1}, \ldots, z^{(d)} \cdot \ell_{i,d}). \tag{20}$$

On the right-hand side of (20), observe that the j-th mode of the instance of U appearing in the i-th term in the sum obtains as an input precisely the i-th row of $z^{(j)} \cdot \Lambda^{(j)}$, which is a column vector of dimension m. To implement (20) as a circuit, we place a fan-in r $+$-gate at the top, contributing r wires. The i-th input of this topmost $+$-gate is given by a copy of the circuit for U, contributing $r \cdot \mathrm{C}(U)$ to the circuit size overall. In addition, we need to compute the inputs to the individual r copies of U (which might sound negligible, but becomes crucial in the asymptotic regime.) As mentioned, for this it suffices to compute the d matrix products $z^{(1)} \cdot \Lambda^{(1)}, \ldots, z^{(d)} \cdot \Lambda^{(d)}$. Finally, we need to wire each of the computed entries of $z^{(j)} \cdot \Lambda^{(j)}$ to the respective input of U, amounting to a total of $r \cdot m \cdot d$ wires. This yields the claimed bound. □

To become useful in our applications, we need an asymptotic variant of Lemma 33, that is, for higher Kronecker powers of T and U. As mentioned in the proof, here the complexity bound on the computation of the matrix-vector products that comprise the inputs of U becomes relevant.

Theorem 5 (Mixed asymptotic rank and circuit complexity)**.** *Given d-mode tensors T and U, where T has asymptotic rank exponent $r \geq 1$ and U has asymptotic circuit complexity exponent $s \geq 1$, the asymptotic circuit complexity exponent of $T \otimes U$ is at most $r + s$.*

Proof. We need to show that $\underset{\sim}{\mathrm{C}}(U \otimes T) \leq \underset{\sim}{\mathrm{R}}(T) \cdot \underset{\sim}{\mathrm{C}}(U)$ holds. Let $T \in (\mathbb{F}^n)^{\otimes d}$ and $U \in (\mathbb{F}^m)^{\otimes d}$. For $T \in (\mathbb{F}^n)^{\otimes d}$, a rank-$r$ decomposition of $T = \sum_{i=1}^{r} \prod_{j=1}^{d} \ell_{ij}(x^{(j)})$ leads to a rank-r^k decomposition

$$T^{\otimes k} = \sum_{i_1,\ldots,i_k=1}^{r} \prod_{j=1}^{d} \left(\bigotimes_{\kappa=1}^{k} \ell_{\kappa j}(x^{(j)}) \right)$$

of $T^{\otimes k} \in (\mathbb{F}^{n^k})^{\otimes d}$, where the rank-1 term corresponding to indices $i_1, \ldots, i_k$ has as its j-th mode the Kronecker product $\ell_{i_1 j} \otimes \cdots \otimes \ell_{i_k, j} \in \mathbb{F}^{n^k}$. Hence, the matrices constituting the j-th mode of the rank decomposition become $n^k \times r^k$-matrices given by $(\Lambda^{(j)})^{\otimes k}$. Now, when applying Lemma 33, the naive circuit for the product $z^{(j)} \cdot (\Lambda^{(j)})^{\otimes k}$ would simply spell out each of the $m^k \cdot r^k$ inner products of dimension-n^k vectors, and hence be of size at least $(rnm)^k$. (Here, $z^{(i)}$ now has become a $m^k \times n^k$ matrix.) However, we can now use the bound (7) provided by Yates's algorithm from Theorem 8 in its most basic form for $d = 2$ to compute a single $n^k \times n^k \times r^k$ matrix-vector product of $z_j^{(i)} \cdot \Lambda^{(i)}$ using a circuit of size $O(r^{k+1}k)$ (as opposed to the naive $\Omega((nr)^k)$), assuming $r \geq n$. This gives the improved bound of $O((mr)^k rk)$ for $\mathrm{C}(z^{(i)} \cdot \Lambda^{(i)})$.

Therefore, using $(U \otimes T)^{\otimes k} \equiv U^{\otimes k} \otimes T^{\otimes k}$ and $\mathrm{C}(U^{\otimes k}) > d \cdot m^k$ simply by counting the number of variables appearing in U, we find from Lemma 33 that

$$\mathrm{C}((U \otimes T)^{\otimes k}) \leq r^k \cdot (\mathrm{C}(U^{\otimes k}) + 1) + O(kd(mr)^{k+1}) \leq O(kdr \cdot r^k \, \mathrm{C}(U^{\otimes k})),$$

which implies the claim, seeing that $(kdr)^{1/k} \to 1$ as $k \to \infty$. □

5.2 Fractional Graph Tensors

As it turns out, our arguments can be stated cleanly using *fractional graphs*, i.e., graphs with edge weights from $\mathbb{Q}$ (as opposed to multigraphs, which can be interpreted as graphs with edge-weights from $\mathbb{N}$.)

In the following, we will identify multigraphs with integer weights with ordinary multigraphs where every weighted edge e has been duplicated $w(e)$ times.

More specifically, let $G = (V, E, w)$ be a weighted multigraph with $w\colon E \to \mathbb{Q}_{>0}$ a weight function taking on only positive values. We call such graphs *fractional* in the following. Define $d_G \in \mathbb{N}$ to be the least common denominator of w, that is, the smallest positive integer such that $d_w \cdot w(e) \in \mathbb{N}$ for all $e \in E$. Then, we define

$$T_{G,n} \coloneqq T_{d_G \cdot G, n}.$$

Consequently, for a *fractional* graph G, we set

$$\omega(G) \coloneqq \omega(d_G \cdot G)/d_G,\ \tau(G) \coloneqq \tau(d_G \cdot G).$$

Remark 34. Note that by Lemma 17, the choice of d_G is immaterial as long as the resulting graph is a proper multigraph, that is, has integer edge weights, in that

$$\omega(G) = \omega(d \cdot G)/d$$

for all $d \in \mathbb{N}$ such that $d \cdot G$ is a multigraph (which implies that d is a multiple of d_G).

Clearly, we can extend the notion of multigraph sums and multiplication with positive rationals to weighted multigraphs via pointwise addition and scaling of the weight functions. In particular, we can identify an ordinary multigraph with the weighted multigraph that has all edge weights equal to 1, and the two definitions of ω coincide in this case. We observe that Lemma 17 carries over.

Lemma 35. *Let G be a multigraph, and let $\alpha \in \mathbb{Q}_{>0}$. We denote by $\alpha \cdot G$ the fractional graph with edge weights multiplied by α. Then,*

$$\omega(\alpha \cdot G) = \alpha\omega(G).$$

Proof. Pick $q \in \mathbb{N}$ such that $q\alpha \in \mathbb{N}$. By definition of ω for fractional graphs and Lemma 17, we have

$$\omega(\alpha \cdot G) = \omega(q\alpha \cdot G)/q = q\alpha \cdot \omega(G)/q = \alpha\omega(G). \qquad \square$$

Remark 36. Lemma 17 also implies that

$$\omega\left(\sum_i \alpha_i \cdot G_i\right) \leq \sum_i \alpha_i \cdot \omega(G) \tag{21}$$

holds for any finite sum of arbitrary fractional graphs $\alpha_i \cdot G_i$.

For a fractional graph $G = (V, E, w)$, we define a *conic decomposition* of G to be a finite sequence $\mathcal{H}$ of fractional graphs H such that

$$\sum_{H \in \mathcal{H}} H = G,$$

and we define

$$\omega(\mathcal{H}) = \sum_{H \in \mathcal{H}} \omega(H).$$

We denote the set of all such conic decompositions of G by $\mathcal{C}(G)$. Note that we consider G as its own conic decomposition, hence the 1-element sequence (G) is a valid conic decomposition of G.

Lemma 37. *Let G be a multigraph. Then*

$$\omega(G) = \min_{\mathcal{H} \in \mathcal{C}(G)} \omega(\mathcal{H}).$$

Proof. This is simply by (21), observing that the minimum is attained by definition at $\omega(G)$ itself. □

Remark 38. Lemma 37 seems trivial, but the point is that it can be used to obtain upper bounds on $\omega(G)$ for graphs G by decomposing G into fractional subgraphs H for which we have good bounds on $\omega(H)$, such as, for instance, fractional multiples of triangles.

5.3 Deriving Concrete Bounds

The above notions and results we will now use to prove tighter upper bounds in the intermediate regime of $d = 4, 5$. Let us reiterate here that this selection of results is to be understood as a basic demonstration of the viability and utility of the approach, rather than a systematic and exhaustive exploitation of its merits.

Sums of Stars. Recall from the proof of Lemma 20 that, for any tensor $T \in (\mathbb{F}^n)^{\otimes d}$, we have that $T \leq T_{S_d(d),n}$, where $S_d(d)$ is the star graph on d vertices with the central vertex labeled by d. Define $S_d^k = T_{\sum_{u=1}^k S_d(u),n}$. Now, consider some $k \leq d$, and recall that

$$T^{\otimes k} \leq T_{S_d^k,n} \tag{22}$$

where we can visualize S_d^k as the result of overlaying k star graphs with different central vertices.

One can view S_d^k as a k-clique with double edges between each pair of vertices, with $d - k$ extra vertices that are connected to each vertex in this k-clique, but not to each other. For example, Fig. 6 shows the graphs S_4^2 and S_5^3. Then, our upper bounds follow by using Theorem 5 to separate a given S_d^k into two graphs G_1, G_2 such that $S_d^k = G_1 + G_2$. Specifically, we have that

$$\underset{\sim}{\mathrm{C}}(T_{S_d^k,n}) \leq \underset{\sim}{\mathrm{R}}(T_{G_1,n}) \cdot \underset{\sim}{\mathrm{C}}(T_{G_2,n}).$$

Thus, to upper bound the asymptotic circuit complexity of a sum of star graphs, we just need to upper bound the asymptotic rank and asymptotic circuit complexities of a partitioning of the edges into two parts. Upper-bounding $\underset{\sim}{\mathrm{C}}(T_{G_2,n})$ is conceptually easy using Lemma 30 and (18), as we simply need to compute the line graph and then find its tree width.

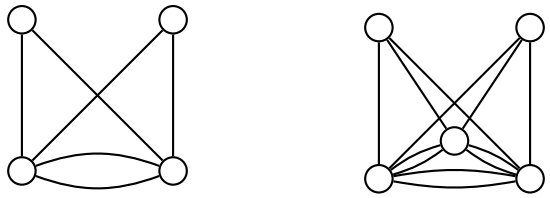


Figure 6: The graph S_4^2 (left) and S_5^3 (right).

Fractional Triangle Coverings. Now, we need a way to upper bound $\underset{\sim}{\mathrm{R}}(T_{G_1,n})$, which we will do using fractional triangle coverings. Beyond the fractional tensors introduced in Section 5.2, this requires one additional insight. Consider $\Delta_{d,t}(i,j,k)$, the fractional d-vertex graph with a single edge of weight $t \in \mathbb{Q}$ between j,k, and an edge of weight 1 between i,j and k,i each. We call this graph a *t-triangle*. Moreover, let $\mathrm{MM}_n(t)$ denote the matrix multiplication tensor of an $n \times n^t$ matrix and an $n^t \times n$ matrix. It is easy to see that for all choices of d,t,i,j,k and n, we have that

$$T_{\Delta_{d,t}(i,j,k),n} \equiv \mathrm{MM}_n(t)$$

holds. As is common in the literature, we write $\omega(t)$ as a shorthand for $\omega(\mathrm{MM}_n(t))$, the tensor exponent of $\mathrm{MM}_n(t)$. As an application of Lemma 37, we hence find that

$$\omega(T_{G,n}) \leq \sum_{i=1}^{\ell} \omega(\alpha_i)$$

holds for every conic decomposition of G into fractional graphs $\Delta_1, \ldots, \Delta_\ell$, where Δ_i is an α_i-triangle for $i = 1, \ldots, \ell$.

It is therefore enough to find a partition of the graph into two parts, one being decomposed into such a set of fractional triangles, and the other according to the bound on the asymptotic circuit complexity via line graph treewidth, as in Lemma 30. The objective is to minimize the sum of the resulting asymptotic exponents in both parts. From a practical computational perspective, this can be done by a combination of brute-force and a simple linear program. For the upper bounds on $\omega(k)$, we used the values calculated in [24]:

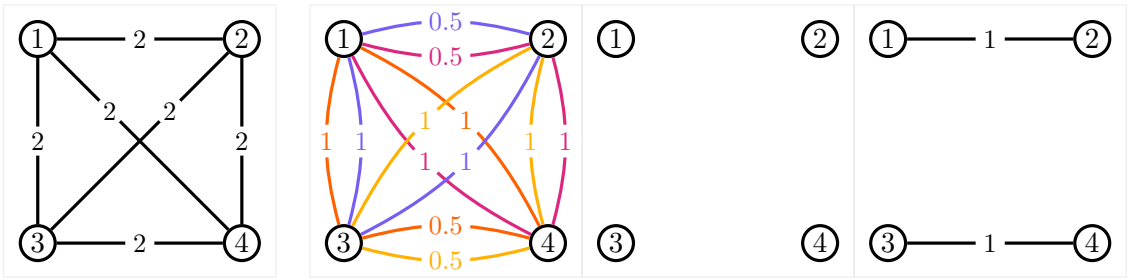
$$\omega(0.5) \leq 2.046681, \qquad \omega(2) \leq 3.256689.$$

We can then formalize this idea into the following lemma. Here, the graph G_2 can be viewed as the "leftover edges" that are handled in a brute-force manner.

Lemma 39. *Let G be a weighted graph and G_1, G_2, G_3 such that $G = G_1 + G_2 + G_3$, G_1 has a conic decomposition $\mathcal{H}$ into fractional triangles and G_2 as well as G_3 are ordinary multigraphs, then we have that*

$$\underset{\sim}{\mathrm{C}}(T_G) \leq n^{\omega(\mathcal{H})+|E(G_2)|+\mathrm{ltw}(G_3)+1}.$$

Proof. By Theorem 5 and Lemma 30, $\underset{\sim}{\mathrm{C}}(T_{G,n}) \leq \underset{\sim}{\mathrm{R}}(T_{G_1,n}) \cdot \underset{\sim}{\mathrm{R}}(T_{G_2,n}) \cdot \underset{\sim}{\mathrm{C}}(T_{G_3,n}) \leq n^{\omega(\mathcal{H})} \cdot n^{|E(G_2)|} \cdot n^{\mathrm{ltw}(G_3)+1}$. □



$$
\begin{array}{ccccccc}
2 \cdot K_4 & = & G_1 & + & G_2 & + & G_3 \\
\underset{\sim}{\mathrm{C}}(T_{2\cdot K_4,n}) & \leq & \underset{\sim}{\mathrm{R}}(T_{G_1,n}) & \cdot & \underset{\sim}{\mathrm{R}}(T_{G_2,n}) & \cdot & \underset{\sim}{\mathrm{C}}(T_{G_3,n}) \\
& \leq & n^{4\cdot\omega(0.5)} & \cdot & n^0 & \cdot & n^1
\end{array}
$$

Figure 7: Decomposition of $2 \cdot K_4$ into G_1, G_2, G_3, where G_1 has a conic decomposition into four 0.5-triangles indicated by edge colors, G_2 is an empty graph, and G_3 is a matching.

Bounds. Our experimental programs were able to obtain results on 4 and 5 vertices. Larger number of vertices seemed to take too long to run to completion. For these two cases, we provide our results in the following two theorems.

Theorem 40. *Every 4-mode tensor* $T \in (\mathbb{F}^n)^{\otimes 4}$ *satisfies* $\underset{\sim}{\mathrm{C}}(T) \leq n^{1/4+\omega(0.5)} \leq n^{2.296681}$.

Proof. From (22), we see that $T^{\otimes 4} \leq T_{2\cdot K_4,n}$, where $2 \cdot K_4 = S_4^4$ is the doubled complete 4-vertex graph. Using Lemma 39 and the decomposition of Figure 7, we observe that $\underset{\sim}{\mathrm{C}}(T_{2\cdot K_4,n}) \leq n^{1+4\cdot\omega(0.5)}$. Hence, $\underset{\sim}{\mathrm{C}}(T) \leq \underset{\sim}{\mathrm{C}}(T^{\otimes 4})^{1/4} \leq n^{1/4+\omega(0.5)}$. □

Theorem 41. *Every 5-mode tensor* $T \in (\mathbb{F}^n)^{\otimes 5}$ *satisfies* $\underset{\sim}{\mathrm{C}}(T) \leq n^{(3+\omega+\omega(2))/3} \leq n^{2.877389}$.

Proof. By (22), $T^{\otimes 3} \leq T_{S_5^3}$. Then, using the decomposition given in Figure 8 that picks $G_1 = \Delta_{5,1}(1,3,5) + \Delta_{5,2}(2,1,3)$, G_2 as the single edge $4,2$, and G_3 as the 4-edge path $3,4,1,2,5$, we observe via Lemma 39 that

$$\underset{\sim}{\mathrm{C}}(T_{S_5^3,n}) \leq n^{\omega+\omega(2)+|E(G_2)|+\mathrm{ltw}(G_3)+1} \leq n^{\omega+\omega(2)+3},$$

given that the line graph of a path is a path and hence has treewidth 1. Hence, we have that

$$\underset{\sim}{\mathrm{C}}(T) \leq \underset{\sim}{\mathrm{C}}(T^{\otimes 3})^{1/3} \leq \underset{\sim}{\mathrm{C}}(T_{S_5^3,n})^{1/3} \leq \underset{\sim}{\mathrm{C}}(T_{S_5^3})^{1/3} \leq n^{1+(\omega+\omega(2))/3}.$$ □

The preceding theorems together constitute Theorem 7 from the introduction. We remark here that our results do not rely on the improved bounds obtained for the rank of $T_{K_4,n}$ referred to in Theorem 1. Judging from computational evidence, the choice of $d = 4, 5$ seems to be too small to profit from the edge afforded by this bound. For higher values of d than those handled by our code, this could become more relevant.

References

[1] Per Austrin, Petteri Kaski, and Kaie Kubjas. Tensor network complexity of multilinear maps. *Theory Comput.*, 18:1–54, 2022. doi:10.4086/TOC.2022.V018A016.

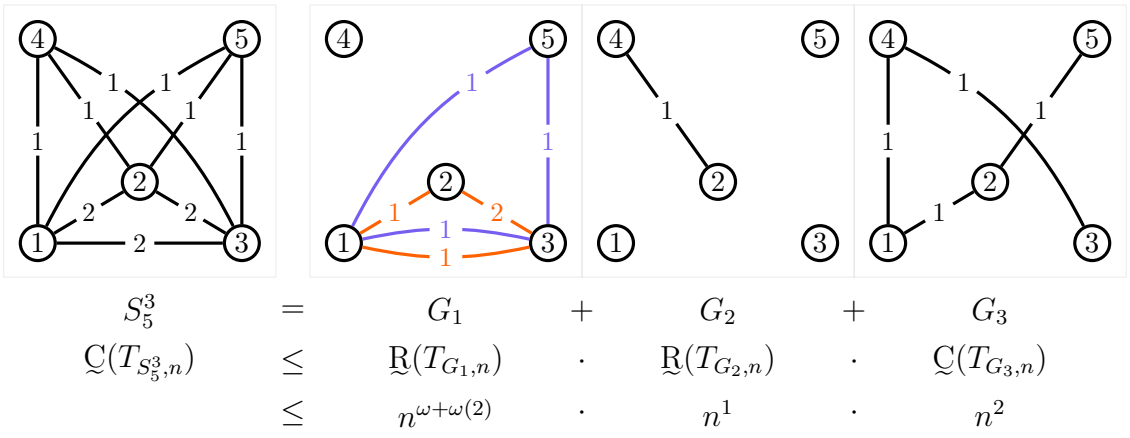


Figure 8: Decomposition of S_5^3 into G_1, G_2, G_3, where G_1 has a conic decomposition into a 2-triangle and a 1-triangle, G_2 contains a single edge, and G_3 is a path.

[2] Bozhena Bidyuk and Rina Dechter. Cutset sampling for Bayesian networks. *J. Artif. Intell. Res.*, 28:1–48, 2007. doi:10.1613/JAIR.2149.

[3] Andreas Björklund, Radu Curticapean, Thore Husfeldt, Petteri Kaski, and Kevin Pratt. Fast deterministic chromatic number under the asymptotic rank conjecture. In Yossi Azar and Debmalya Panigrahi, editors, *Proceedings of the 2025 Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2025, New Orleans, LA, USA, January 12-15, 2025*, pages 2804–2818. SIAM, 2025. doi:10.1137/1.9781611978322.91.

[4] Andreas Björklund and Petteri Kaski. The asymptotic rank conjecture and the set cover conjecture are not both true. In Bojan Mohar, Igor Shinkar, and Ryan O'Donnell, editors, *Proceedings of the 56th Annual ACM Symposium on Theory of Computing, STOC 2024, Vancouver, BC, Canada, June 24-28, 2024*, pages 859–870. ACM, 2024. doi:10.1145/3618260.3649656.

[5] Andreas Björklund, Petteri Kaski, Tomohiro Koana, and Jesper Nederlof. Kronecker scaling of tensors with applications to arithmetic circuits and algorithms. *CoRR*, abs/2504.05772, 2025. arXiv:2504.05772, doi:10.48550/ARXIV.2504.05772.

[6] Cornelius Brand, Radu Curticapean, Baitian Li, and Kevin Pratt. Faster convolutions: Yates and Strassen revisited. In *Proceedings of the 2026 SIAM Symposium on Simplicity in Algorithms (SOSA)*, pages 328–339, 2026. doi:10.1137/1.9781611978964.25.

[7] Karl Bringmann and Jasper Slusallek. Current algorithms for detecting subgraphs of bounded treewidth are probably optimal. In Nikhil Bansal, Emanuela Merelli, and James Worrell, editors, *48th International Colloquium on Automata, Languages, and Programming, ICALP 2021, Glasgow, Scotland (Virtual Conference), July 12-16, 2021*, volume 198 of *LIPIcs*, pages 40:1–40:16. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2021. doi:10.4230/LIPICS.ICALP.2021.40.

[8] Peter Bürgisser, Michael Clausen, and M. Amin Shokrollahi. *Algebraic Complexity Theory*, volume 315 of *Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1997. With the collaboration of Thomas Lickteig. doi:10.1007/978-3-662-03338-8.

[9] Jin-Yi Cai and Xi Chen. *Complexity Dichotomies for Counting Problems: Volume 1, Boolean Domain*. Cambridge University Press, 2017. doi:10.1017/9781107477063.

[10] Jin-Yi Cai, Heng Guo, and Tyson Williams. A complete dichotomy rises from the capture of vanishing signatures. *SIAM J. Comput.*, 45(5):1671–1728, 2016. doi:10.1137/15M1049798.

[11] Jin-yi Cai and Pinyan Lu. Holographic algorithms: From art to science. *J. Comput. Syst. Sci.*, 77(1):41–61, 2011. doi:10.1016/J.JCSS.2010.06.005.

[12] Jin-Yi Cai and Ben Young. Vanishing signatures, orbit closure, and the converse of the holant theorem. In Shubhangi Saraf, editor, *17th Innovations in Theoretical Computer Science Conference, ITCS 2026, Bocconi University, Milan, Italy, January 27-30, 2026*, volume 362 of *LIPIcs*, pages 32:1–32:20. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2026. doi:10.4230/LIPICS.ITCS.2026.32.

[13] Matthias Christandl, Péter Vrana, and Jeroen Zuiddam. Asymptotic tensor rank of graph tensors: beyond matrix multiplication. *Comput. Complex.*, 28(1):57–111, 2019. doi:10.1007/S00037-018-0172-8.

[14] Matthias Christandl and Jeroen Zuiddam. Tensor surgery and tensor rank. *Comput. Complex.*, 28(1):27–56, 2019. doi:10.1007/S00037-018-0164-8.

[15] J. Ignacio Cirac, David Pérez-García, Norbert Schuch, and Frank Verstraete. Matrix product states and projected entangled pair states: Concepts, symmetries, theorems. *Rev. Mod. Phys.*, 93:045003, Dec 2021. doi:10.1103/RevModPhys.93.045003.

[16] Don Coppersmith and Shmuel Winograd. Matrix multiplication via arithmetic progressions. *Journal of Symbolic Computation*, 9(3):251–280, 1990. doi:10.1016/S0747-7171(08)80013-2.

[17] Radu Curticapean, Simon Döring, Daniel Neuen, and Jiaheng Wang. Can you link up with treewidth? In Olaf Beyersdorff, Michal Pilipczuk, Elaine Pimentel, and Kim Thang Nguyen, editors, *42nd International Symposium on Theoretical Aspects of Computer Science, STACS 2025, Jena, Germany, March 4-7, 2025*, volume 327 of *LIPIcs*, pages 28:1–28:24. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2025. doi:10.4230/LIPICS.STACS.2025.28.

[18] Radu Curticapean and Dániel Marx. Complexity of counting subgraphs: Only the boundedness of the vertex-cover number counts. In *55th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2014, Philadelphia, PA, USA, October 18-21, 2014*, pages 130–139. IEEE Computer Society, 2014. doi:10.1109/FOCS.2014.22.

[19] Marek Cygan, Fedor V. Fomin, Lukasz Kowalik, Daniel Lokshtanov, Dániel Marx, Marcin Pilipczuk, Michal Pilipczuk, and Saket Saurabh. *Parameterized Algorithms*. Springer, 2015. doi:10.1007/978-3-319-21275-3.

[20] Anuj Dawar, Benedikt Pago, and Tim Seppelt. Symmetric algebraic circuits and homomorphism polynomials, 2025. URL: https://arxiv.org/abs/2502.06740v3, arXiv:2502.06740v3.

[21] Rina Dechter. Bucket elimination: A unifying framework for reasoning. *Artif. Intell.*, 113(1-2):41–85, 1999. doi:10.1016/S0004-3702(99)00059-4.

[22] Praeteek Dwivedi, Benedikt Pago, and Tim Seppelt. Lower bounds in algebraic complexity via symmetry and homomorphism polynomials, 2026. URL: http://arxiv.org/abs/2601.09343.

[23] Jörg Flum and Martin Grohe. The parameterized complexity of counting problems. *SIAM J. Comput.*, 33(4):892–922, 2004. doi:10.1137/S0097539703427203.

[24] Francois Le Gall and Florent Urrutia. Improved rectangular matrix multiplication using powers of the Coppersmith–Winograd tensor. In Artur Czumaj, editor, *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2018, New Orleans, LA, USA, January 7-10, 2018*, pages 1029–1046. SIAM, 2018. doi:10.1137/1.9781611975031.67.

[25] Egor Gorbachev and Marvin Künnemann. Combinatorial designs meet hypercliques: Higher lower bounds for Klee's measure problem and related problems in dimensions $d \geq 4$. In Erin W. Chambers and Joachim Gudmundsson, editors, *39th International Symposium on Computational Geometry, SoCG 2023, Dallas, Texas, USA, June 12-15, 2023*, volume 258 of *LIPIcs*, pages 36:1–36:14. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2023. doi:10.4230/LIPICS.SOCG.2023.36.

[26] Daniel J. Harvey and David R. Wood. Treewidth of the line graph of a complete graph. *J. Graph Theory*, 79(1):48–54, 2015. doi:10.1002/JGT.21813.

[27] Daniel J. Harvey and David R. Wood. The treewidth of line graphs. *J. Comb. Theory B*, 132:157–179, 2018. doi:10.1016/J.JCTB.2018.03.007.

[28] William Huggins, Piyush Patil, Bradley Mitchell, K Birgitta Whaley, and E Miles Stoudenmire. Towards quantum machine learning with tensor networks. *Quantum Science and Technology*, 4(2):024001, Jan 2019. doi:10.1088/2058-9565/aaea94.

[29] Karthik C. S., Dániel Marx, Marcin Pilipczuk, and Uéverton S. Souza. Conditional lower bounds for sparse parameterized 2-CSP: A streamlined proof. In Merav Parter and Seth Pettie, editors, *2024 Symposium on Simplicity in Algorithms, SOSA 2024, Alexandria, VA, USA, January 8-10, 2024*, pages 383–395. SIAM, 2024. doi:10.1137/1.9781611977936.35.

[30] Donald E. Knuth. *The Art of Computer Programming. Vol. 2.* Addison-Wesley, Reading, MA, third edition, 1998. Seminumerical Algorithms.

[31] Daphne Koller and Nir Friedman. *Probabilistic Graphical Models.* Adaptive Computation and Machine Learning. MIT Press, Cambridge, MA, 2009. Principles and techniques.

[32] Frank R. Kschischang, Brendan J. Frey, and Hans-Andrea Loeliger. Factor graphs and the sum-product algorithm. *IEEE Trans. Inform. Theory*, 47(2):498–519, 2001. doi:10.1109/18.910572.

[33] Marvin Künnemann. A tight (non-combinatorial) conditional lower bound for Klee's measure problem in 3D. In *63rd IEEE Annual Symposium on Foundations of Computer Science, FOCS 2022, Denver, CO, USA, October 31 - November 3, 2022*, pages 555–566. IEEE, 2022. doi:10.1109/FOCS54457.2022.00059.

[34] J. M. Landsberg. *Tensors: Geometry and Applications*, volume 128 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2012. doi:10.1090/gsm/128.

[35] S. L. Lauritzen and D. J. Spiegelhalter. Local computations with probabilities on graphical structures and their application to expert systems. *J. R. Stat. Soc. Ser. B. Stat. Methodol.*, 50(2):157–194, 1988. doi:10.1111/j.2517-6161.1988.tb01721.x.

[36] Andrea Lincoln, Virginia Vassilevska Williams, and R. Ryan Williams. Tight hardness for shortest cycles and paths in sparse graphs. In Artur Czumaj, editor, *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2018, New Orleans, LA, USA, January 7-10, 2018*, pages 1236–1252. SIAM, 2018. doi:10.1137/1.9781611975031.80.

[37] Dániel Marx. Can you beat treewidth? *Theory Comput.*, 6(1):85–112, 2010. doi:10.4086/TOC.2010.V006A005.

[38] Román Orús. A practical introduction to tensor networks: Matrix product states and projected entangled pair states. *Ann. Physics*, 349:117–158, 2014. doi:10.1016/j.aop.2014.06.013.

[39] Judea Pearl. Fusion, propagation, and structuring in belief networks. *Artif. Intell.*, 29(3):241–288, 1986. doi:10.1016/0004-3702(86)90072-X.

[40] Roger Penrose and Wolfgang Rindler. *Spinors and Space-Time. Vol. 1.* Cambridge Monographs on Mathematical Physics. Cambridge University Press, Cambridge, 1987. Two-spinor calculus and relativistic fields. doi:10.1017/CBO9780511564048.

[41] Kevin Pratt. A stronger connection between the asymptotic rank conjecture and the set cover conjecture. In Bojan Mohar, Igor Shinkar, and Ryan O'Donnell, editors, *Proceedings of the 56th Annual ACM Symposium on Theory of Computing, STOC 2024, Vancouver, BC, Canada, June 24-28, 2024*, pages 871–874. ACM, 2024. doi:10.1145/3618260.3649620.

[42] Irina Rish and Rina Dechter. Resolution versus search: Two strategies for SAT. *J. Autom. Reason.*, 24(1/2):225–275, 2000. doi:10.1023/A:1006303512524.

[43] Ross D. Shachter, Stig K. Andersen, and Peter Szolovits. Global conditioning for probabilistic inference in belief networks. In Ramón López de Mántaras and David Poole, editors, *UAI '94: Proceedings of the Tenth Annual Conference on Uncertainty in Artificial Intelligence, Seattle, Washington, USA, July 29-31, 1994*, pages 514–522. Morgan Kaufmann, 1994. doi:10.1016/B978-1-55860-332-5.50070-5.

[44] Glenn Shafer and Prakash P. Shenoy. Probability propagation. *Ann. Math. Artif. Intell.*, 2:327–351, 1990. doi:10.1007/BF01531015.

[45] Shuai Shao and Jin-Yi Cai. A dichotomy for real boolean holant problems. In Sandy Irani, editor, *61st IEEE Annual Symposium on Foundations of Computer Science, FOCS 2020, Durham, NC, USA, November 16-19, 2020*, pages 1091–1102. IEEE, 2020. doi:10.1109/FOCS46700.2020.00105.

[46] V. Strassen. The asymptotic spectrum of tensors. *J. Reine Angew. Math.*, 384:102–152, 1988. doi:10.1515/crll.1988.384.102.

[47] V. Strassen. Algebra and complexity. In *First European Congress of Mathematics, Vol. II (Paris, 1992)*, volume 120 of *Progr. Math.*, pages 429–446. Birkhäuser, Basel, 1994. doi:10.1007/978-3-0348-9112-7_18.

[48] Leslie G. Valiant. Completeness classes in algebra. In Michael J. Fischer, Richard A. DeMillo, Nancy A. Lynch, Walter A. Burkhard, and Alfred V. Aho, editors, *Proceedings of the 11h Annual ACM Symposium on Theory of Computing, April 30 - May 2, 1979, Atlanta, Georgia, USA*, pages 249–261. ACM, 1979. doi:10.1145/800135.804419.

[49] Leslie G. Valiant. Holographic algorithms. *SIAM J. Comput.*, 37(5):1565–1594, 2008. doi:10.1137/070682575.

[50] Avi Wigderson and Jeroen Zuiddam. Asymptotic spectra: Theory, applications and extensions. *Manuscript*, 2022. URL: https://www.math.ias.edu/~avi/PUBLICATIONS/WigdersonZu_Final_Draft_Oct2023.pdf.

[51] Frank Yates. *The Design and Analysis of Factorial Experiments.* Imperial Bureau of Soil Science. Technical Communication, no. 35. Imperial Bureau of Soil Science, Harpenden, 1937.

[52] Yitong Yin and Chihao Zhang. Approximate counting via correlation decay on planar graphs. In Sanjeev Khanna, editor, *Proceedings of the Twenty-Fourth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2013, New Orleans, Louisiana, USA, January 6-8, 2013*, pages 47–66. SIAM, 2013. doi:10.1137/1.9781611973105.4.

[53] N. L. Zhang and D. L. Poole. A simple approach to Bayesian network computations. In *Proc. Tenth Canadian Conference on Artificial Intelligence*, pages 171–178, 1994.

A Properties of Graph Tensors

This appendix collects and proves some further basic facts about graph tensors in relation to their graphs, including proofs of Lemma 15, 16 and 17.

Lemma 42 (Isomorphism implies equivalence)**.** *Let G and H be graphs and let $n \in \mathbb{N}$. If $G \equiv H$, then $T_{G,n} \equiv T_{H,n}$.*

Proof. Recall Eq. (8) and write $x_g^{(v)}$ for $v \in V(G)$ and $g : I_G(v) \to [n]$ for the indeterminates of the polynomial $T_{G,n}$. Similarly, write $y_h^{(v)}$ for $v \in V(H)$ and $g : I_H(v) \to [n]$ for the indeterminates of the polynomial $T_{H,n}$. Let the bijections $\phi : V(H) \to V(H)$ and $\psi : E(G) \to E(H)$ be an isomorphism from G to H. The substitution $x_g^{(v)} \coloneqq y_{\psi(g)}^{(\phi(v))}$ obtains $T_{G,n}$ from $T_{H,n}$, and the substitution $y_g^{(v)} \coloneqq x_{\psi^{-1}(g)}^{(\phi^{-1}(v))}$ obtains $T_{H,n}$ from $T_{G,n}$. □

Lemma 43 (Projection under length reduction)**.** *Let G be a graph. If $m \leq n$, then $T_{G,m} \leq_{\mathrm{s}} T_{G,n}$.*

Proof. In the polynomial $T_{G,n}$, substitute 0 to every indeterminate $x_g^{(v)}$ such that $v \in V(G)$, $g : I_G(v) \to [n]$, and there is an $e \in I_G(v)$ such that $g(e) \geq m+1$. The resulting polynomial equals the polynomial $T_{G,m}$. □

Lemma 44 (Projection under subgraphs)**.** *Let G and H be graphs and let $n \in \mathbb{N}$. If H is a subgraph of G, then $T_{H,n} \leq_{\mathrm{s}} T_{G,n}$.*

Proof. Recall that

$$T_{G,n} = \sum_{f:E(G)\to[n]} \prod_{v\in V(G)} x^{(v)}_{f|_{I_G(v)}} .$$

Now $T_{G,n}$ can be projected by substituting, for all $v \in V(G)$ and $g : I_G(v) \to [n]$,

$$x_g^{(v)} := \begin{cases} 1, & \text{if } v \notin V(H), \\ x^{(v)}_{g|_{I_H(v)}}, & \text{if } v \in V(H) \text{ and } g(e) = 1 \text{ for all } e \in I_G(v) \setminus I_H(v), \\ 0, & \text{otherwise.} \end{cases}$$

Indeed, under this substitution, it holds that

$$T_{H,n} = \sum_{f::E(H)\to[n]} \prod_{v\in V(H)} x^{(v)}_{f|_{I_H(v)}} \leq_{\mathrm{s}} \sum_{\substack{f:E(H)\to[n] \\ f':E(G)\setminus E(H)\to[n]}} \prod_{v\in V} x^{(v)}_{f\cup f'|_{I_G(v)}} = T_{G,n}$$

as desired. □

We are now ready to prove Lemma 15.

Proof of Lemma 15. From Lemma 44 it follows that it suffices to consider the case when G is obtained from H by subdividing some edge $e \in E(H)$ with ends $u, v \in V(G)$. Write $w \in V(G)$ for the vertex introduced by this subdivision. Let us also write uv and vw for the edges introduced by the subdivision. We thus have $V(G) = V(H) \cup \{w\}$ and $E(G) = (E(H) \setminus \{e\}) \cup \{uw, vw\}$. It holds that

$$\begin{aligned} T_{G,n} &= \sum_{f:E(G)\to[n]} \prod_{a\in V(G)} x^{(a)}_{f|_{I_G(a)}} \\ &= \sum_{f:E(H)\setminus\{e\}\to[n]} \sum_{i,j\in[n]} x^{(w)}_{uw\mapsto i,vw\mapsto j} \cdot x^{(u)}_{\{uw\mapsto i\}\cup f|_{I_H(u)\setminus\{e\}}} \cdot x^{(v)}_{\{vw\mapsto j\}\cup f|_{I_H(v)\setminus\{e\}}} \cdot \prod_{b\in V(H)\setminus\{u,v\}} x^{(b)}_{f|_{I_H(b)}} . \end{aligned}$$

By substituting $x^{(w)}_{uw\mapsto i,vw\mapsto j} := \delta_{ij}$ into the polynomial above, we obtain the polynomial

$$\sum_{f:E\to[n]} \prod_{v\in V} x^{(v)}_{f|_{I_H}(v)} = T_{H,n} ,$$

implying $T_{H,n} \leq_{\mathrm{s}} T_{G,n}$ as desired. □

With a standard argument, we can also prove Lemma 16 on contractions:

Proof of Lemma 16. Let $E \coloneqq E(G)$. Write $A = E(G[U])$ for the edges within U, and let w be the vertex replacing U in G/U; this vertex is incident with the edges in G that have exactly one endpoint in U.

Given functions f, g on disjoint domains, we write fg for the function that agrees with f and g on their respective domains. We have

$$\begin{aligned} T_{G,n} &= \sum_{f\colon E\to[n]} \prod_{v\in V} x^{(v)}_{f|_{I(v)}} \\ &= \sum_{f\colon E\setminus A\to[n]} \underbrace{\left(\sum_{g\colon A\to[n]} \prod_{v\in U} x^{(v)}_{fg|_{I(v)}} \right)}_{=h_{f|_{I(w)}}} \left(\prod_{v\in V\setminus U} x^{(v)}_{f|_{I(v)}} \right) \\ &= \sum_{f\colon E\setminus A\to[n]} h_{f|_{I(w)}} \prod_{v\in V\setminus U} x^{(v)}_{f|_{I(v)}}. \end{aligned}$$

Each polynomial $h_{f|_{I(w)}}$ admits a circuit of size $O(|U|\cdot n^{|A|})$, and the last right-hand side is obtained from $T_{G/U,n}$ by the substitution $x^{(w)}_{f|_{I(w)}} \leftarrow h_{f|_{I(w)}}$ in the mode for w, which represents U. □

We also include for completeness:

Proof Lemma 17. By Lemma 12, we have $\mathrm{R}(T_{G+H,n}) = \mathrm{R}(T_{G,n} \otimes T_{H,n}) \le \mathrm{R}(T_{G,n}) \cdot \mathrm{R}(T_{H,n})$, so $\omega(G+H) \le \omega(G) + \omega(H)$ and $\omega(k\cdot G) \le k\cdot\omega(G)$ follow. For the other inequality, let $\beta = \omega(k\cdot G)$, and hence $\mathrm{R}(T_{k\cdot G,n}) \le cn^\beta$ for some $c > 0$ and all large enough n. By the sum rule (Remark 14) this means that also $\mathrm{R}(T_{G,n^k}) \le cn^\beta$ for large enough n. For $\bar{n} = \lceil n^{1/k}\rceil^k \ge n$ we have that $T_{G,n} \le_{\mathrm{s}} T_{G,\bar{n}}$. Hence, for large enough n, $\mathrm{R}(T_{G,n}) \le \mathrm{R}(T_{G,\bar{n}}) \le c\lceil n^{1/k}\rceil^\beta \le c(n^{1/k}+1)^\beta \le c2^\beta n^{\beta/k} = O(n^{\beta/k})$. By definition of $\omega(G)$ we conclude $\omega(G) \le \beta/k = \omega(k\cdot G)/k$. □

B Facts on Asymptotic Circuit Complexity

This appendix collects select facts on asymptotic circuit complexity.

Fact 45. *Let $U, V \in (\mathbb{F}^n)^{\otimes d}$ with $V \le U$. Then*

$$\underline{\mathrm{C}}(V) \le \underline{\mathrm{C}}(U)$$

holds.

Proof. The claim follows by an argument close to the proof of Lemma 33 and Theorem 5 as an application of tensor contraction. Let $L_1, \dots, L_d$ be the linear maps that witness $V \le U$, that is, $V = U(L_1(x^{(1)}), \dots, L_d(x^{(d)}))$. Evaluation can be shown to be compatible with Kronecker powering, so that $V^{\otimes k} = U^{\otimes k}(L_1^{\otimes k}(z^{(1)}), \dots, L_d^{\otimes k}(z^{(d)}))$. The claim then follows by Yates's algorithm (7) as in the proof of Theorem 5. □

Fact 46. *Let $U \in (\mathbb{F}^n)^{\otimes d}, V \in (\mathbb{F}^m)^{\otimes d}$ with $V \neq 0$. Then, $U \le U \otimes V$, and $\mathrm{C}(U) \le \mathrm{C}(U \otimes V)$.*

Proof. Suppose the coefficients of U and V are

$$U = \sum_{i_1,\ldots,i_d=1}^{n} u_{i_1,\ldots,i_d} \cdot x_{i_1}^{(1)} \cdots x_{i_d}^{(d)}, \quad V = \sum_{j_1,\ldots,j_d=1}^{m} v_{j_1,\ldots,j_d} \cdot y_{j_1}^{(1)} \cdots y_{j_d}^{(d)}$$

and assume without loss of generality that $v_{1,\ldots,1} \neq 0$.

Let C be circuit computing $U \otimes V$ with modes $Z^{(i)}$ containing variables $z_{k,\ell}^{(i)}$, that is

$$U \otimes V = \sum_{i_1,\ldots,i_d=1}^{n} \sum_{j_1,\ldots,j_d=1}^{m} u_{i_1,\ldots,i_d} \cdot v_{j_1,\ldots,j_d} \cdot z_{i_1,j_1}^{(1)} \cdots z_{i_d,j_d}^{(d)}.$$

Define the substitution μ via $\mu(z_{k,1}^{(1)}) = \frac{1}{v_{1,\ldots,1}} x_k^{(1)}$, $\mu(z_{k,1}^{(i)}) = x_k^{(i)}$ for $i > 1$ and all k, and $\mu(z_{k,\ell}^{(i)}) = 0$ for all i and $\ell > 1$. Since only terms with $j_1, \ldots, j_d = 1$ survive in $U \otimes V$ under this substitution, we find

$$(U \otimes V)^{\mu} = \sum_{i_1,\ldots,i_d=1}^{n} u_{i_1,\ldots,i_d} \cdot v_{1\ldots1} \cdot \frac{1}{v_{1\ldots1}} \cdot x_{i_1}^{(1)} \cdots x_{i_d}^{(d)} = U.$$

Note that in our circuit model, this substitution can be performed without any size increase in C, by either replacing inputs with 0, or substituting inputs for another with the scalar factor $1/v_{1\ldots1}$ possibly at the appropriate wires, which is free. The claim follows. (Note that for general projections as afforded via $\leq$ instead of $\leq_{\mathrm{s}}$, the translation on the circuit sizes is not automatic, necessitating a closer look as given in this proof.) □

Fact 47. *If* C *is submultiplicative on d'-mode tensors, for any fixed $d' \in \mathbb{N}$, then* C *is submultiplicative on d-mode tensors for all $1 \leq d \leq d'$.*

Proof. By induction, it suffices to consider the case when $d' = d + 1$. Suppose that C is submultiplicative on $(d+1)$-mode tensors. Let $S \in \mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_d}$ and $T \in \mathbb{C}^{m_1} \otimes \cdots \otimes \mathbb{C}^{m_d}$ for $n_1, \ldots, n_1, m_1, \ldots, m_d \geq 0$. Extend S and T to $(d+1)$-mode tensors $\widehat{S}$ and $\widehat{T}$ by tensoring with 1, i.e. appending a 0-dimensional mode. It holds that $\mathrm{C}(S) = \mathrm{C}(\widehat{S})$ and furthermore

$$\mathrm{C}(S \otimes T) = \mathrm{C}(\widehat{S} \otimes \widehat{T}) \leq \mathrm{C}(\widehat{S}) \cdot \mathrm{C}(\widehat{T}) = \mathrm{C}(S) \cdot \mathrm{C}(T),$$

as desired. □

Yates's Algorithm. We now briefly recall Yates's algorithm and give a generalized exposition that includes the case $d \geq 4$. Suppose $T \in (\mathbb{F}^n)^{\otimes d}$ is concise, such that $T = \sum_{i=1}^{r} \prod_{j=1}^{d} \ell_i^{(j)}(x^{(j)})$, where $\ell_i^{(j)}(x^{(j)}) = \sum_{p=1}^{n} \ell_{ip}^{(j)} x_p^{(j)}$ are linear forms and $r \geq n$. For $j \in [d]$ and $k \in \mathbb{N}$, let us write $z^{(j)}$ for the indeterminates of mode j of $T^{\otimes k}$. We can thus view the tensor $T^{\otimes k}$ as the polynomial

$$T^{\otimes k}(z^{(1)}, \ldots, z^{(d)}) = \sum_{i \in [r]^k} \prod_{j=1}^{d} \ell_i^{(j)}(z^{(j)}), \tag{23}$$

where for $i = (i_1, \ldots, i_k) \in [r]^k$ we define $\ell_i^{(j)} = \ell_{i_1}^{(j)} \otimes \cdots \otimes \ell_{i_k}^{(j)}$. Yates's algorithm gives a circuit for evaluating the polynomial $T^{\otimes k}$ as follows. Let us write ε for the empty tuple. For $j \in [d]$, $q \in \{0, 1, \ldots, k\}$, $i \in [r]^q$, and $h \in [n]^{k-q}$, define the intermediate polynomial

$$L_{h,i}^{(j)}(z^{(j)}) = \sum_{p \in [n]^q} \ell_{i_1 p_1}^{(j)} \cdots \ell_{i_q p_q}^{(j)} z_{h,p}^{(j)}.$$

We observe in particular that for $q = 0$ we have $L^{(j)}_{h,\varepsilon}(z^{(j)}) = z^{(j)}_h$ for $h \in [n]^k$. These indeterminates $z^{(j)}_h$ for all $j \in [d]$ and $h \in [n]^k$ are the dn^k inputs to the circuit. We construct the circuit one stage $q \in \{0, 1, \ldots, k\}$ at a time, ensuring by induction that stage q has gates that evaluate to the polynomial $L^{(j)}_{h,i}(z^{(j)})$ for each $h \in [n]^{k-q}$ and $i \in [r]^q$. The input gates at stage $q = 0$ form the base case. Assume stage $q - 1$ has been constructed and construct stage q using the recurrence

$$L^{(j)}_{h,it}(z^{(j)}) = \sum_{p=1}^{n} \ell^{(j)}_{tp} L^{(j)}_{hp,i}(z^{(j)}) \tag{24}$$

for all $t \in [r]$, $i \in [r]^{q-1}$, and $h \in [n]^{k-q+1}$. From (24) we immediately observe that stage $q \in [k]$ can be constructed using at most $O(n^{k-q+1} \cdot r^q) = O(n^{k+1} \cdot (r/n)^q)$ wires. Furthermore, stage $q = k$ has gates that evaluate to $L^{(j)}_{\varepsilon,i}(z^{(j)}) = \ell^{(j)}_i(z^{(j)})$ for all $i \in [r]^k$. Complete the circuit for $T^{\otimes k}$ using (23) and wiring in from gates at stage k; at most $O(dr^k)$ wires suffice for this completion. We thus have a circuit of size at most $O(dkr^{k+1})$ for $T^{\otimes k}$, as desired for (7), immediately implying also Theorem 8.

C Improved Asymptotic Rank of the 4-Clique

The purpose of this appendix is to present an improved upper bound on $\tau(K_4)$.

Theorem 48. *For the 4-clique, we have $\tau(K_4) < 0.772318$.*

This improves upon the previous bound $\tau(K_4) < 0.772943$ due to Christandl, Vrana, and Zuiddam [13]. The improvement arises from a refinement of the analysis of Coppersmith–Winograd tensors (CW tensors) [16] in the k-mode setting. While [13] extends the classical 3-mode CW construction to k modes, their analysis is restricted to the *small* CW tensors (corresponding to the *simple construction* in [16]). Here we incorporate the *big* CW tensors (the *complicated construction*), which leads to a slightly improved bound.

Throughout this appendix we follow the terminology and general framework of [13]. Our notation differs slightly: our small CW tensor cw^k_q corresponds to the CW tensor analyzed in [13].

C.1 CW tensors in k modes

Definition 49 (Small and big CW k-tensors)**.** Let $k, q \geq 2$ be integers. For each mode $u \in [k]$, let $A^{(u)} \cong \mathbb{F}^{q+2}$ be equipped with a basis

$$x^{(u)}_0, x^{(u)}_1, \ldots, x^{(u)}_q, x^{(u)}_{q+1}.$$

The *small* CW k-tensor cw^k_q is the element of $\bigotimes_{u=1}^k \langle x^{(u)}_0, \ldots, x^{(u)}_q \rangle \subseteq \bigotimes_{u=1}^k A^{(u)}$ defined by

$$\mathrm{cw}^k_q \;:=\; \sum_{1 \leq u < v \leq k} \left(\sum_{i=1}^{q} x^{(u)}_i x^{(v)}_i \right) x^{(1)}_0 \cdots \overline{x^{(u)}_0} \cdots \overline{x^{(v)}_0} \cdots x^{(k)}_0,$$

where the overlines indicate that the corresponding factors are omitted.

The *big* CW k-tensor is

$$\mathrm{CW}^k_q \;:=\; \mathrm{cw}^k_q \;+\; \sum_{u=1}^{k} x^{(u)}_{q+1} \cdot x^{(1)}_0 \cdots \overline{x^{(u)}_0} \cdots x^{(k)}_0 \;\in\; \bigotimes_{u=1}^{k} A^{(u)}.$$

Block structure and outer support. We partition the basis in each mode $u \in [k]$ into three blocks

$$b_0^{(u)} = \{x_0^{(u)}\}, \qquad b_1^{(u)} = \{x_1^{(u)}, \dots, x_q^{(u)}\}, \qquad b_2^{(u)} = \{x_{q+1}^{(u)}\}.$$

This induces a block decomposition of CW_q^k indexed by the alphabet $B = \{0, 1, 2\}$. The set of block indices of nonzero blocks (the *outer support*) is

$$\Phi \;=\; \big\{(i_1, \dots, i_k) \in B^k : i_1 + \cdots + i_k = 2\big\}.$$

There are two types of nonzero blocks:

- $(1+1$ *blocks*$)$ If $i_u = i_v = 1$ for some $u < v$ and $i_w = 0$ for $w \notin \{u, v\}$, then the block is

$$T_{\{u,v\},q} \;=\; \left(\sum_{i=1}^{q} x_i^{(u)} x_i^{(v)}\right) x_0^{(1)} \cdots \overline{x_0^{(u)}} \cdots \overline{x_0^{(v)}} \cdots x_0^{(k)}.$$

- $(2$ *blocks*$)$ If $i_u = 2$ for some u and $i_w = 0$ for $w \neq u$, then the block is rank-one and (up to relabelling) is the graph tensor of the empty graph, denoted $T_{\varnothing,q}$.

Lemma 50. *The border rank of* CW_q^k *is at most* $q + 2$.

Proof. The following degeneration is a direct k-mode generalization of the original Coppersmith–Winograd construction [16]:

$$\begin{aligned}
&\sum_{i=1}^{q} \varepsilon(x_0^{(1)} + \varepsilon^2 x_i^{(1)}) \cdots (x_0^{(k)} + \varepsilon^2 x_i^{(k)}) \\
&\quad - \left(x_0^{(1)} + \varepsilon^3 \sum_{i=1}^{q} x_i^{(1)}\right) \cdots \left(x_0^{(k)} + \varepsilon^3 \sum_{i=1}^{q} x_i^{(k)}\right) \\
&\quad + (1 - q\varepsilon)(x_0^{(1)} + \varepsilon^5 x_{q+1}^{(1)}) \cdots (x_0^{(k)} + \varepsilon^5 x_{q+1}^{(k)}) \\
&= \varepsilon^5 \mathrm{CW}_q^k + O(\varepsilon^6).
\end{aligned}$$

The left-hand side is a sum of $q + 2$ rank-one tensors, hence $\underline{\mathrm{R}}(\mathrm{CW}_q^k) \le q + 2$. □

C.2 Laser Method on CW k-tensors

We work with the block decomposition of CW_q^k induced by the alphabet $B = \{0, 1, 2\}$. This decomposition lifts to tensor powers as follows. For each mode $u \in [k]$ and each word $w = (w_1, \dots, w_n) \in B^n$, define the block

$$b_w^{(u)} \;:=\; b_{w_1}^{(u)} \otimes \cdots \otimes b_{w_n}^{(u)} \;\subseteq\; (A^{(u)})^{\otimes n}.$$

A block subtensor of $(\mathrm{CW}_q^k)^{\otimes n}$ is indexed by a k-tuple of words $(w^{(1)}, \dots, w^{(k)}) \in (B^n)^k$ and equals

$$T_{w^{(1)},\dots,w^{(k)}} \;=\; T_{w_1^{(1)},\dots,w_1^{(k)}} \;\otimes\; \cdots \;\otimes\; T_{w_n^{(1)},\dots,w_n^{(k)}}.$$

Whenever this block is nonzero, each factor $T_{w_p^{(1)},\dots,w_p^{(k)}}$ is (up to isomorphism) a graph tensor. Consequently $T_{w^{(1)},\dots,w^{(k)}}$ is isomorphic to the graph tensor of the multigraph obtained by the multiset union of the one-edge factors.

Type classes and principal subtensors. A standard step in the Laser Method is to restrict to a principal subtensor in which the relevant edge count is fixed. Let $\alpha, \beta, \gamma \in \mathbb{Q}_{\geq 0}$ with $\alpha+\beta+\gamma = 1$, and assume n is chosen such that $\alpha n, \beta n, \gamma n$ are integers. Let

$$W_{\alpha,\beta,\gamma} \;:=\; \left\{\, w \in B^n : \#\{p : w_p = 0\} = \alpha n,\ \#\{p : w_p = 1\} = \beta n,\ \#\{p : w_p = 2\} = \gamma n \,\right\}$$

be the corresponding type class. We restrict $(\mathrm{CW}_q^k)^{\otimes n}$ to the subtensor obtained by keeping, in each mode $u \in [k]$, only the blocks indexed by $W_{\alpha,\beta,\gamma}$.

Now consider a nonzero block $T_{w^{(1)},\ldots,w^{(k)}}$ with $w^{(u)} \in W_{\alpha,\beta,\gamma}$. At each position $p \in [n]$, the factor $T_{w_p^{(1)},\ldots,w_p^{(k)}}$ is either a one-edge tensor or an empty-graph tensor. Let m be the number of one-edge positions. Counting symbols across all k modes yields:

- each one-edge factor contributes two symbols "1" (and an empty factor contributes none), hence $2m = k\beta n$;
- each empty factor contributes one symbol "2" (and a one-edge factor contributes none), hence $n - m = k\gamma n$.

Therefore

$$\frac{1}{2}\,\beta + \gamma = \frac{1}{k}, \qquad \text{so} \qquad \beta = \frac{2}{k} - 2\gamma, \qquad \alpha = 1 - \frac{2}{k} + \gamma.$$

In particular, feasibility forces $0 \leq \gamma \leq 1/k$, and the associated multigraph has exactly

$$m \;=\; (1 - k\gamma)\, n$$

edges. We abbreviate $W_{\alpha,\beta,\gamma}$ by W_γ, and write $(\mathrm{CW}_q^k)^{\otimes n}[\gamma]$ for the resulting principal subtensor.

Principal subtensors with prescribed marginals. For later use, we record a convenient general definition. (In our application, $B = \{0,1,2\}$ and the marginals are all equal to (α, β, γ).)

Definition 51 (Principal subtensor of prescribed marginals)**.** Let T be a k-tensor written in a fixed basis B in each mode. Let $P_1, \ldots, P_k$ be probability distributions on B with rational entries. For n such that $nP_u(b) \in \mathbb{Z}$ for all $u \in [k]$ and $b \in B$, let $W_{P_u} \subseteq B^n$ be the type class of words whose empirical distribution equals P_u. We define the *principal subtensor* $T^{\otimes n}[P_1, \ldots, P_k]$ as the block subtensor of $T^{\otimes n}$ obtained by restricting, in mode u, to the direct sum of blocks indexed by W_{P_u} and zeroing out all other blocks.

A fixed-marginal variant of [13, Theorem 1.52]. We next isolate the form of the Christandl–Vrana–Zuiddam bound that we use. It is already implicit in the proof of [13, Theorem 1.52]: once the marginals are fixed, the proof constructs a large diagonal tensor inside the corresponding principal subtensor, and the only additional maximization in the theorem statement is over the choice of marginals.

Theorem 52 (Fixed-marginal Laser-Method bound)**.** *Let T be a tight k-tensor written in a fixed basis B, with support $\Phi \subseteq B^k$. Fix marginal distributions $P_1, \ldots, P_k$ on B such that there exists a distribution on Φ with these marginals, and let $P^\star$ be a maximum-entropy distribution on Φ with marginals $P_1, \ldots, P_k$. Write $\mathrm{H}(\cdot)$ for Shannon entropy in bits.*

Define the diagonal set $\Delta_\Phi := \{(x,x) : x \in \Phi\}$ and, for each $i \in [k]$,

$$R_i \;:=\; \{(x,y) \in \Phi \times \Phi : x_i = y_i\}.$$

Let

$$\mathscr{R}_\Phi \;:=\; \{\, R \subseteq \Phi \times \Phi : R \not\subseteq \Delta_\Phi \;\; \textit{and} \;\; R \subseteq R_i \textit{ for some } i \in [k] \,\}.$$

For $R \in \mathscr{R}_\Phi$, let $\mathscr{Q}_{\Phi,(P_1,\dots,P_k),R}$ be the set of all probability distributions Q on R such that the $2k$ coordinate-marginals satisfy

$$Q_i = Q_{k+i} = P_i \qquad (i \in [k]),$$

where $(Q_1, \dots, Q_k)$ are the marginals on the first component $x \in \Phi$ and $(Q_{k+1}, \dots, Q_{2k})$ those on the second component $y \in \Phi$. Finally, let $\alpha_1, \dots, \alpha_k : B \to \mathbb{Z}$ be tightness maps for T (unrelated to the scalar parameter α above), and define $r_\alpha(R)$ to be the rank over $\mathbb{Q}$ of the matrix with rows $\alpha(x) - \alpha(y)$ for $(x, y) \in R$, where $\alpha(x) := (\alpha_1(x_1), \dots, \alpha_k(x_k))$.

Set

$$\mu(P_1, \dots, P_k) \;:=\; \mathrm{H}(P^\star) \;-\; (k-2) \cdot \max_{R \in \mathscr{R}_\Phi} \; \max_{Q \in \mathscr{Q}_{\Phi,(P_1,\dots,P_k),R}} \frac{\mathrm{H}(Q) - \mathrm{H}(P^\star)}{r_\alpha(R)}.$$

Then for every sufficiently large n, the principal subtensor $T^{\otimes n}[P_1, \dots, P_k]$ can be zeroed out to a diagonal tensor of size

$$2^{\,(\mu(P_1,\dots,P_k) - o(1))\, n}.$$

Remark 53. For computations it is convenient to rewrite the expression for μ as a *minimum* over R. For each $R \in \mathscr{R}_\Phi$, let $Q_R^\star \in \arg\max_{Q \in \mathscr{Q}_{\Phi,(P_1,\dots,P_k),R}} \mathrm{H}(Q)$ and define

$$F_R(P_1, \dots, P_k) \;:=\; \mathrm{H}(P^\star) \;-\; (k-2) \cdot \frac{\mathrm{H}(Q_R^\star) - \mathrm{H}(P^\star)}{r_\alpha(R)}.$$

Then $\mu(P_1, \dots, P_k) = \min_{R \in \mathscr{R}_\Phi} F_R(P_1, \dots, P_k)$.

C.3 Analysis in 4 modes

In this subsection we specialize to $k = 4$. With respect to $B = \{0, 1, 2\}$, the outer support of CW_q^4 is

$$\Phi \;=\; \big\{(i_1, i_2, i_3, i_4) \in \{0,1,2\}^4 : i_1 + i_2 + i_3 + i_4 = 2\big\}.$$

We fix identical marginals

$$P_1 = P_2 = P_3 = P_4 = (\alpha, \beta, \gamma), \qquad \alpha = \tfrac{1}{2} + \gamma, \qquad \beta = \tfrac{1}{2} - 2\gamma, \qquad \gamma \in \Big[0, \tfrac{1}{4}\Big].$$

We write $\mu(\gamma)$ for the value of $\mu(P_1, \dots, P_4)$ in Theorem 52. Our goal is to show that

$$\mu(\gamma) = \mathrm{H}(P_1) = \mathrm{H}(\alpha, \beta, \gamma),$$

i.e. the rank-one candidates R do not decrease the minimum in Remark 53 below the value attained by R_1.

Reducing the enumeration of R. Since the marginals are identical and Φ is invariant under permuting the four modes, the quantity $F_R(P_1, \dots, P_4)$ is unchanged under simultaneous permutation of coordinates in $(x, y) \in \Phi \times \Phi$. Therefore, when enumerating $R \in \mathscr{R}_\Phi$, we may assume

$$R \subseteq R_1 := \{(x, y) \in \Phi \times \Phi : x_1 = y_1\}.$$

Moreover, by [13, Lemma 3.18] it suffices to consider sets R that are maximal (w.r.t. inclusion) among those with fixed rank $r_\alpha(R)$. Since for $k = 4$ one always has $1 \le r_\alpha(R) \le k - 2 = 2$, this yields the following reduction:

- for $r_\alpha(R) = 2$, it suffices to take the unique maximal set $R = R_1$;
- for $r_\alpha(R) = 1$, it suffices to take maximal rank-one relations inside R_1.

In the rank-one case, all difference vectors $\alpha(x) - \alpha(y)$ with $(x, y) \in R$ must lie on a common line. Up to permuting modes, this line is generated by one of the two types of difference vectors

$$(0, 1, 1, -2) = (0, 1, 1, 0) - (0, 0, 0, 2), \qquad \text{or} \qquad (0, 1, -1, 0) = (1, 1, 0, 0) - (1, 0, 1, 0),$$

(the second type also occurs as $\frac{1}{2}\big((0, 2, 0, 0) - (0, 0, 2, 0)\big)$). Consequently, it suffices to consider the following three representatives, written as equivalence relations on Φ (we list only the nontrivial classes; all remaining points are singleton classes):

$$R_{(1)} := R_1, \qquad R_{(2)} : \ (0, 1, 1, 0) \sim (0, 0, 0, 2),$$

and

$$R_{(3)} : \ (0, 0, 1, 1) \sim (0, 1, 0, 1), \quad (1, 1, 0, 0) \sim (1, 0, 1, 0), \quad (0, 2, 0, 0) \sim (0, 0, 2, 0).$$

In particular $r_\alpha(R_{(1)}) = 2$ and $r_\alpha(R_{(2)}) = r_\alpha(R_{(3)}) = 1$. The proofs that $F_{R_{(2)}}(\gamma) \geq \mathrm{H}(P_1)$ and $F_{R_{(3)}}(\gamma) \geq \mathrm{H}(P_1)$ (for all $\gamma \in [0, 1/4]$) will be given in the subsequent lemmas; together with the explicit evaluation $F_{R_{(1)}}(\gamma) = \mathrm{H}(P_1)$, this implies $\mu(\gamma) = \mathrm{H}(P_1)$.

Analysis of $R_{(1)}$. Recall $R_{(1)} := R_1 = \{(x, y) \in \Phi \times \Phi : \ x_1 = y_1\}$ and $r_\alpha(R_{(1)}) = 2$. We first show that the rank-two candidate $R_{(1)}$ yields the baseline value $F_{R_{(1)}}(\gamma) = \mathrm{H}(P_1)$.

Lemma 54. *Let $k = 4$ and fix identical marginals $P_1 = \cdots = P_4 = (\alpha, \beta, \gamma)$. Let $P^\star$ be the maximum-entropy distribution on Φ with these marginals. Then for $R_{(1)} = R_1$ we have*

$$F_{R_{(1)}}(\gamma) = \mathrm{H}(P_1) = \mathrm{H}(\alpha, \beta, \gamma).$$

Proof. Let $Q \in \mathscr{Q}_{\Phi,(P_1,\ldots,P_4),R_{(1)}}$ and let $(X, Y) \sim Q$. By definition of $R_{(1)}$ we have $X_1 = Y_1$ almost surely; write $Z := X_1 = Y_1$. Then, by the chain rule and subadditivity,

$$\mathrm{H}(X, Y) = \mathrm{H}(Z) + \mathrm{H}(X_{2,3,4}, Y_{2,3,4} \mid Z) \leq \mathrm{H}(Z) + \mathrm{H}(X_{2,3,4} \mid Z) + \mathrm{H}(Y_{2,3,4} \mid Z).$$

Moreover, $\mathrm{H}(Z) = \mathrm{H}(P_1)$ since $Z = X_1$ has marginal P_1. Using again the chain rule,

$$\mathrm{H}(X) = \mathrm{H}(Z) + \mathrm{H}(X_{2,3,4} \mid Z), \qquad \mathrm{H}(Y) = \mathrm{H}(Z) + \mathrm{H}(Y_{2,3,4} \mid Z).$$

Since X (resp. Y) is supported on Φ and has marginals $P_1, \ldots, P_4$, maximality of $P^\star$ gives $\mathrm{H}(X) \leq \mathrm{H}(P^\star)$ and $\mathrm{H}(Y) \leq \mathrm{H}(P^\star)$. Hence

$$\mathrm{H}(X_{2,3,4} \mid Z) \leq \mathrm{H}(P^\star) - \mathrm{H}(P_1), \qquad \mathrm{H}(Y_{2,3,4} \mid Z) \leq \mathrm{H}(P^\star) - \mathrm{H}(P_1),$$

and therefore

$$\mathrm{H}(Q) = \mathrm{H}(X, Y) \leq 2\,\mathrm{H}(P^\star) - \mathrm{H}(P_1). \tag{25}$$

This upper bound is tight: sample $Z \sim P_1$, and then sample X and Y *independently conditioned on* Z, with $X \mid (Z = z) \sim P^\star(\,\cdot \mid X_1 = z)$ and $Y \mid (Z = z) \sim P^\star(\,\cdot \mid Y_1 = z)$. This produces a feasible Q supported on $R_{(1)}$ and with $\mathrm{H}(Q) = 2\,\mathrm{H}(P^\star) - \mathrm{H}(P_1)$, hence (25) is optimal and

$$\mathrm{H}(Q^\star_{R_{(1)}}) = 2\,\mathrm{H}(P^\star) - \mathrm{H}(P_1).$$

Finally, since $r_\alpha(R_{(1)}) = 2$ and $k - 2 = 2$,

$$F_{R_{(1)}}(\gamma) = \mathrm{H}(P^\star) - 2 \cdot \frac{\mathrm{H}(Q^\star_{R_{(1)}}) - \mathrm{H}(P^\star)}{2} = \mathrm{H}(P^\star) - \big(\mathrm{H}(Q^\star_{R_{(1)}}) - \mathrm{H}(P^\star)\big) = \mathrm{H}(P_1),$$

as claimed. □

Analysis of $R_{(2)}$. We next treat the rank-one relation $R_{(2)}$, whose only nontrivial class is

$$a := (0,1,1,0) \;\sim\; b := (0,0,0,2).$$

Lemma 55. *For all $\gamma \in [0,1/4]$ we have*

$$F_{R_{(2)}}(\gamma) \;\geq\; \mathrm{H}(P_1) = \mathrm{H}(\alpha,\beta,\gamma).$$

Proof. Let $Q \in \mathscr{Q}_{\Phi,(P_1,\dots,P_4),R_{(2)}}$ and $(X,Y) \sim Q$. Since $R_{(2)}$ is diagonal except on the class $\{a,b\}$, we have: if $X \notin \{a,b\}$ then $Y = X$, while if $X \in \{a,b\}$ then $Y \in \{a,b\}$. Therefore $H(Y \mid X = x) \leq 1$ if $x \in \{a,b\}$ and 0 otherwise, hence

$$\mathrm{H}(Y \mid X) \;\leq\; \Pr[X \in \{a,b\}].$$

By the chain rule,

$$\mathrm{H}(Q) = \mathrm{H}(X,Y) = \mathrm{H}(X) + \mathrm{H}(Y \mid X) \;\leq\; \mathrm{H}(X) + \Pr[X \in \{a,b\}].$$

The marginal constraints imply that X is supported on Φ and has coordinate marginals $P_1,\dots,P_4$, hence $\mathrm{H}(X) \leq \mathrm{H}(P^\star)$ by maximality of $P^\star$. Moreover, $X = a$ forces $X_2 = 1$ and $X = b$ forces $X_4 = 2$, so

$$\Pr[X = a] \leq \Pr[X_2 = 1] = \beta, \qquad \Pr[X = b] \leq \Pr[X_4 = 2] = \gamma,$$

and thus $\Pr[X \in \{a,b\}] \leq \beta + \gamma = \frac{1}{2} - \gamma$. Altogether,

$$\mathrm{H}(Q^\star_{R_{(2)}}) \;\leq\; \mathrm{H}(P^\star) + \Big(\tfrac{1}{2} - \gamma\Big). \tag{26}$$

Since $r_\alpha(R_{(2)}) = 1$ and $k - 2 = 2$, we have

$$F_{R_{(2)}}(\gamma) = \mathrm{H}(P^\star) - 2\big(\mathrm{H}(Q^\star_{R_{(2)}}) - \mathrm{H}(P^\star)\big) \;\geq\; \mathrm{H}(P^\star) - 2\Big(\tfrac{1}{2} - \gamma\Big) = \mathrm{H}(P^\star) - 1 + 2\gamma.$$

It remains to show $\mathrm{H}(P^\star) - 1 + 2\gamma \geq \mathrm{H}(P_1)$, i.e.

$$\mathrm{H}(P^\star) - \mathrm{H}(P_1) \;\geq\; 1 - 2\gamma. \tag{27}$$

Let $X^\star \sim P^\star$. Then $\mathrm{H}(P^\star) - \mathrm{H}(P_1) = \mathrm{H}(X^\star_{2,3,4} \mid X^\star_1)$. If $X^\star_1 = 1$, then exactly one of $X^\star_2, X^\star_3, X^\star_4$ equals 1, hence $\mathrm{H}(X^\star_{2,3,4} \mid X^\star_1 = 1) = \lg 3$. If $X^\star_1 = 0$, then $(X^\star_2, X^\star_3, X^\star_4)$ lies either in the three "2-type" outcomes or in the three "$1+1$-type" outcomes; these two sets are disjoint and each is uniform under $P^\star$ by symmetry, so $\mathrm{H}(X^\star_{2,3,4} \mid X^\star_1 = 0) \geq \lg 3$. Finally, if $X^\star_1 = 2$ the remaining coordinates are deterministic, so the conditional entropy is 0. Therefore

$$\mathrm{H}(P^\star) - \mathrm{H}(P_1) = \mathrm{H}(X^\star_{2,3,4} \mid X^\star_1) \;\geq\; \Pr[X^\star_1 \neq 2] \cdot \lg 3 = (1-\gamma)\lg 3.$$

Since $\lg 3 > 1$ and $2 - \lg 3 > 0$, we have

$$(1-\gamma)\lg 3 - (1 - 2\gamma) = (\lg 3 - 1) + \gamma(2 - \lg 3) \geq 0,$$

which implies (27) and completes the proof. □

Analysis of $R_{(3)}$. We finally treat the remaining rank-one representative, where (listing only the nontrivial classes)

$$R_{(3)}:\ (0,0,1,1)\sim(0,1,0,1),\quad (1,1,0,0)\sim(1,0,1,0),\quad (0,2,0,0)\sim(0,0,2,0).$$

Lemma 56. *For all $\gamma\in[0,1/4]$ we have*

$$F_{R_{(3)}}(\gamma)\ \geq\ \mathrm{H}(P_1)=\mathrm{H}(\alpha,\beta,\gamma).$$

Proof. Let $Q\in\mathscr{Q}_{\Phi,(P_1,\dots,P_4),R_{(3)}}$ and $(X,Y)\sim Q$. Let

$$S:=\{(0,0,1,1),(0,1,0,1),(1,1,0,0),(1,0,1,0),(0,2,0,0),(0,0,2,0)\}\subseteq\Phi$$

be the set of points lying in the nontrivial $R_{(3)}$-classes. If $X\notin S$ then $Y=X$ deterministically, while if $X\in S$ then Y has at most two possible values. Hence $\mathrm{H}(Y\mid X)\leq\Pr[X\in S]$ and

$$\mathrm{H}(Q)=\mathrm{H}(X,Y)=\mathrm{H}(X)+\mathrm{H}(Y\mid X)\ \leq\ \mathrm{H}(X)+\Pr[X\in S]. \tag{28}$$

We now write the law of X explicitly. Since each coordinate takes the value 2 with probability γ, and in Φ the event $\{X_u=2\}$ uniquely identifies the point $2e_u$, we have

$$\Pr[X=(2,0,0,0)]=\Pr[X=(0,2,0,0)]=\Pr[X=(0,0,2,0)]=\Pr[X=(0,0,0,2)]=\gamma.$$

For the remaining six points (of type $(1,1,0,0)$), define

$$x_{12}:=\Pr[X=(1,1,0,0)],\ x_{13}:=\Pr[X=(1,0,1,0)],\ x_{14}:=\Pr[X=(1,0,0,1)],$$

$$x_{23}:=\Pr[X=(0,1,1,0)],\ x_{24}:=\Pr[X=(0,1,0,1)],\ x_{34}:=\Pr[X=(0,0,1,1)].$$

The marginal constraint $\Pr[X_u=1]=\beta$ for $u=1,2,3,4$ yields

$$\begin{aligned}
x_{12}+x_{13}+x_{14}&=\beta,\\
x_{12}+x_{23}+x_{24}&=\beta,\\
x_{13}+x_{23}+x_{34}&=\beta,\\
x_{14}+x_{24}+x_{34}&=\beta,
\end{aligned}$$

from which it follows that $x_{12}=x_{34}$, $x_{13}=x_{24}$, and $x_{14}=x_{23}$. Thus there exist $x,y,z\geq0$ with $x+y+z=\beta$ such that

$$x_{12}=x_{34}=x,\qquad x_{13}=x_{24}=y,\qquad x_{14}=x_{23}=z.$$

In these variables,

$$\mathrm{H}(X)=-4\gamma\lg\gamma\ -\ 2x\lg x\ -\ 2y\lg y\ -\ 2z\lg z,$$

and, since S contains two of the "2-type" points and four of the "$1+1$-type" points,

$$\Pr[X\in S]=2\gamma+2(x+y).$$

Substituting into (28) gives the bound

$$\mathrm{H}(Q^{\star}_{R_{(3)}})\ \leq\ -4\gamma\lg\gamma\ -\ 2x\lg x\ -\ 2y\lg y\ -\ 2z\lg z\ +\ 2\gamma+2(x+y), \tag{29}$$

where $x, y, z \geq 0$ and $x + y + z = \beta$.

The right-hand side of (29) is concave in (x, y, z) over the simplex $x + y + z = \beta$. A short Lagrange-multiplier computation shows that it is maximized at

$$x = y = \frac{2\beta}{5}, \qquad z = \frac{\beta}{5},$$

which yields the explicit upper bound

$$U(\gamma) := -4\gamma \lg \gamma \;-\; \frac{8\beta}{5} \lg\Big(\frac{2\beta}{5}\Big) \;-\; \frac{2\beta}{5} \lg\Big(\frac{\beta}{5}\Big) \;+\; 2\gamma + \frac{8\beta}{5}, \qquad \beta = \tfrac{1}{2} - 2\gamma, \tag{30}$$

so that $\mathrm{H}(Q^\star_{R_{(3)}}) \leq U(\gamma)$.

Next note that all difference vectors $\alpha(x) - \alpha(y)$ for $(x, y) \in R_{(3)}$ are multiples of $(0, 1, -1, 0)$, hence $r_\alpha(R_{(3)}) = 1$. Therefore

$$F_{R_{(3)}}(\gamma) = \mathrm{H}(P^\star) - 2\big(\mathrm{H}(Q^\star_{R_{(3)}}) - \mathrm{H}(P^\star)\big) = 3\,\mathrm{H}(P^\star) - 2\,\mathrm{H}(Q^\star_{R_{(3)}}) \;\geq\; 3\,\mathrm{H}(P^\star) - 2U(\gamma).$$

It remains to show that $3\,\mathrm{H}(P^\star) - 2U(\gamma) \geq \mathrm{H}(P_1)$ for $\gamma \in [0, 1/4]$.

Using $\beta = \frac{1}{2} - 2\gamma$ and the explicit form of $P^\star$ on Φ, we have

$$\mathrm{H}(P^\star) = -4\gamma \lg \gamma \;-\; 2\beta \lg\Big(\frac{\beta}{3}\Big).$$

Define

$$D(\gamma) := \big(3\,\mathrm{H}(P^\star) - 2U(\gamma)\big) - \mathrm{H}(\alpha, \beta, \gamma), \qquad \alpha = \tfrac{1}{2} + \gamma,\ \beta = \tfrac{1}{2} - 2\gamma.$$

A direct differentiation yields, for $\gamma \in (0, 1/4)$,

$$D''(\gamma) = \frac{3}{\gamma(2\gamma + 1)(4\gamma - 1)\ln 2} \;<\; 0,$$

so D is concave on $[0, 1/4]$ and hence attains its minimum at an endpoint. At $\gamma = 0$ one has $\mathrm{H}(P^\star) = \lg 6$, $U(0) = \lg 10$, and $\mathrm{H}(\alpha, \beta, 0) = 1$, giving

$$D(0) = 3 \lg 6 - 2 \lg 10 - 1 = \lg\Big(\frac{27}{25}\Big) > 0.$$

At $\gamma = \frac{1}{4}$ one has $\beta = 0$, $\mathrm{H}(P^\star) = 2$, $U(\frac{1}{4}) = \frac{5}{2}$, and $\mathrm{H}(\frac{3}{4}, 0, \frac{1}{4}) = -\frac{3}{4}\lg\frac{3}{4} - \frac{1}{4}\lg\frac{1}{4}$, giving

$$D\Big(\tfrac{1}{4}\Big) = 1 - \mathrm{H}\Big(\tfrac{3}{4}, 0, \tfrac{1}{4}\Big) = -1 + \frac{3}{4}\lg 3 \;> 0.$$

Thus $D(\gamma) \geq 0$ on $[0, 1/4]$, which implies $F_{R_{(3)}}(\gamma) \geq \mathrm{H}(P_1)$. □

Coda. Combining Lemmas 54 to 56 with $\mu(\gamma) = \min_{R \in \mathscr{R}_\Phi} F_R(\gamma)$ (cf. Remark 53) yields:

Corollary 57. *For all $\gamma \in [0, 1/4]$,*

$$\mu(\gamma) = \mathrm{H}(\alpha, \beta, \gamma), \qquad \alpha = \tfrac{1}{2} + \gamma, \qquad \beta = \tfrac{1}{2} - 2\gamma.$$

We now explain how the fixed-marginal laser bound (Theorem 52) specializes to an explicit upper bound on $\tau(K_4)$. Recall that for $k = 4$ we write

$$P_1 = P_2 = P_3 = P_4 = (\alpha, \beta, \gamma), \qquad \alpha = \tfrac{1}{2} + \gamma, \qquad \beta = \tfrac{1}{2} - 2\gamma, \qquad 0 \le \gamma \le \tfrac{1}{4},$$

and that the associated principal subtensor $(\mathrm{CW}_q^4)^{\otimes n}[\gamma]$ corresponds to a multigraph on K_4 with

$$m = (1 - 4\gamma)n$$

edge-factors (and $n - m = 4\gamma n$ empty factors).

Proposition 58. *Let $q \ge 2$ and let $0 \le \gamma < \frac{1}{4}$ be rational. Then*

$$\tau(K_4) \;\le\; \frac{1}{1-4\gamma}\,\log_q\!\left(\frac{q+2}{2^{\mu(\gamma)}}\right), \qquad \mu(\gamma) = H\big(\tfrac{1}{2}+\gamma,\ \tfrac{1}{2}-2\gamma,\ \gamma\big)\,. \tag{31}$$

In particular, taking $\gamma = 0$ recovers the Christandl–Vrana–Zuiddam bound

$$\tau(K_4) \le \min_{q \ge 2} \log_q\!\left(\frac{q+2}{2}\right) = \log_7(\tfrac{9}{2}).$$

Proof. Fix $q \ge 2$ and $\gamma \in [0, \frac{1}{4})$ rational, and take n such that $\alpha n, \beta n, \gamma n$ are integers. Equip CW_q^4 with the 3-block partition described in Definition 49 (blocks indexed by $B = \{0,1,2\}$), so that the outer support is

$$\Phi = \{(i_1, i_2, i_3, i_4) \in \{0,1,2\}^4 : i_1 + i_2 + i_3 + i_4 = 2\}.$$

Let φ denote the corresponding outer-structure tensor on alphabet B.

By Corollary 57, the principal outer subtensor $\varphi^{\otimes n}[P_1, P_2, P_3, P_4]$ can be zeroed out to a diagonal 4-tensor of size

$$p \;=\; 2^{(\mu(\gamma) - o(1))n}, \qquad \mu(\gamma) = H(P_1) = H(\alpha, \beta, \gamma).$$

It lifts to a zeroing-out on the full tensor, yielding a direct sum

$$\phi_1 \oplus \cdots \oplus \phi_p \;\le\; (\mathrm{CW}_q^4)^{\otimes n}$$

where each ϕ_i is an inner-structure tensor obtained as a tensor product of n block tensors of CW_q^4. For our fixed principal subtensor $[\gamma]$, each ϕ_i is isomorphic to a graph tensor $T_{G_i,q}$ with $m = (1-4\gamma)n$ edges in the (multi)graph G_i.

On the other hand, Lemma 50 gives $\underline{\mathrm{R}}(\mathrm{CW}_q^4) \le q + 2$, and therefore

$$\underline{\mathrm{R}}\,(\phi_1 \oplus \cdots \oplus \phi_p) \;\le\; \underline{\mathrm{R}}\,\big((\mathrm{CW}_q^4)^{\otimes n}\big) \;\le\; (q+2)^n.$$

Applying the generalized asymptotic sum inequality [13, Theorem 2.1.6] to K_4 yields

$$\tau(K_4) \;\le\; \log_{q^m}\!\left(\frac{(q+2)^n}{p}\right) \;=\; \frac{n}{m}\,\log_q\!\left(\frac{q+2}{2^{\mu(\gamma)-o(1)}}\right).$$

Since $m = (1-4\gamma)n$, letting $n \to \infty$ gives (31). □

Proof of Theorem 48. By Proposition 58, for any $q \ge 2$ and any rational $\gamma \in [0, \frac{1}{4})$ we have

$$\tau(K_4) \;\le\; \frac{1}{1-4\gamma}\,\log_q\!\left(\frac{q+2}{2^{\mathrm{H}(\frac{1}{2}+\gamma,\ \frac{1}{2}-2\gamma,\ \gamma)}}\right).$$

Substituting $q = 7$ and $\gamma = 0.0012105179$ into (31) gives

$$\tau(K_4) \;\approx\; 0.77231702 \;<\; 0.772318.$$

This proves $\tau(K_4) < 0.772318$. □

ECCC ISSN 1433-8092

https://eccc.weizmann.ac.il