The Entropy Law for the Asymptotic Subrank of Dicke Tensors

Wenqi Zhu* Chunlin Liu† Kai Zou‡ Ping Nie§

November 2026

Abstract

Let $\mathbf{d} = (d_1, \dots, d_n)$ be any tuple of positive integers, where $n \geq 2$, and put $k = \sum_a d_a$ and $p_a = d_a/k$. The Dicke tensor $D_{\mathbf{d}} \in (\mathbb{C}^n)^{\otimes k}$ is the sum of all basis tensors with these multiplicities. We prove that its ordinary asymptotic subrank and its asymptotic subrank under coordinate restrictions both equal $\exp \mathrm{H}(p)$, where H is Shannon entropy with natural logarithms. All k tensor factors remain separate. This proves the all-multiplicity entropy formula for Dicke tensors. The lower bound combines rank-dependent hashing and global isolation with an entropy inequality for finite distributions invariant under an irreducible linear action. A full joint-type class supplies the candidates, while every edge of the original support is included in the collision count. A slice cover gives the finite-power upper bound $Q(D_{\mathbf{d}}^{\boxtimes N}) \leq k \exp(N\,\mathrm{H}(p))$ for every $N \geq 1$, including arbitrary complex-linear restriction maps. The proof includes the average-free construction, the isolation argument, and the passage from an unbounded prime-indexed subsequence to the asymptotic limit. Finite exact checks accompany the written proof.

1 The main theorem

Fix integers $n \geq 2$ and $d_1, \dots, d_n \geq 1$, and set

$$k = \sum_{a=1}^{n} d_a, \qquad p_a = \frac{d_a}{k}, \qquad \mathrm{H}(p) = -\sum_{a=1}^{n} p_a \log p_a.$$

All logarithms are natural unless a base is indicated; in particular, $\mathrm{H}_2(p) = \mathrm{H}(p)/\log 2$. Write $e_1, \dots, e_n$ for the standard basis of $\mathbb{C}^n$. Let $\mathcal{W}_{\mathbf{d}}$ be the set of words $(a_1, \dots, a_k) \in \{1, \dots, n\}^k$ in which symbol a occurs exactly d_a times. The *Dicke tensor* is

$$D_{\mathbf{d}} = \sum_{(a_1,\dots,a_k) \in \mathcal{W}_{\mathbf{d}}} e_{a_1} \otimes \cdots \otimes e_{a_k}. \tag{1.1}$$

Under the usual symmetric-tensor interpretation it corresponds to a nonzero scalar multiple of the monomial $x_1^{d_1} \cdots x_n^{d_n}$. This normalization does not affect the quantities below.

For a nonzero k-tensor $T \in V_1 \otimes \cdots \otimes V_k$, its *subrank* $Q(T)$ is the largest integer q for which linear maps $A_i : V_i \to \mathbb{C}^q$, chosen independently for the k factors, satisfy

$$(A_1 \otimes \cdots \otimes A_k)T = \Delta_q := \sum_{j=1}^{q} e_j \otimes \cdots \otimes e_j.$$

Relative to specified bases of the V_i, write $Q_{\mathrm{M}}(T)$ for the largest such diagonal obtainable by deleting local basis vectors and then independently relabelling and rescaling the retained basis vectors. This is the *coordinate, or monomial, restriction* convention. In particular, $Q_{\mathrm{M}}(T) \leq Q(T)$. For the all-one tensor (1.1), it amounts to finding an induced matching in its support.

The grouped power $T^{\boxtimes N}$ is the ordinary Nth tensor power with corresponding factors collected together:

$$T^{\boxtimes N} \in V_1^{\otimes N} \otimes \cdots \otimes V_k^{\otimes N}.$$

*Corresponding author. Mathematical Institute, University of Oxford, Andrew Wiles Building, Woodstock Road, Oxford OX2 6GG, United Kingdom. wenqi.zhu@maths.ox.ac.uk

†School of Mathematical Sciences, University of Science and Technology of China, Hefei 230026, Anhui, China.

‡Mathematical & Engineering Sciences, King's College London, UK. kz@netmind.ai

§David R. Cheriton School of Computer Science, University of Waterloo, 200 University Avenue West, Waterloo, ON N2L 3G1, Canada. ping.nie@uwaterloo.ca

There are still exactly k factors. Define

$$\widetilde{Q}(T) = \lim_{N\to\infty} Q(T^{\boxtimes N})^{1/N}, \qquad \widetilde{Q}_{\mathrm{M}}(T) = \lim_{N\to\infty} Q_{\mathrm{M}}(T^{\boxtimes N})^{1/N}$$

when these limits exist. Their existence for $D_{\mathbf{d}}$ is proved in Lemma 3.1.

Theorem 1.1. *For every $n \geq 2$ and every tuple $\mathbf{d} \in \mathbb{Z}^n_{>0}$, over $\mathbb{C}$ and with all $k = \sum_a d_a$ original factors separate,*

$$\widetilde{Q}_{\mathrm{M}}(D_{\mathbf{d}}) = \widetilde{Q}(D_{\mathbf{d}}) = \exp \mathrm{H}(\mathbf{d}/k) = 2^{\mathrm{H}_2(\mathbf{d}/k)}. \tag{1.2}$$

Moreover, for every integer $N \geq 1$,

$$Q(D_{\mathbf{d}}^{\boxtimes N}) \leq k \exp\bigl(N \mathrm{H}(\mathbf{d}/k)\bigr). \tag{1.3}$$

The ordinary restriction maps in the theorem need not be symmetric or monomial. The assertion is pointwise in each fixed multiplicity tuple; it supplies neither a convergence rate uniform in $\mathbf{d}$ nor an efficiency bound for finding the restrictions. It makes no claim that finite subrank equals the entropy expression, and concerns neither tensor rank nor border rank. Merging different original modes is not part of the statement or proof.

For example, when all d_a are equal, the entropy is $\log n$, so both asymptotic subranks equal n. For $\mathbf{d} = (k-1, 1)$ the value is

$$\exp \mathrm{H}\bigl((k-1)/k, 1/k\bigr) = \frac{k}{(k-1)^{(k-1)/k}}.$$

These are specializations of the same formula, with the same separation of the original tensor factors.

2 Relation to earlier work

The entropy formula for Dicke tensors was proposed in the study of asymptotic multipartite entanglement conversion. Vrana and Christandl [4, Section VI, equation (40)] suggested that the asymptotic number of copies of a Dicke tensor needed per two-level GHZ tensor is $1/\mathrm{H}_2(p)$, where $p = \mathbf{d}/k$ and H_2 denotes entropy with base-two logarithms. Equivalently, the logarithm of its asymptotic subrank should be $\mathrm{H}_2(p)$. Oneto and Ventura explicitly restated the ordinary subrank formula as Conjecture 4.26 in [2]. Their symmetric-tensor convention for a monomial differs from our Dicke tensor by a nonzero scalar, which does not affect subrank. Permuting basis symbols also identifies the partition convention used in some sources with arbitrary positive multiplicity tuples. Throughout this paper, all k tensor factors remain separate; tensor powers group only corresponding factors.

A distinction between restriction and degeneration is needed to compare our conclusion with Question 1.58 of Christandl, Vrana and Zuiddam [1]. Our Q_{M} uses coordinate deletions followed by independent relabellings and nonzero rescalings in the individual factors. These are monomial restrictions in their Definition 1.6. Their monomial subexponent q_{M}, however, is defined using monomial *degeneration*: see Definition 1.7, equation (1.9), and Definition 1.18. Correspondingly, Proposition 1.19 expresses this subexponent using monomial *border* subrank. Remark 1.12 explicitly distinguishes the resulting asymptotic rate from the rate defined using monomial restrictions; the two are not identified there in general. Writing superscripts to distinguish their notation, our theorem gives

$$\begin{aligned} \mathrm{H}_2(p) &= \log_2 \widetilde{Q}_{\mathrm{M}}(D_{\mathbf{d}}) \\ &\leq q_{\mathrm{M}}^{\mathrm{CVZ}}(D_{\mathbf{d}}) \leq q^{\mathrm{CVZ}}(D_{\mathbf{d}}) \\ &= \log_2 \widetilde{Q}(D_{\mathbf{d}}) = \mathrm{H}_2(p). \end{aligned} \tag{2.1}$$

The first comparison uses inclusion of monomial restriction in monomial degeneration, and the second inclusion in ordinary degeneration. The ordinary restriction–degeneration equivalence of [1, Theorem 1.11] justifies the last comparison with ordinary asymptotic subrank. Thus the theorem also answers their Question 1.58, without assuming any general equivalence between monomial restriction and monomial degeneration. This external equivalence is needed only for this comparison of conventions, not for our proof of the ordinary or coordinate-restriction formula.

We use the higher-order Coppersmith–Winograd construction of Christandl, Vrana and Zuiddam [1, Theorem 1.52 and Section 3.3]. The survival probability of two colliding edges is controlled by the rational rank of their coordinate differences. We use this rank-dependent hashing mechanism and give the needed probability

calculation directly. The accompanying average-free sets also have a standard antecedent: Vrana and Christandl prove the fixed-length average-free estimate in [4, Lemma 10 and Appendix A], quoted as Proposition 3.2 in [1]. Our proof includes a no-carry digit construction on a common Euclidean sphere. We make no novelty claim for the hashing or average-free mechanisms.

Counting a specified class of candidates while testing isolation against the entire surviving support likewise has a published precedent. Vrana and Christandl's weighted isolation lemma [5, Lemma 2.3, arXiv version 2] permits an arbitrary nonnegative weight on the full supported set and bounds its sum over globally isolated vertices of the collision graph. An indicator weight selects a prescribed joint type for counting purposes while leaving every possible collider in the graph. Our argument uses this isolation principle with a fixed full type and derives the required collision estimate from permutation symmetry. The selection of a full joint type is not presumed to be a local tensor restriction: the local restriction is obtained only after isolation in the full support.

There is a related distinction in the entropy optimization of [1, Theorem 1.52]. Its feasible pair distributions prescribe only the $2k$ one-coordinate marginals, not both full support-valued marginal laws. An entropy bound proved for a uniform full first law therefore cannot simply be substituted into that optimization. We do not invoke the theorem as a black box or claim to evaluate its displayed optimization. Instead, we count partners of each fixed candidate by exact quotient fibers, and bound their entropy by an invariant-subspace argument. This supplies the collision estimate needed for arbitrary positive multiplicities, without assuming separate all-binary or all-distinct cases.

Finally, diagonal slice rank and entropy upper bounds are established parts of the slice-rank method; see Tao and Sawin [3, Example 5 and Proposition 6]. We include elementary proofs of the facts used here. A score assigned to each local word partitions every support entry into a slice and gives $Q(D_{\mathbf{d}}^{\boxtimes N}) \le k \exp(N\,\mathrm{H}(p))$ for every $N \ge 1$. This upper bound applies to arbitrary complex linear maps and completes the comparison with the coordinate restrictions constructed below.

3 Powers, limits and the two-factor case

Each basis vector in a factor of $D_{\mathbf{d}}^{\boxtimes N}$ is indexed by a word of length N. A support edge is a k-tuple of local words such that each column has multiplicities $\mathbf{d}$. A set of edges is an *induced matching* if they have pairwise distinct local words in every factor and keeping just their local words leaves no other support edge. A coordinate restriction to such a set, followed by relabelling, gives a diagonal of that cardinality.

Lemma 3.1. *For $R = Q$ and for $R = Q_{\mathrm{M}}$, the limit $\lim_{N\to\infty} N^{-1} \log R(D_{\mathbf{d}}^{\boxtimes N})$ exists and equals $\sup_{N\ge 1} N^{-1} \log R(D_{\mathbf{d}}^{\boxtimes N})$. Consequently, a lower bound along any unbounded subsequence also bounds this limit from below.*

Proof. Keeping a single nonzero entry gives $R(D_{\mathbf{d}}^{\boxtimes N}) \ge 1$. Tensoring two restrictions, and identifying a pair of diagonal output indices with one index, gives

$$R(D_{\mathbf{d}}^{\boxtimes (N+L)}) \ge R(D_{\mathbf{d}}^{\boxtimes N}) R(D_{\mathbf{d}}^{\boxtimes L}). \tag{3.1}$$

The flattening along one factor has rank at most n^N, cannot increase under restriction, and has rank q on Δ_q. Thus $R \le n^N$. Put $a_N = \log R(D_{\mathbf{d}}^{\boxtimes N})$ and $a_0 = 0$. Then $0 \le a_N \le N \log n$ and (a_N) is superadditive. Fix $m \ge 1$ and write $N = hm + r$ with $0 \le r < m$. Superadditivity gives $a_N \ge h a_m + a_r \ge h a_m$, whence

$$\liminf_{N\to\infty} \frac{a_N}{N} \ge \frac{a_m}{m}.$$

Take the supremum over m. The reverse upper bound is part of the definition of that supremum, proving the assertion. □

If $k = 2$, positivity and $n \ge 2$ force $n = 2$ and $\mathbf{d} = (1, 1)$. The tensor $D_{(1,1)} = e_1 \otimes e_2 + e_2 \otimes e_1$ is a permutation matrix. Its powers have $Q_{\mathrm{M}} = Q = 2^N$, so (1.2) holds. Henceforth we may assume $k \ge 3$ in the lower-bound construction.

4 An entropy inequality from irreducible invariance

We use finite Shannon entropy with natural logarithms. For a random variable X with finite support, write

$$\mathrm{H}(X) = -\sum_x \mathbb{P}(X = x) \log \mathbb{P}(X = x),$$

where $0 \log 0 = 0$, and define conditional entropy by averaging over the conditioning variable. The chain rule is $\mathrm{H}(X, Y) = \mathrm{H}(Y) + \mathrm{H}(X \mid Y)$. We will use nonnegativity of conditional mutual information. To recall its elementary basis, for probability vectors p and q the relative entropy

$$\mathrm{KL}(p\|q) = \sum_{x:p_x>0} p_x \log \frac{p_x}{q_x}$$

is nonnegative, with value $+\infty$ if some $p_x > 0$ has $q_x = 0$. Otherwise, the inequality $-\log u \geq 1 - u$ gives

$$\mathrm{KL}(p\|q) \geq \sum_{x:p_x>0} (p_x - q_x) \geq 0.$$

Conditional mutual information is the average, over the conditioning variable, of the relative entropy of a conditional joint law with respect to the product of its marginals. Thus

$$\mathrm{I}(A; B \mid C) = \mathrm{H}(A \mid C) + \mathrm{H}(B \mid C) - \mathrm{H}(A, B \mid C) \geq 0.$$

In particular, if C is a deterministic function of A and also of B, the chain rule yields

$$\mathrm{I}(A; B) = \mathrm{H}(C) + \mathrm{I}(A; B \mid C) \geq \mathrm{H}(C). \tag{4.1}$$

Let E be a finite-dimensional vector space over $\mathbb{Q}$, and let X be an E-valued random variable with finite support. For a rational linear subspace $V \leq E$, the notation $X \bmod V$ denotes its coset label, which is a finite-valued random variable. Put

$$F_X(V) = \mathrm{H}(X \mid X \bmod V) = \mathrm{H}(X) - \mathrm{H}(X \bmod V). \tag{4.2}$$

The last equality holds because the coset label is a function of X. In particular, $F_X(0) = 0$ and $F_X(E) = \mathrm{H}(X)$.

Lemma 4.1 (Supermodularity)**.** *For rational linear subspaces $U, V \leq E$,*

$$F_X(U) + F_X(V) \leq F_X(U + V) + F_X(U \cap V). \tag{4.3}$$

Proof. Set $A = X \bmod U$, $B = X \bmod V$, and $C = X \bmod (U+V)$. The pair (A, B) and the label $X \bmod (U \cap V)$ determine each other on the finite support: two support points have the same pair of labels if and only if their difference belongs to both U and V. Moreover, C is a deterministic function of each of A and B. Equation (4.1) therefore gives

$$\mathrm{H}(X \bmod U) + \mathrm{H}(X \bmod V) - \mathrm{H}(X \bmod (U \cap V)) \geq \mathrm{H}(X \bmod (U + V)).$$

Substituting (4.2) proves the assertion. □

Lemma 4.2 (Irreducible invariance)**.** *Suppose that $\dim E = t > 0$, a group G acts $\mathbb{Q}$-linearly and irreducibly on E, and gX has the same law as X for every $g \in G$. Then every rational subspace $V \leq E$ satisfies*

$$F_X(V) \leq \frac{\dim V}{t}\, \mathrm{H}(X). \tag{4.4}$$

Proof. A finite set has only finitely many partitions. The partition of the support into cosets of V determines $F_X(V)$, so this quantity takes only finitely many values as V varies. Since the positive dimensions also belong to a finite set, the maximum

$$\lambda = \max_{0 \neq V \leq E} \frac{F_X(V)}{\dim V}$$

exists. Choose a maximizing subspace V of largest dimension. Invariance of the law of X gives $F_X(gV) = F_X(V)$, so gV is another maximizer for every $g \in G$. Supermodularity and the definition of λ imply

$$\begin{aligned} \lambda\big(\dim V + \dim gV\big) &\leq F_X(V + gV) + F_X(V \cap gV) \\ &\leq \lambda\big(\dim(V + gV) + \dim(V \cap gV)\big). \end{aligned}$$

The bound on the intersection also holds when it is zero because $F_X(0) = 0$. The two outer expressions coincide by the dimension formula. Thus $V + gV$ attains the maximum slope as well. The choice of V forces $\dim(V + gV) = \dim V$, whence $gV = V$. Since this holds for every g, irreducibility and $V \neq 0$ imply $V = E$. Consequently $\lambda = F_X(E)/t = \mathrm{H}(X)/t$, proving the bound for every nonzero subspace; the zero subspace satisfies it directly. The argument also applies when $\mathrm{H}(X) = 0$. □

5 Permutation fibers and counts over the full support

Fix the positive multiplicity tuple $\mathbf{d} = (d_1, \dots, d_n)$, with $n \geq 2$ and $k = \sum_a d_a \geq 3$. All constants in the remainder of the proof may depend on this fixed tuple. Define

$$p_a = \frac{d_a}{k}, \qquad M = \frac{k!}{\prod_{a=1}^{n} d_a!}, \qquad h = \log M, \qquad \alpha_a = ka - \sum_{c=1}^{n} c\, d_c. \tag{5.1}$$

The integers α_a are distinct. Let $S \subseteq \mathbb{Z}^k$ be the set of tuples containing α_a exactly d_a times for every a. Replacing the symbol a by α_a is only an injective labelling of the basis symbols of $D_{\mathbf{d}}$; it is not an operation on the tensor. We have $|S| = M$ and

$$\sum_{i=1}^{k} s_i = 0 \qquad (s \in S).$$

Let Z be uniform on S. Each coordinate Z_j has distribution p on the integer labels, so

$$\mathrm{H}(Z) = h, \qquad \mathrm{H}(Z_j) = \mathrm{H}(p), \qquad \mathrm{H}(p) = -\sum_{a=1}^{n} p_a \log p_a.$$

For $1 \leq j \leq k$, put

$$E_j = \left\{ v \in \mathbb{Q}^k : \sum_{i=1}^{k} v_i = 0,\ v_j = 0 \right\}, \qquad \dim E_j = k - 2, \qquad \beta = \frac{h - \mathrm{H}(p)}{k-2} \geq 0. \tag{5.2}$$

Nonnegativity follows from $h - \mathrm{H}(p) = \mathrm{H}(Z \mid Z_j)$.

Lemma 5.1 (Entropy in a fixed-coordinate fiber)**.** *If $V \leq E_j$ is a rational subspace of dimension r, then*

$$F_Z(V) \leq r\beta. \tag{5.3}$$

Proof. Since every vector of V has coordinate j equal to zero, the quotient label $Z \bmod V$ determines Z_j. Condition on $Z_j = \alpha_a$. The other $k-1$ coordinates are uniformly distributed among permutations of the multiset obtained by removing one copy of α_a; their sum is $-\alpha_a$. Set the jth coordinate to zero and replace every other coordinate Z_i by $Z_i + \alpha_a/(k-1)$. This translates the conditional support into E_j. The translation preserves differences, entropy, and the conditional entropy modulo V. The translated law is invariant under all permutations of the $k-1$ coordinates different from j.

The resulting standard action of the symmetric group on E_j is irreducible over $\mathbb{Q}$. Indeed, let $W \leq E_j$ be a nonzero invariant subspace and choose $v \in W \setminus \{0\}$. Among the $k-1$ moving coordinates, two entries v_u, v_w must differ, since they sum to zero. Subtracting the image of v under their transposition yields

$$(v_u - v_w)(e_u - e_w) \in W.$$

The coefficient is nonzero. Permuting this difference gives all coordinate differences among the moving coordinates, which span E_j. Thus $W = E_j$. The same argument includes $k = 3$, when E_j is one-dimensional.

Lemma 4.2, applied to each translated conditional law, now gives

$$\mathrm{H}(Z \mid Z \bmod V,\ Z_j = \alpha_a) \leq \frac{r}{k-2}\ \mathrm{H}(Z \mid Z_j = \alpha_a).$$

A conditional law supported at one point has entropy zero and obeys the same inequality. Average with weights $\mathbb{P}(Z_j = \alpha_a) = p_a$. Since $Z \bmod V$ determines Z_j, the left side averages to $F_Z(V)$. The right side averages to

$$\frac{r}{k-2}\ \mathrm{H}(Z \mid Z_j) = \frac{r}{k-2}\big(h - \mathrm{H}(p)\big) = r\beta,$$

which proves the claim. □

For $s \in S$ and a rational subspace $V \le E_j$, define

$$C_V(s) = (s + V) \cap S.$$

Conditioned on a coset, the uniform variable Z is uniform on the corresponding fiber. Therefore

$$F_Z(V) = \frac{1}{M} \sum_{s \in S} \log |C_V(s)|. \tag{5.4}$$

Let N be a positive multiple of M. Identify a support edge of $D_{\mathbf{d}}^{\boxtimes N}$ with its sequence of N support columns, and define

$$\mathcal{A}_N = \left\{ I = (I(1), \dots, I(N)) \in S^N : \#\{t : I(t) = s\} = N/M \text{ for every } s \in S \right\}.$$

We call these edges *candidates*. They form a counting class; we do not assume that the full joint type $\mathcal{A}_N$ can itself be selected by local linear maps.

For a fixed candidate $I \in \mathcal{A}_N$, each column of a partner $J \in S^N$ satisfying $J(t) - I(t) \in V$ can be chosen independently from $C_V(I(t))$. Hence the count over the *entire original support* is exactly

$$\begin{aligned} \#\left\{ J \in S^N : J(t) - I(t) \in V \text{ for all } 1 \le t \le N \right\} &= \prod_{s \in S} |C_V(s)|^{N/M} \\ &= \exp\bigl(N F_Z(V)\bigr) \le \exp(Nr\beta), \qquad r = \dim V. \end{aligned} \tag{5.5}$$

No joint-type or marginal-type condition is imposed on J. In particular, (5.5) also bounds the smaller class of partners whose column differences span exactly V.

Finally, the size of the candidate class satisfies

$$\frac{M^N}{(N+1)^M} \le |\mathcal{A}_N| = \frac{N!}{((N/M)!)^M} \le M^N. \tag{5.6}$$

There are at most $(N+1)^M$ count vectors for words in an M-letter alphabet. The balanced vector maximizes its multinomial coefficient when M divides N: if two counts differ by at least two, transferring one occurrence from the larger to the smaller strictly decreases the factorial denominator. At a maximum, all counts therefore equal N/M. Since the multinomial classes partition the M^N words, their largest class has size at least $M^N/(N+1)^M$. The upper bound follows from $\mathcal{A}_N \subseteq S^N$. For fixed $\mathbf{d}$, equation (5.6) yields

$$\log |\mathcal{A}_N| = Nh - o(N).$$

6 Large average-free sets

The hashing argument uses the following form of the no-carry sphere construction. Generalized average-free sets of this kind occur in [1, 4]. We include the construction and its size estimate to specify exactly the property needed here.

Lemma 6.1 (Average-free sets)**.** *Fix an integer $L \ge 2$. For every sufficiently large integer q there is a set*

$$B \subseteq \left\{ 0, \dots, \left\lfloor \frac{q-1}{L} \right\rfloor \right\}$$

such that

$$x_1 + \cdots + x_L = Ly, \qquad x_1, \dots, x_L, y \in B, \quad \Longrightarrow \quad x_1 = \cdots = x_L = y,$$

and

$$\log |B| = \log q - O_L\left(\sqrt{\log q}\right).$$

The integer q need not be prime in this lemma.

Proof. Put

$$\ell = \left\lfloor \sqrt{\log q} \right\rfloor, \qquad m = \left\lfloor \frac{(q/(2L))^{1/\ell}}{L} \right\rfloor, \qquad b = Lm.$$

For sufficiently large q, we have $\ell \geq 1$, $m \geq 2$, and $b^\ell \leq q/(2L)$. There are m^ℓ vectors in $\{0, \ldots, m-1\}^\ell$ and at most $\ell(m-1)^2 + 1$ possible values of their squared Euclidean norms. Choose a norm level $\mathcal{C}$ with

$$|\mathcal{C}| \geq \frac{m^\ell}{\ell(m-1)^2 + 1},$$

and encode its vectors in base b:

$$B = \left\{ \sum_{t=0}^{\ell-1} a_t b^t : (a_0, \ldots, a_{\ell-1}) \in \mathcal{C} \right\}.$$

The encodings are distinct and lie between 0 and $b^\ell - 1$. Since $b^\ell \leq q/(2L)$, they belong to the required interval.

Suppose that $x_1 + \cdots + x_L = Ly$ for elements of B, and let $a^{(1)}, \ldots, a^{(L)}, c \in \mathcal{C}$ be their digit vectors. Every digit sum on either side is at most $L(m-1) < b$. Thus neither side has carries, and uniqueness of the base-b expansion gives

$$a^{(1)} + \cdots + a^{(L)} = Lc.$$

All these vectors have the same squared norm. Consequently

$$\frac{1}{L} \sum_{i=1}^{L} \|a^{(i)} - c\|^2 = \frac{1}{L} \sum_{i=1}^{L} \|a^{(i)}\|^2 - \|c\|^2 = 0.$$

Each $a^{(i)}$ equals c, proving the average-free property.

For the size estimate, write $t = \log q$. Then $\ell \asymp \sqrt{t}$ and

$$\log m = \frac{t - \log(2L)}{\ell} - \log L + O_L(m^{-1}) = \sqrt{t} + O_L(1).$$

In particular $m \to \infty$, and

$$\ell \log m = t - \log(2L) - \ell \log L + O_L(\ell/m) = t - O_L(\sqrt{t}).$$

Moreover,

$$\log\big(\ell(m-1)^2 + 1\big) \leq \log(\ell + 1) + 2 \log m = O_L(\sqrt{t}).$$

The lower bound for $|\mathcal{C}| = |B|$ therefore gives $\log |B| \geq \log q - O_L(\sqrt{\log q})$. The opposite estimate $|B| \leq q$ completes the proof. □

7 Hashes and exact survival probabilities

We retain the notation $S \subseteq \mathbb{Z}^k$, M, h, E_j, and β from Section 5, with $k \geq 3$ fixed. In particular, every column in S has coordinate sum zero. We use the tight-support hashing construction of [1]; the uniform rank bound and the exact probabilities needed here are established below.

Lemma 7.1 (Uniform reduction of ranks)**.** *Let*

$$C = \max_{s,s' \in S} \max_{1 \leq i \leq k} |s_i - s'_i|, \qquad B_0 > \max\{k,\, k! \max\{1, C\}^k\}.$$

For every prime $q > B_0$, any integer matrix whose columns belong to $S - S$ has the same rank over $\mathbb{Q}$ and over $\mathbb{F}_q$. The threshold is independent of the number of columns.

Proof. If the rational rank is zero, every entry is zero and the assertion is immediate. Otherwise, let that rank be r. There is a nonzero integer $r \times r$ minor, of absolute value at most

$$r! C^r \leq k! \max\{1, C\}^k < q.$$

It remains nonzero modulo q, so the rank over $\mathbb{F}_q$ is at least r. Every larger minor is zero as an integer and remains zero modulo q, proving the opposite inequality. The argument depends only on the entries and the number k of rows. □

Fix a prime $q > B_0$ large enough for Lemma 6.1 with $L = k-1$, and fix a positive multiple N of M. Choose the set B from that lemma, viewed also as a subset of $\mathbb{F}_q$ by reduction modulo q. Choose

$$v = (v_1, \dots, v_N) \in \mathbb{F}_q^N, \qquad u_1, \dots, u_{k-1} \in \mathbb{F}_q$$

independently and uniformly. For an integer-labelled local word $w = (w_1, \dots, w_N)$, define

$$b_i(w) = u_i + \sum_{t=1}^{N} w_t v_t \qquad (1 \le i < k),$$
$$b_k(w) = \frac{\sum_{i<k} u_i - \sum_{t=1}^{N} w_t v_t}{k-1}.$$

These expressions are in $\mathbb{F}_q$; the denominator is invertible because $q > k$.

An original support edge $I \in S^N$ may be described by its columns $I(t) \in S$ or by its k local words $I_1, \dots, I_k$. The zero-sum condition on every column gives

$$\sum_{i<k} b_i(I_i) = (k-1) b_k(I_k) \qquad \text{in } \mathbb{F}_q. \tag{7.1}$$

Keep in each mode exactly the local basis words whose hash belongs to B. This is a coordinate restriction of all k factors. If I survives, represent each of its hashes by the corresponding integer in B. Both sides of (7.1) then belong to $[0, q-1]$, so their congruence is an equality of integers. Lemma 6.1 implies that all k hashes equal a single element of B. Conversely, this condition implies survival.

For a fixed edge I and a prescribed common hash $b \in B$, each choice of v permits exactly one offset tuple: set $u_i = b - \sum_t I_i(t) v_t$ for $i < k$. The last hash then also equals b by (7.1). The events for distinct b are disjoint. Of the q^{N+k-1} choices of all hashing variables, exactly $|B| q^N$ make I survive. Thus

$$\mathbb{P}[I \text{ survives}] = |B| q^{-(k-1)}. \tag{7.2}$$

Now let distinct original edges $I, J \in S^N$ share a local word, say $I_j = J_j$. Put

$$V = \operatorname{span}_{\mathbb{Q}} \{ I(t) - J(t) : 1 \le t \le N \}, \qquad r = \dim V.$$

Every column difference has coordinate sum zero and coordinate j equal to zero. Hence $V \subseteq E_j$ and $r \le k-2$. Also $r \ge 1$, since the integer labelling is injective and $I \ne J$.

If both edges survive, their common hash values agree: the shared local word has the same hash for both. For a prescribed common $b \in B$, the offsets are uniquely determined by I once v is chosen. The additional conditions on J are precisely

$$\sum_{t=1}^{N} \big(I_i(t) - J_i(t) \big) v_t = 0, \qquad 1 \le i < k.$$

The rows of the full $k \times N$ difference matrix sum to zero. Its last row is therefore redundant, over both $\mathbb{Q}$ and $\mathbb{F}_q$. Deleting that row leaves rank r by Lemma 7.1. The displayed system has exactly q^{N-r} solutions for v, each with one associated offset tuple. Consequently

$$\mathbb{P}[I \text{ and } J \text{ survive}] = |B| q^{-(k-1+r)}. \tag{7.3}$$

This argument also applies when the shared mode is $j = k$: its shared hash still forces equality of the two common values, and the last row remains redundant. Neither edge is required to have uniform joint type for (7.2) or (7.3).

8 Isolation against the full original support

We now distinguish the counting class $\mathcal{A}_N$ from the support S^N of the full tensor power. Let X count surviving candidates $I \in \mathcal{A}_N$. Let Y count ordered pairs (I, J) such that I is a surviving candidate, $J \ne I$ is any surviving original edge, and I and J share a local word in at least one mode. In particular, no type condition is imposed on J. Equation (7.2) gives

$$\mathbb{E}X = |\mathcal{A}_N| \, |B| q^{-(k-1)}. \tag{8.1}$$

Let $\mathcal{V}$ be the collection of distinct nonzero rational spans of subsets of $S - S$ that are contained in at least one E_j, and put $K = |\mathcal{V}|$. This collection is finite because $S - S$ is finite; it depends on $\mathbf{d}$, but not on N. Every $V \in \mathcal{V}$ has dimension between 1 and $k - 2$, and each pair counted by Y has exactly one such difference span.

Fix $I \in \mathcal{A}_N$ and $V \in \mathcal{V}$ of dimension r. The number of partners $J \in S^N$ whose column differences with I span exactly V is at most the number whose differences all belong to V. By (5.5) and Lemma 5.1, this latter number is at most $\exp(Nr\beta)$. We use a mode j with $V \subseteq E_j$ only to apply that entropy estimate; we do not sum over choices of j. Summing (7.3) over the candidates and their possible difference spans therefore yields

$$\mathbb{E}Y \le \mathbb{E}X \sum_{V \in \mathcal{V}} \exp\bigl(N \dim(V)\beta\bigr) q^{-\dim(V)}. \tag{8.2}$$

The larger fiber count is an upper bound even though it also allows partners of smaller difference span. Each actual pair is assigned to its unique exact span in this sum, so there is no extra multiplicity for the shared mode.

For a realization of the hashing variables, call a surviving candidate *isolated* if it shares no local word with any other surviving original edge. Denote the set of these candidates by $\mathcal{I}$. Every nonisolated candidate contributes at least one ordered pair to Y, and hence

$$|\mathcal{I}| \ge X - Y.$$

This is the candidate-indicator instance of the weighted global-isolation principle from [5]. The point needed for a coordinate restriction is that the comparison uses the entire original support, as follows.

If $\mathcal{I}$ is empty, then $X - Y \le 0$ and the corresponding lower estimate follows already from $Q_{\mathrm{M}} \ge 1$. Otherwise, retain exactly the local words belonging to the edges in $\mathcal{I}$. Distinct edges of $\mathcal{I}$ have distinct local words in every mode, since they are isolated. Any original edge remaining after this final coordinate restriction would have survived the earlier hash restriction. It also shares a retained word with some $I \in \mathcal{I}$. If it differed from I, it would contradict the definition of isolation. Thus the remaining support is exactly $\mathcal{I}$, and independent relabelling in the k modes gives the diagonal tensor of size $|\mathcal{I}|$.

The hashing probability space is finite, so some realization has $X - Y \ge \mathbb{E}(X - Y)$. We have proved

$$Q_{\mathrm{M}}\bigl(D_{\mathbf{d}}^{\boxtimes N}\bigr) \ge \mathbb{E}(X - Y). \tag{8.3}$$

No local restriction selecting $\mathcal{A}_N$ was used. Isolation only within the candidate class would not justify the conclusion: an edge of $S^N \setminus \mathcal{A}_N$ could remain on the retained local words.

9 The entropy lower bound

Fix $\delta > 0$ and set $\lambda = \beta + \delta > 0$. Choose primes q tending to infinity, and for sufficiently large q define

$$N = N(q) = M \left\lfloor \frac{\log q}{M\lambda} \right\rfloor.$$

Then N is a positive multiple of M, tends to infinity with q, and satisfies

$$N\lambda \le \log q < (N + M)\lambda. \tag{9.1}$$

The primes are chosen first and the tensor powers second. Thus unboundedness of the primes suffices; no estimate for prime gaps is needed. For all sufficiently large q, both the rank threshold and Lemma 6.1 apply.

For $r = \dim V \ge 1$, equation (9.1) gives

$$\exp(Nr\beta) q^{-r} \le \exp(-Nr\delta) \le \exp(-N\delta).$$

It follows from (8.2) that

$$\mathbb{E}Y \le \mathbb{E}X \, K \exp(-N\delta).$$

Combining this inequality with (8.1) and (8.3) yields

$$Q_{\mathrm{M}}\bigl(D_{\mathbf{d}}^{\boxtimes N}\bigr) \ge \bigl(1 - Ke^{-N\delta}\bigr) |\mathcal{A}_N| \, |B| q^{-(k-1)} \tag{9.2}$$

for these powers and all sufficiently large q. The prefactor is positive and tends to one.

All of $\mathbf{d}, k, M, K, \beta, \delta$ are fixed before q tends to infinity. By (5.6), $\log|\mathcal{A}_N| = Nh - o(N)$. Lemma 6.1 gives $\log|B| = \log q - O_k(\sqrt{\log q})$, and (9.1) gives $\log q = N(\beta+\delta) + O_{\mathbf{d},\delta}(1)$. In particular $\log q = O_{\mathbf{d},\delta}(N)$, so the error in the average-free estimate is $o(N)$. The logarithm of the right side of (9.2) is consequently

$$\begin{aligned} Nh - (k-2)\log q - o(N) &= N\big[h - (k-2)(\beta+\delta)\big] - o(N) \\ &= N\big[\mathrm{H}(p) - (k-2)\delta\big] - o(N). \end{aligned}$$

Here we used $(k-2)\beta = h - \mathrm{H}(p)$.

Lemma 3.1 transfers this lower bound along the unbounded sequence of powers to the asymptotic limit, giving

$$\log \widetilde{Q}_{\mathrm{M}}(D_{\mathbf{d}}) \geq \mathrm{H}(p) - (k-2)\delta.$$

Since $\delta > 0$ was arbitrary, and every coordinate restriction is a linear restriction, we conclude that

$$\widetilde{Q}(D_{\mathbf{d}}) \geq \widetilde{Q}_{\mathrm{M}}(D_{\mathbf{d}}) \geq \exp\big(\mathrm{H}(p)\big). \tag{9.3}$$

10 The upper bound for arbitrary linear restrictions

A *slice in mode i* is a tensor of the form $a \otimes B$, with a in the ith factor and B in the tensor product of the other factors, placed in their original order. The *slice rank* $\mathrm{sr}(T)$ is the least number of slices summing to T; their distinguished modes may differ. Applying linear maps independently to the factors takes a slice to a slice or zero. We first prove the diagonal value needed to bound ordinary subrank. This fact and the entropy upper-bound method have precedents in Tao and Sawin [3]; the required arguments follow.

Lemma 10.1. *For subspaces $U, V \subseteq \mathbb{C}^q$, let $U * V$ be the linear span of their coordinatewise products. Then*

$$\dim(U * V) \geq \dim U + \dim V - q. \tag{10.1}$$

Consequently, for any $\ell \geq 1$ and $U_1, \dots, U_\ell \subseteq \mathbb{C}^q$,

$$\dim \operatorname{span}\{u_1 * \cdots * u_\ell : u_i \in U_i\} \geq q - \sum_{i=1}^{\ell} \operatorname{codim} U_i. \tag{10.2}$$

Proof. Let A be the union of the coordinate supports of vectors in U. There is $u \in U$ nonzero in every coordinate of A: each such coordinate is a nonzero linear function on U, and a finite union of their proper kernels cannot cover a vector space over the infinite field $\mathbb{C}$. Equivalently, the product of these linear functions is a nonzero polynomial and is nonzero somewhere. If $U = 0$, take A empty and $u = 0$.

Coordinatewise multiplication by u on $\mathbb{C}^q$ has kernel of dimension $q - |A|$. Since $\dim U \leq |A|$, its restriction to V gives

$$\dim(U * V) \geq \dim(u * V) \geq \dim V - (q - |A|) \geq \dim V + \dim U - q.$$

Iterate this inequality, using bilinearity of coordinatewise multiplication to identify the successive spans. The case $\ell = 1$ of (10.2) is an equality. □

Lemma 10.2. *For every $k \geq 2$ and $q \geq 1$, $\mathrm{sr}(\Delta_q) = q$. In particular, $Q(T) \leq \mathrm{sr}(T)$ for every k-tensor T over $\mathbb{C}$.*

Proof. Consider any slice decomposition of Δ_q. For each i, let r_i be the number of its slices in mode i, and let $U_i \subseteq \mathbb{C}^q$ annihilate their distinguished coefficient vectors under the bilinear pairing $\sum_j x_j y_j$. Then $\operatorname{codim} U_i \leq r_i$. Evaluating the decomposition on $U_1 \times \cdots \times U_k$ kills every slice, so

$$\sum_{j=1}^{q} u_{1,j} \cdots u_{k,j} = 0 \qquad (u_i \in U_i).$$

Thus the span of $U_1 * \cdots * U_{k-1}$ lies in $U_k^{\perp}$, whose dimension is $\operatorname{codim} U_k$. By (10.2),

$$q - \sum_{i<k} \operatorname{codim} U_i \leq \operatorname{codim} U_k, \qquad q \leq \sum_i \operatorname{codim} U_i \leq \sum_i r_i.$$

The defining expression for Δ_q is a sum of q slices, proving equality. Monotonicity of slice rank under independent linear maps now gives $q = \mathrm{sr}(\Delta_q) \leq \mathrm{sr}(T)$ for every diagonal restriction of T, and hence $Q(T) \leq \mathrm{sr}(T)$. □

Proposition 10.3. *For all $N \geq 1$,*

$$Q(D_{\mathbf{d}}^{\boxtimes N}) \leq \mathrm{sr}(D_{\mathbf{d}}^{\boxtimes N}) \leq k \exp(N\,\mathrm{H}(p)).$$

Proof. For a local word $w = (w_1, \ldots, w_N) \in \{1, \ldots, n\}^N$, set

$$s(w) = -\sum_{t=1}^{N} \log p_{w_t}, \qquad P(w) = \prod_{t=1}^{N} p_{w_t} = e^{-s(w)}, \qquad L_N = \{w : s(w) \leq N\,\mathrm{H}(p)\}.$$

All p_a are positive. Since $\sum_w P(w) = 1$ and every $w \in L_N$ has $P(w) \geq e^{-N\,\mathrm{H}(p)}$, we have

$$|L_N| \leq e^{N\,\mathrm{H}(p)}. \tag{10.3}$$

For every support edge $(w^{(1)}, \ldots, w^{(k)})$ of the full tensor power, each column has multiplicities $\mathbf{d}$. Therefore

$$\sum_{i=1}^{k} s(w^{(i)}) = -N \sum_{a=1}^{n} d_a \log p_a = Nk\,\mathrm{H}(p).$$

At least one local word belongs to L_N. Assign the edge to the least such mode. For each i and $w \in L_N$, collect all assigned terms whose ith word is w into one slice with distinguished factor e_w. This partitions the entire support and yields at most $k|L_N|$ slices. Apply Lemma 10.2 and (10.3). □

Taking Nth roots in Proposition 10.3 and using Lemma 3.1 gives $\widetilde{Q}(D_{\mathbf{d}}) \leq e^{\mathrm{H}(p)}$. Together with the coordinate-restriction lower bound (9.3) and $\widetilde{Q}_{\mathrm{M}} \leq \widetilde{Q}$, this proves Theorem 1.1 for $k \geq 3$. The two-factor case was handled in Section 3. Every restriction in the argument retains the same k factors.

11 Exact finite verification

The proof above establishes the theorem for every multiplicity tuple. Two accompanying standard-library Python programs provide finite exact checks of its entropy and hashing interfaces. They use integer and rational arithmetic, explicit failure exits and negative controls; no floating-point comparison or numerical tolerance enters the checks. The finite computations are supporting regressions, not premises of the universal theorem or an extrapolation to larger multiplicities.

11.1 Collision subspaces and the entropy inequality

The program `check_entropy.py` enumerates all nonzero difference-generated collision subspaces for one fixed shared mode in the five supports shown in Table 1. The other modes are equivalent by permutation symmetry. It forms the quotient fibers and independently checks their cardinalities by exact difference membership. The zero subspace has the immediate zero-entropy bound and is omitted.

For a checked support of size M, let $A = \prod_{s \in S} |C_V(s)|$ and $r = \dim V$. Rather than approximate logarithms in Lemma 5.1, the program checks the equivalent integer inequality

$$A^{k(k-2)} k^{kMr} \leq M^{kMr} \left(\prod_{a=1}^{n} d_a^{d_a} \right)^{Mr}. \tag{11.1}$$

Indeed, $\mathrm{H}(p) = \log k - k^{-1} \sum_a d_a \log d_a$, and multiplying $M^{-1} \log A \leq r(\log M - \mathrm{H}(p))/(k-2)$ by $Mk(k-2)$ and exponentiating gives (11.1).

Each case includes a negative control that decreases the full collision space's claimed dimension by one while leaving its fibers fixed. The same inequality validator must reject the resulting false claim. All five controls reject. This tests selected dimensions and supports; the invariant-entropy argument, rather than this enumeration, proves the all-subspace statement.

Multiplicities $\mathbf{d}$	Support size M	Nonzero collision subspaces
$(1,1,1)$	6	1
$(2,1,1)$	12	7
$(1,1,1,1)$	24	13
$(2,2,1)$	30	191
$(3,1,1)$	20	176

Table 1: Exact finite checks of the conditional entropy bound.

11.2 Hashing and isolation in the original support

The program `check_hashing.py` checks 405 rank-one collision-subspace cases, counting the shared mode as part of the case, across multiplicities

$$(2,1),\ (1,1,1),\ (2,2),\ (2,1,1),\ (1,1,1,1),\ (3,1,1),\ (2,2,1).$$

It also exhausts a complete small hashing experiment. Take

$$S=\{(-1,-1,2),(-1,2,-1),(2,-1,-1)\},\quad N=3,\quad q=11,\quad B=\{0,1,3\}.$$

There are $11^5 = 161{,}051$ assignments of the common hash vector and two offsets, 27 original support edges, six candidates of uniform joint type, and 54 ordered candidate/collider pairs sharing a local word. For every assignment, the program checks equal hashes on surviving edges, the $X-Y$ isolation inequality, and that retaining the isolated candidates' vertices leaves exactly those edges in the *full* original support.

Each edge survives in 3,993 assignments and each listed pair in 363, matching (7.2) and (7.3) exactly. The totals are $\sum X = 23{,}958$ and $\sum Y = 19{,}602$, so the average of $X-Y$ is positive. The largest isolated matching obtained has two edges. These small numbers test the mechanism, not its asymptotic rate.

Three further controls reject: a two-point distribution on a line in dimension two refutes omission of invariance from the entropy-slope bound; replacing B by $\{0,1,2\}$ violates average-freeness through $0+2=2\cdot 1$; and keeping only candidate edges $(0,0)$ and $(1,1)$ while the original support also contains $(0,1)$ exposes a hidden mixed edge. The first is a counterexample to dropping an assumption, not an algorithm for detecting all invariant distributions.

Both checkers pass with byte-identical mathematical output under normal Python and optimized Python (`-O`); all eight negative controls reject in both modes. Reproduction instructions and exact input/output digests accompany the manuscript. These computations make no claim about a uniform convergence rate, efficient extraction of asymptotic restrictions, or formal proof-assistant certification.

Statements and Declarations

Funding. The research of Wenqi Zhu was supported by the Engineering and Physical Sciences Research Council under grant EP/Y028872/1, *Mathematical Foundations of Intelligence: An "Erlangen Programme" for AI.* Ping Nie gratefully acknowledges the support of AI4ALPHA and NetMind.AI.

Competing interests. The authors have no competing interests to declare that are relevant to the content of this article.

Data availability. The exact verification scripts, certificates and Lean 4 files supporting the finite computations are available from the corresponding author on reasonable request.

Use of generative AI and AI-assisted technologies. During the preparation of this work, the authors used AI tools (Claude, Anthropic; Codex, OpenAI; Qwen APP, Alibaba) to generate candidate constructions and code, perform computations, draft mathematical arguments, and improve the readability of the text. After using these tools, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication. Finite exact checks and a Lean 4 formalization of the algebraic and combinatorial core accompany the written proof, as described in Section 11.

References

[1] Matthias Christandl, Péter Vrana, and Jeroen Zuiddam. Asymptotic tensor rank of graph tensors: beyond matrix multiplication. *Computational Complexity*, 28:57–111, 2019. doi: 10.1007/s00037-019-0172-8.

[2] Alessandro Oneto and Emanuele Ventura. Ranks of tensors: geometry and applications. *Bollettino dell'Unione Matematica Italiana*, 18:751–778, 2025. doi: 10.1007/s40574-025-00472-9.

[3] Terence Tao and Will Sawin. Notes on the "slice rank" of tensors. https://terrytao.wordpress.com/2016/08/24/notes-on-the-slice-rank-of-tensors/, August 2016. 24 August 2016.

[4] Péter Vrana and Matthias Christandl. Asymptotic entanglement transformation between W and GHZ states. *Journal of Mathematical Physics*, 56(2):022204, 2015. doi: 10.1063/1.4908106. Author version: arXiv:1310.3244v2.

[5] Péter Vrana and Matthias Christandl. Distillation of Greenberger–Horne–Zeilinger states by combinatorial methods. *IEEE Transactions on Information Theory*, 65(9):5945–5958, 2019. doi: 10.1109/TIT.2019.2908646. Author version: arXiv:1805.09096v2 (2020).