Universal points in the asymptotic spectrum of tensors

Christandl, M.; Vrana, P.; Zuiddam, J.

DOI
10.1145/3188745.3188766

Publication date
2018

Document Version
Other version

Published in
STOC'18

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Universal points in the asymptotic spectrum of tensors

Matthias Christandl, Péter Vrana and Jeroen Zuiddam

Abstract

The asymptotic restriction problem for tensors is to decide, given tensors \( s \) and \( t \), whether the \( n \)th tensor power of \( s \) can be obtained from the \( (n + o(n)) \)th tensor power of \( t \) by applying linear maps to the tensor legs (this we call restriction), when \( n \) goes to infinity. In this context, Volker Strassen, striving to understand the complexity of matrix multiplication, introduced in 1986 the asymptotic spectrum of tensors. Essentially, the asymptotic restriction problem for a family of tensors \( \mathcal{X} \), closed under direct sum and tensor product, reduces to finding all maps from \( \mathcal{X} \) to the nonnegative reals that are monotone under restriction, normalised on diagonal tensors, additive under direct sum and multiplicative under tensor product, which Strassen named spectral points. Spectral points are by definition an upper bound on asymptotic subrank and a lower bound on asymptotic rank. Strassen created the support functionals, which are spectral points for oblique tensors, a strict subfamily of all tensors.

Universal spectral points are spectral points for the family of all tensors. The construction of nontrivial universal spectral points has been an open problem for more than thirty years. We construct for the first time a family of nontrivial universal spectral points over the complex numbers, using the theory of quantum entropy and covariants: the quantum functionals. In the process we connect the asymptotic spectrum of all tensors to the quantum marginal problem and to the entanglement polytope. In entanglement theory, our results amount to the first construction of additive entanglement monotones for the class of stochastic local operations and classical communication.

To demonstrate the asymptotic spectrum, we reprove (in hindsight) recent results on the cap set problem by reducing this problem to computing the lowest point in the asymptotic spectrum of the reduced polynomial multiplication tensor, a prime example of Strassen. A better understanding of our universal spectral points construction may lead to further progress on related combinatorial questions. We additionally show that the quantum functionals characterise asymptotic slice rank for complex tensors.

Keywords. asymptotic restriction, asymptotic spectrum, tensors, fast matrix multiplication, cap set problem, reduced polynomial multiplication, stochastic local operations and classical communication (slocc), entanglement monotones, quantum entropy, moment polytope

MSC. 15A69, 14L24, 68Q17
Contents

1 Introduction 3
  1.1 The asymptotic restriction problem 3
  1.2 Asymptotic spectra for tensors 6
  1.3 This paper 8

2 Strassen support functionals 9
  2.1 Upper support functional 10
  2.2 Lower support functional 11
  2.3 Comparing the support functionals 13
  2.4 Robust and oblique tensors; spectral points 15

3 Quantum functionals 16
  3.1 Upper quantum functional 17
  3.2 Lower quantum functional 23
  3.3 Comparing the quantum functionals 25
  3.4 Universal spectral points 27

4 Several families of tensors 30
  4.1 Tight tensors 31
  4.2 The cap set problem 35
  4.3 Free tensors 38
  4.4 Generic tensors 41

5 Subrank, slice rank and multi-slice rank 46
  5.1 Definition 46
  5.2 Key properties 47
  5.3 Upper bound by support and quantum functionals 47

References 48
1. Introduction

1.1. The asymptotic restriction problem

We study the asymptotic restriction problem, following the pioneering work of Volker Strassen [Str86, Str87, Str88, Str91]. The asymptotic restriction problem is a problem about multilinear maps \( f : \mathbb{F}^{n_1} \times \cdots \times \mathbb{F}^{n_k} \to \mathbb{F} \) over an arbitrary field \( \mathbb{F} \). Letting \((e_1, \ldots, e_{n_i})\) be the standard basis of \( \mathbb{F}^{n_i} \), one may equivalently think of \( f \) as the \( k \)-tensor \( t \in \mathbb{F}^{n_1} \otimes \cdots \otimes \mathbb{F}^{n_k} \) defined by \( t = \sum f(e_{a_1}, \ldots, e_{a_k}) e_{a_1} \otimes \cdots \otimes e_{a_k} \) where \( a_i \) goes over \( \{1, \ldots, n_i\} \). To state the asymptotic restriction problem we need the concepts restriction and tensor product. Let \( f : \mathbb{F}^{n_1} \times \cdots \times \mathbb{F}^{n_k} \to \mathbb{F} \) and \( g : \mathbb{F}^{m_1} \times \cdots \times \mathbb{F}^{m_k} \to \mathbb{F} \) be multilinear maps. We say \( f \) restricts to \( g \), and write \( f \geq g \), if there are linear maps \( A_i : \mathbb{F}^{m_i} \to \mathbb{F}^{n_i} \) such that \( g = f \circ (A_1, \ldots, A_k) \) where \( \circ \) denotes composition. We naturally define the tensor product \( f \otimes g \) as the multilinear map \((\mathbb{F}^{n_1} \otimes \mathbb{F}^{m_1}) \times \cdots \times (\mathbb{F}^{n_k} \otimes \mathbb{F}^{m_k}) \to \mathbb{F} \) defined by \((v_1 \otimes w_1, \ldots, v_k \otimes w_k) \mapsto f(v_1, \ldots, v_k)g(w_1, \ldots, w_k)\). We say \( f \) restricts asymptotically to \( g \), written \( f \gtrsim g \), if there is a sequence of natural numbers \( a(n) \in o(n) \) such that

\[
f^{\otimes n + a(n)} \geq g^{\otimes n} \text{ when } n \to \infty.
\]

The asymptotic restriction problem is: given \( f \) and \( g \), decide whether \( f \gtrsim g \).

Applications of the asymptotic restriction problem include computing the computational complexity of matrix multiplication in algebraic complexity theory [AS81, Blä01, BI11, Lan14, Str69, CW90, Sto10, Wil12, LG14, CU03, CKSU05] (see also [BCS97, Lan12, Blä13, Lan17]), deciding the feasibility of an asymptotic transformation between pure quantum states via stochastic local operations and classical communication (slocc) in quantum information theory [BPR+00, DVC00, VDDMV02, HHHH09], bounding the size of combinatorial structures like cap sets and tri-colored sum-free sets in additive combinatorics [Ede04, Tao08, ASU13, CLP17, EG17, Tao16, BCC+17, KSS16, TS16], and bounding the query complexity of certain properties in algebraic property testing [KS08, BCSX10, Sha09, BX15, HX17, FK14].

There are naturally two directions in the asymptotic restriction problem, namely finding (1) constructions, i.e. matrices that carry out \( f \gtrsim g \), and (2) obstructions, i.e. certificates that prohibit \( f \gtrsim g \). For constructions one should think of fast matrix multiplication algorithms or efficient quantum protocols. For obstructions one should think of lower bounds in the sense of computational complexity theory. Strassen introduced in 1986 the theory of asymptotic spectra of tensors to understand the asymptotic restriction problem [Str86, Str88]. Deferring the details to the next subsection, this can be viewed as the theory of obstructions in the above sense. A remarkable result of this theory is that the asymptotic restriction problem for a family of tensors \( \mathcal{X} \) that is closed under direct sum and tensor product and contains the diagonal tensors \( \langle n \rangle \), reduces to finding all maps \( \mathcal{X} \to \mathbb{R}_{\geq 0} \) that are

(a) monotone under restriction \( \geq \)

(b) multiplicative under tensor product \( \otimes \)

(c) additive under direct sum \( \oplus \)

3
(d) normalised to have value $n$ at the unit tensor $\langle n \rangle$.

Such maps are called spectral points. Here the direct sum $f \oplus g$ is defined naturally as the multilinear map $(\mathbb{F}^{n_1} \oplus \mathbb{F}^{n_1}) \times \cdots \times (\mathbb{F}^{n_k} \oplus \mathbb{F}^{m_k}) \to \mathbb{F}$ such that

$$(v_1 + w_1, \ldots, v_k + w_k) \mapsto f(v_1, \ldots, v_k) + f(w_1, \ldots, w_k)$$

where $v_i, w_i \in \mathbb{F}^{n_i}$, and the unit tensor $\langle n \rangle$ for $n \in \mathbb{N}$ is defined as the multilinear map $(\mathbb{F}^n) \times k \to \mathbb{F}$ that maps $(e_1, \ldots, e_k)$ to 1 if $i_1 = \cdots = i_k$ and to 0 otherwise.

Properties (a) and (b) are natural properties to obtain an obstruction. Namely, suppose $\xi$ is such a map, and let $f, g \in \mathcal{X}$. If $f \succeq g$, then by definition $f \otimes n + o(n) \geq g \otimes n$, and (a) and (b) imply $\xi(f) + o(n) = \xi(f^n) \geq \xi(g^n) = \xi(g)$, which implies $\xi(f) \geq \xi(g)$. Turning this around, if $\xi(f) < \xi(g)$ then not $f \succeq g$, so $\xi$ yields an obstruction to $f \succeq g$.

Strassen in [Str91] created a family of spectral points. Let $\Theta$ be the set of all probability distributions on $\{1, 2, \ldots, k\}$. Strassen defined a family of maps $\zeta^\theta : \{k\text{-tensors}\} \to \mathbb{R}_{\geq 0}$ parametrised by $\theta \in \Theta$, named the support functionals, and he proved that the $\zeta^\theta$ are spectral points for the family $\mathcal{X}$ of oblique tensors, tensors whose support in some basis is an antichain, a strict (and nongeneric) subfamily of all tensors. In [Str05], for such tensors, $\zeta^\theta$ has been given a formulation in terms of moment polytopes.

**Our main result.** Universal spectral points are spectral points for the family of all tensors. The construction of nontrivial universal spectral points has been an open problem for more than thirty years. We introduce maps $F_\theta : \{\text{complex } k\text{-tensors}\} \to \mathbb{R}_{\geq 0}$ called the quantum functionals and we prove that they are universal spectral points. The quantum functionals are defined as follows (we will carefully define the quantum concepts later). For any $\theta \in \Theta$, define

$$F_\theta(t) = 2E_\theta(t)$$

$$E_\theta(t) = \sup_{g_i} \sum_{i=1}^k \theta(i) H(\text{Tr}_i (g_1, \ldots, g_k) \cdot t)$$

where the supremum goes over invertible maps $g_i \in \text{GL}(\mathbb{C}^{d_i})$. $H$ denotes the quantum entropy i.e. von Neumann entropy, and $\text{Tr}_i (g_1, \ldots, g_k) \cdot t$ denotes a partial trace of $(g_1, \ldots, g_k) \cdot t$ interpreted as a pure quantum state. To prove properties (a)–(d) we draw from the theory on invariants, quantum entropy, entanglement polytopes, Kronecker coefficients and Littlewood–Richardson coefficients. Let us briefly sketch the connection to moment polytopes, which in this context are called entanglement polytopes. Given a tensor $s$, let $r_1(s), \ldots, r_k(s)$ be the single-system quantum marginal entropies of the tensor $s$ viewed as a quantum state. Let $G$ be the group $\text{GL}(\mathbb{C}^{n_1}) \times \cdots \times \text{GL}(\mathbb{C}^{n_k})$. For a tensor $t$, define the set $\Delta_t = \{ (r_1(s), \ldots, r_k(s)) \mid s \in G \cdot t \}$, where $G \cdot t$ denotes the Euclidean closure (or equivalently Zariski closure) of the orbit $G \cdot t$. It is a nontrivial fact that $\Delta_t$ is a polytope, named the entanglement polytope. Then $E_\theta(t)$ equals the following optimisation over the entanglement polytope,

$$E_\theta(t) = \sup_{\lambda} \left\{ \sum_{i=1}^k \theta(i) H(\lambda^{(i)}) \mid \lambda^{(1)}, \ldots, \lambda^{(k)} \in \Delta_t \right\},$$

which is a convex optimisation problem.
The definition of $F_0$ can in fact be extended to probability distributions $\theta$ on subsets of $\{1, 2, \ldots, k\}$ in several ways. We explore these extensions and study the properties (a)–(d). One such extension, the lower quantum functional, satisfies (a) and (d) and is super-additive and super-multiplicative. Another extension, the upper quantum functional, satisfies (a) and (d) and is sub-additive and sub-multiplicative for $\theta$ that are what we call noncrossing. (The Strassen support functionals in fact similarly come in an upper and a lower version.)

**Asymptotic rank and asymptotic subrank.** In applications (e.g. the complexity of matrix multiplication) one is often interested in asymptotic restriction to or from a unit tensor. One commonly defines the tensor rank $R(f) = \min\{r \in \mathbb{N} \mid f \leq \langle r \rangle\}$, and the subrank $Q(f) = \max\{s \in \mathbb{N} \mid \langle s \rangle \leq f\}$, and asymptotically the asymptotic rank $\overline{R}(f) = \lim_{n \to \infty} R(f \otimes n)^{1/n}$ and the asymptotic subrank $\overline{Q}(f) = \lim_{n \to \infty} Q(f \otimes n)^{1/n}$. Clearly, maps that satisfy (a), (d) and super-multiplicativity are lower bounds on asymptotic rank, and maps that satisfy (a), (d) and sub-multiplicativity are upper bounds on asymptotic subrank.

Spectral points (maps satisfying (a)–(d)) are thus between asymptotic subrank and asymptotic rank. The defining expression of asymptotic rank does not suggest any algorithm for computing its value other than computing the rank for high tensor powers. In information theory, such an expression is called a multi-letter formula, and maps satisfying (a)–(d) are called single-letter formulas.

**Cap sets and slice rank.** To demonstrate an application of the asymptotic spectrum we go on a brief combinatorial excursion to the cap set problem. (Full details are in Section 4.2.) A subset $A \subseteq (\mathbb{Z}/3\mathbb{Z})^n$ is called a cap set if any line in $A$ is a point, a line being a triple of points of the form $(u, u + v, u + 2v)$. The *cap set problem* is to decide whether the maximal size of a cap set in $(\mathbb{Z}/3\mathbb{Z})^n$ grows like $3^{n-o(n)}$ or like $c^{n+o(n)}$ for some $c < 3$ when $n \to \infty$. Gijswijt and Ellenberg in [EG17], inspired by the work of Croot, Lev and Pach in [CLP17], settled this problem, showing that $c \leq 3(207 + 33\sqrt{33})^{1/3}/8 \approx 2.755$. Tao realised in [Tao16] that the cap set problem may naturally be phrased as the problem of computing the size of the largest main diagonal in powers of the *cap set tensor* $\sum_\alpha e_{\alpha_1} \otimes e_{\alpha_2} \otimes e_{\alpha_3}$ where the sum is over $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{Z}/3\mathbb{Z}$ with $\alpha_1 + \alpha_2 + \alpha_3 = 0$. Here main diagonal refers to a subset $A$ of the basis elements such that restricting to $A \times A \times A$ gives the tensor $\sum_{v \in A} v \otimes v \otimes v$. A main diagonal is essentially a unit tensor, so it is sufficient to upper bound the asymptotic subrank of the cap set tensor interpreted as a tensor over $\mathbb{F}_3$. We show (in hindsight) that the cap set tensor is in the $GL_3(\mathbb{F}_3) \times \mathbb{Z}$-orbit of the structure tensor of the algebra $\mathbb{F}_3[x]/(x^3)$. This implies that the asymptotic spectrum of the cap set tensor and the asymptotic spectrum of $\mathbb{F}_3[x]/(x^3)$ coincide. The tensor $\mathbb{F}_3[x]/(x^3)$ is the prime example in [Str91] for which Strassen computes the whole asymptotic spectrum. The minimal point in this asymptotic spectrum, which corresponds to the asymptotic subrank, is computed by Strassen to be $3(207 + 33\sqrt{33})^{1/3}/8 \approx 2.755$ (see [Str91, Table 1]). This reproves the bound by Ellenberg–Gijswijt.

Although our universal spectral points, the quantum functionals, do not play a role for cap sets (we cannot work over $\mathbb{C}$ for this problem), we think that a better understanding of our universal spectral points construction may lead to further progress on related combinatorial questions.

In the study of the cap set problem, slice rank and multi-slice rank were introduced as upper bounds on subrank [Tao16, Nas17]. We show that asymptotically the (upper) quantum functionals and the Strassen (upper) support functionals
are an upper bound on slice rank and multi-slice rank. As a consequence we prove that for the family of tight 3-tensors asymptotic subrank and asymptotic slice rank coincide. For complex tensors we characterise the asymptotic slice rank in terms of the quantum functionals.

1.2. Asymptotic spectra for tensors

To understand why we focus on maps satisfying the properties (a)–(d) we have to give a brief introduction to Strassen’s theory of asymptotic spectra for tensors.

We begin by putting an equivalence relation on tensors to get rid of trivialities. Restriction $\geq$ and asymptotic restriction $\gtrsim$ are both preorders (reflexive and transitive). We say $f$ is isomorphic to $g$, and write $f \cong g$, if there are bijective linear maps $A_i : F^{n_i} \to F^{n_i}$ such that $g = f \circ (A_1, \ldots, A_k)$. We say $f$ and $g$ are equivalent, and write $f \sim g$, if there are null maps $f_0 : F^{n_1} \times \cdots \times F^{n_k} \to F : x \mapsto 0$ and $g_0 : F^{b_1} \times \cdots \times F^{b_k} \to F : x \mapsto 0$ such that $f \oplus f_0 \cong g \oplus g_0$. The equivalence relation $\sim$ is in fact the equivalence relation generated by the restriction preorder $\geq$. Let $\mathcal{T}$ be the set of $\sim$-equivalence classes of $F$-multilinear maps of order $k$. Direct sum and tensor product naturally carry over to $\mathcal{T}$, and $\mathcal{T}$ becomes a semiring with additive unit $\langle 0 \rangle$ and multiplicative unit $\langle 1 \rangle$ (more precisely, the equivalence classes of those tensors). Restriction $\geq$ induces a partial order on $\mathcal{T}$ (reflexive, antisymmetric, transitive), and asymptotic restriction $\gtrsim$ induces a preorder on $\mathcal{T}$. Both behave well with respect to the semiring operations, and naturally $n \geq m$ if and only if $\langle n \rangle \geq \langle m \rangle$.

The theory of asymptotic spectra revolves around the following theorem proved by Strassen [Str86, Str88]. Given a topological space $\Delta$ we denote by $C^+(\Delta)$ the semiring of continuous maps $\Delta \to \mathbb{R}_{\geq 0}$.

**Theorem 1.1 (Spectral theorem).** For any semiring $\mathcal{X} \subseteq \mathcal{T}$ there are

1. a compact space $\Delta$
2. a homomorphism $\phi : \mathcal{X} \to C^+(\Delta)$ of semirings

such that $\phi(\mathcal{X})$ separates points and such that $a \preceq b$ if and only if $\phi(a) \leq \phi(b)$ pointwise on $\Delta$, and the pair $(\Delta, \phi)$ is essentially unique. We call $(\Delta, \phi)$ an asymptotic spectrum for $\mathcal{X}$. Explicitly, $(\Delta, \phi)$ may be taken as follows:

1. $\Delta = \{ \geq$-monotone homomorphisms $\mathcal{X} \to \mathbb{R}_{\geq 0} \}$.
2. $\phi : \mathcal{X} \to C^+(\Delta) : a \mapsto \hat{a}$ where $\hat{a} : \Delta \to \mathbb{R}_{\geq 0} : \xi \mapsto \xi(a)$.

We call this pair $(\Delta, \phi)$ the asymptotic spectrum of $\mathcal{X}$ and refer to it as $\Delta(\mathcal{X})$.

Theorem 1.1 is proved by a nontrivial reduction to the structure theory of Stone–Kadison–Dubois for a certain class of ordered rings (see [BS83]). We will not go into the details of this proof here, nor do we elaborate on how $(\Delta, \phi)$ is unique. We have taken the name spectral theorem from the talk [Str12].

**Remark 1.2.** We note that $\Delta(\mathcal{X})$ may equivalently be defined with degeneration $\supseteq$ instead of restriction $\geq$. Over $\mathbb{C}$, we say $f$ degenerates to $g$, written, $f \supseteq g$, if $g$ is in the Euclidean closure (or equivalently Zariski closure) of the orbit $\text{GL}(\mathbb{C}^{n_1}) \times \cdots \times \text{GL}(\mathbb{C}^{n_k}) : f$. It is a nontrivial fact from algebraic geometry (see [Kra84, Lemma III.2.3.1] or [BCS97]) that there is a degeneration $f \supseteq g$ if and only if there are linear maps $A_i \in \text{GL}(\mathbb{C}^{n_i})$ such that
We call \( \Delta(\phi) \) the asymptotic spectrum of \( X \), and write \( \Delta(X) \). When \( X = \{a\} \) we will denote \( \Delta(X) \) by \( \Delta(a) \).

We finish by stating the important relationship between the asymptotic spectrum and the asymptotic (sub)rank, which follows from Theorem 1.1 (see [Str88, Theorem 3.8]).
Proposition 1.6. Let $\mathcal{X} \subseteq \mathcal{T}$ be a semiring. Let $(\Delta, \phi)$ be an asymptotic spectrum of $\mathcal{X}$. Let $a \in \mathcal{X}$. Then

$$Q(a) = \min_{\xi \in \Delta} \phi(a)(\xi) \quad R(a) = \max_{\xi \in \Delta} \phi(a)(\xi).$$

Remark 1.7. Obviously $R$ and $Q$ are $\geq$-monotones and have value $n$ on $\langle n \rangle$. They are not universal spectral points however. Namely, the asymptotic rank of $e_1 \otimes e_1 \otimes 1 + e_2 \otimes e_2 \otimes 1$, $e_1 \otimes 1 \otimes e_1 + e_2 \otimes 1 \otimes e_2$ and $1 \otimes e_1 \otimes e_1 + 1 \otimes e_2 \otimes e_2$ is 2, whereas the tensor product equals the matrix multiplication tensor whose asymptotic rank is strictly smaller than $2^3$. With the same tensors one shows that asymptotic subrank is not multiplicative.

1.3. This paper

Main results. The main results of this paper are summarised as follows.

1. We construct for the first time a nontrivial family of maps

$$F_\theta : \{k\text{-tensors}\} \to \mathbb{R}_{\geq 0},$$

named quantum functionals, that are nonincreasing under restriction $\geq$, normalised on the unit tensor $\langle r \rangle$, additive under direct sum $\oplus$ and multiplicative under tensor product $\otimes$, i.e. universal spectral points, advancing Strassen’s theory of asymptotic spectra.

2. We connect Strassen’s asymptotic spectra to entanglement polytopes and the quantum marginal problem. This is to our knowledge the first time that information about asymptotic transformations is obtained from entanglement polytopes, as opposed to information about single-copy transformations.

3. We put recent progress on the cap set problem in the framework of Strassen, which may prove useful in solving variations on the cap set problem. We characterise asymptotic slice rank in terms of the quantum functionals and show that asymptotic slice rank coincides with asymptotic subrank for tight 3-tensors.

Paper overview. In Section 2 we set the scene by reviewing the theory of the Strassen upper and lower support functionals $\zeta^\theta$ and $\zeta^\theta$ introduced in [Str91]. This is important background material for our work. The support functionals allow us to compute elements in the asymptotic spectrum for a class of tensors called oblique tensors. Our exposition follows closely the exposition in [Str91] with the exception that we consider multilinear maps $V_1 \times \cdots \times V_k \to F$ with $k \geq 3$ instead of bilinear maps $V_1 \times V_2 \to V_3$. Strassen already observed that his results generalise to the multilinear regime, but we will make this explicit for the sake of understanding and comparison to our results.

Section 3 contains our main result. Here we introduce two new families of functionals over the complex numbers. These functionals are denoted by $F^\theta$ and $F_\theta$ and called the upper and lower quantum functionals. Both functionals are $\geq$-monotone. The upper quantum functional $F^\theta$ is sub-additive and sub-multiplicative when $\theta$ is what we call noncrossing. The lower quantum functional $F_\theta$ is super-additive and super-multiplicative. For singleton $\theta$ the two
functionals coincide and thus yield an additive, multiplicative \( \geq \)-monotone, a universal spectral point.

In Section 4 we consider several families of tensors. We introduce the family of free tensors. We show that for free tensors, the Strassen upper support functional coincides with the quantum functionals for singleton \( \theta \). This is useful computationally, since the upper support functional is defined as a minimisation while the quantum functionals are defined as a maximisation. We next compute generic values of the quantum functionals for tensor formats in which a quantum state with completely mixed marginals exists. Our results extend over \( \mathbb{C} \) the results by Verena Töbler [Tob91] on the generic value of the upper support functional. Finally, we reprove recent results on the cap set problem, by reducing this problem to a result of Strassen on reduced polynomial multiplication.

In Section 5 we show that asymptotically the Strassen upper support functional and the upper quantum functional are upper bounds for slice rank and multi-slice rank, recently introduced \( \geq \)-monotones that are neither super-multiplicative nor sub-multiplicative. As a consequence, we find that slice rank coincides asymptotically with subrank for tight 3-tensors. We additionally show that the quantum functionals characterise asymptotic slice rank for complex tensors.

2. Strassen support functionals

In this section the field \( \mathbb{F} \) is arbitrary. After having introduced the concept of the asymptotic spectrum of tensors in [Str88], Strassen constructed a nontrivial family of spectral points in the asymptotic spectrum of oblique tensors in [Str91]. Oblique tensors are tensors for which the support is an antichain in some basis. In plain English, Strassen constructed a family of functions from \( k \)-tensors to \( \mathbb{R} \) that are \( \geq \)-monotone, are normalised to attain value \( r \) at \( \langle r \rangle \), and, when restricted to oblique tensors, are additive under \( \oplus \) and multiplicative under \( \otimes \). He calls his functions the support functionals, because the support of a tensor plays an important role in the definition. Not all tensors are oblique, and obliqueness is not a generic property. Many tensors that are of interest in the field of algebraic complexity theory, however, turn out to be oblique, notably the structure tensor of the algebra of \( n \times n \) matrices. This section is devoted to explaining the construction of these spectral points for oblique tensors. We do this not only to set the scene and provide benchmarks for our new functionals, but also because the Strassen support functionals remain relevant today, for example in the context of combinatorial problems like the cap set problem, as we will make clear in Section 4.2.2.

The construction goes in four steps. First we define for probability distributions \( \theta \) on \( \{1, 2, \ldots, k\} \) the upper support functional \( \zeta^\theta \), which is \( \geq \)-monotone, normalised at \( \langle r \rangle \), sub-additive under \( \oplus \) and sub-multiplicative under \( \otimes \). Second we define the lower support functional \( \zeta_\theta \), which is \( \geq \)-monotone, normalised at \( \langle r \rangle \), super-additive under \( \oplus \) and super-multiplicative under \( \otimes \). Third we show that \( \zeta^\theta(t) \geq \zeta_\theta(t) \) for any \( k \)-tensor \( t \). Fourth we show that for oblique tensors \( t \) the upper support functional \( \zeta^\theta(t) \) and lower support functional \( \zeta_\theta(t) \) coincide, which gives a spectral point.

**Notation.** We use the following standard notation. For any natural number \( n \) we write \( [n] \) for the set \( \{1, 2, \ldots, n\} \). For any finite set \( X \), let \( \mathcal{P}(X) \) be the set
of all probability distributions on $X$. For any probability distribution $P \in \mathcal{P}(X)$ the Shannon entropy of $P$ is defined as $H(P) = -\sum_{x \in X} P(x) \log_2 P(x)$ with $0 \log_2 0$ understood as 0. The support of $P$ is supp $P = \{ x \in X \mid P(x) \neq 0 \}$. Given finite sets $I_1, \ldots, I_k$ and a probability distribution $P \in \mathcal{P}(I_1 \times \cdots \times I_k)$ on the product set $I_1 \times \cdots \times I_k$ we denote the marginal distribution of $P$ on $I_i$ by $P_i$, that is, $P_i(a) = \sum_{x: x_i = a} P(x)$ for any $a \in I_i$.

We set some specific notation for this text. Let $k \geq 3$. For $i \in [k]$, let $V_i$ be a vector space over $\mathbb{F}$ of dimension $n_i$. Let $I_i$ be a finite set. Let $f : V_1 \times \cdots \times V_k \to \mathbb{F}$ be a multilinear map. We order the set $[n_i]$ naturally by $1 < 2 < \cdots < n_i$. When the sets $I_1, \ldots, I_k$ are ordered, we give $I_1 \times \cdots \times I_k$ the product order, defined by $x \leq y$ iff for all $i \in [k]$: $x_i \leq y_i$.

### 2.1. Upper support functional

We begin with introducing the upper support functional. We naturally define the notion of the support of a multilinear map $V_1 \times \cdots \times V_k \to \mathbb{F}$.

**Definition 2.1.** Let $f : V_1 \times \cdots \times V_k \to \mathbb{F}$ be a multilinear map. Let $C(f)$ denote the set of $k$-tuples of bases for $V_1, \ldots, V_k$. That is, each element $C \in C(f)$ is a $k$-tuple $((v_{1,1}, v_{1,2}, \ldots, v_{1,n_1}), \ldots, (v_{k,1}, v_{k,2}, \ldots, v_{k,n_k}))$ in which $(v_{i,1}, v_{i,2}, \ldots, v_{i,n_i})$ is a basis of $V_i$. Define the support of $f$ with respect to $C$ as

$$\text{supp}_C f := \{(a_1, \ldots, a_k) \in [n_1] \times \cdots \times [n_k] \mid f(v_{1,a_1}, \ldots, v_{k,a_k}) \neq 0 \}.$$

**Definition 2.2.** Let $\theta \in \mathcal{P}([k])$. Let $P \in \mathcal{P}(I_1 \times \cdots \times I_k)$ be a probability distribution. Define $H_\theta(P)$ as the $\theta$-weighted average of the Shannon entropies of the marginal distributions $P_i$ of $P$,

$$H_\theta(P) = \sum_{i \in [k]} \theta(i) H(P_i).$$

Let $\Phi \subseteq I_1 \times \cdots \times I_k$ be a nonempty subset. Define $H_\theta(\Phi)$ as the maximum $H_\theta(P)$ over all probability distributions $P \in \mathcal{P}(\Phi)$,

$$H_\theta(\Phi) = \max_{P \in \mathcal{P}(\Phi)} H_\theta(P). \quad (1)$$

**Definition 2.3** (Strassen upper support functional). Let $\theta \in \mathcal{P}([k])$. Let $f$ be a nonzero multilinear map. Define $\rho^\theta(f)$ by

$$\rho^\theta(f) = \min_{C \in C(f)} H_\theta(\text{supp}_C f)$$

and define $\zeta^\theta$ by

$$\zeta^\theta(f) = 2^{\rho^\theta(f)}.$$

We call $\zeta^\theta$ the Strassen upper support functional and we call $\rho^\theta$ the logarithmic Strassen upper support functional. When $f$ is 0 we naturally define $\zeta^\theta(f) = 0$ and $\rho^\theta(f) = -\infty$.

The key properties of the upper support functional are as follows.
Theorem 2.4 ([Str91]). Let \( f : V_1 \times \cdots \times V_k \to F \) and \( g : W_1 \times \cdots \times W_k \to F \) be multilinear maps. Let \( \theta \in \mathcal{P}(\{k\}) \). The Strassen upper support functional \( \zeta^\theta \) has the following properties.

1. \( \zeta^\theta(r) = r \) for \( r \in \mathbb{N} \).
2. \( \zeta^\theta(f \oplus g) = \zeta^\theta(f) + \zeta^\theta(g) \).
3. \( \zeta^\theta(f \otimes g) \leq \zeta^\theta(f) \zeta^\theta(g) \).
4. If \( f \geq g \), then \( \zeta^\theta(f) \geq \zeta^\theta(g) \).
5. \( 0 \leq \zeta^\theta(f) \leq (\dim V_1)^{\theta(1)} \cdots (\dim V_k)^{\theta(k)} \).

(Note that \( \zeta^\theta \) is not just sub-additive, but even additive.) One verifies directly that statement 1 and 5 of Theorem 2.4 are true. Statement 2–4 can be proved with generalisations of the arguments in [Str91, Section 2].

In the rest of this section we discuss an important alternative characterisation of the upper support functional in terms of filtrations. This alternative definition we need later to relate the upper and lower support functionals.

Definition 2.5. A proper filtration or complete flag of a vector space \( U \) is a sequence \((U_\alpha)_{1 \leq \alpha \leq a}\) of subspaces of \( U \) with \( U = U_1 \supseteq \cdots \supseteq U_a = \{0\} \) and \( \dim U_\alpha/U_{\alpha+1} = 1 \) for \( 1 \leq \alpha < a \). (In this text all flags will be decreasing.) We denote by \( \mathcal{F}_\ell(f) \) the set of \( k \)-tuples of complete flags \( F = ((V_1, \alpha_1), \ldots, (V_k, \alpha_k)) \) of \( V_1, \ldots, V_k \). For \( F \in \mathcal{F}_\ell(f) \) we define the support of \( f \) with respect to \( F \) as

\[
\text{supp}_F f = \{(\alpha_1, \ldots, \alpha_k) \mid f(V_{1, \alpha_1} \times \cdots \times V_{k, \alpha_k}) \neq 0\} \subseteq [n_1] \times \cdots \times [n_k].
\]

The alternative characterisation of the upper support functional in terms of filtrations is as follows. We refer to [Str91, Section 2] for the proof.

Proposition 2.6. \( \rho^\theta(f) = \min_{F \in \mathcal{F}_\ell(f)} H_\theta(\text{supp}_F f) \).

Remark 2.7. The characterisation of \( \rho^\theta \) as a minimisation over filtrations can be thought of as an asymptotic version of the filtration method explained in [BCS97, Exercise 15.13].

2.2. Lower support functional

The upper support functional \( \zeta^\theta \) has a companion called the lower support functional \( \zeta_\varnothing \) which we introduce here. We shift focus from the support to the maximal points in the support.

Definition 2.8. Let \( \Phi \subseteq [n_1] \times \cdots \times [n_k] \) be a subset. We define the maximal points of \( \Phi \) with respect to the product order of the natural orders on \([n_1], \ldots, [n_k]\) as

\[
\max \Phi = \{\alpha \in \Phi \mid \nexists \beta \in \Phi : \alpha < \beta\}.
\]

Definition 2.9. Let \( f \) be a nonzero multilinear map. Given \( C \in \mathcal{C}(f) \) we set the notation \( M_C f = \max \text{supp}_C f \). Define \( \rho_\varnothing(f) \) by

\[
\rho_\varnothing(f) = \max_{C \in \mathcal{C}(f)} H_\theta(M_C f).
\]
and define $\zeta_\theta(f)$ by

$$\zeta_\theta(f) = 2^{\rho_\theta(f)}.$$ 

We call $\zeta_\theta$ the \textit{Strassen lower support functional} and we call $\rho_\theta$ the \textit{logarithmic Strassen lower support functional}. When $f$ is 0 we naturally define $\zeta_\theta(f) = 0$ and $\rho_\theta(f) = -\infty$.

\textbf{Theorem 2.10} ([Str91]). Let $f : V_1 \times \cdots \times V_k \to F$ and $g : W_1 \times \cdots \times W_k \to F$. Let $\theta \in \mathcal{P}([k])$. The Strassen lower support functional $\zeta_\theta$ has the following properties.

1. $\zeta_\theta((r)) = r$ for $r \in \mathbb{N}$.
2. $\zeta_\theta(f \oplus g) \leq \zeta_\theta(f) + \zeta_\theta(g)$.
3. $\zeta_\theta(f \otimes g) \leq \zeta_\theta(f) \zeta_\theta(g)$.
4. If $f \geq g$, then $\zeta_\theta(f) \leq \zeta_\theta(g)$.
5. $0 \leq \zeta_\theta(f) \leq (\dim V_1)^{\theta(1)} \cdots (\dim V_k)^{\theta(k)}$.

\textbf{Remark 2.11.} Regarding statement 2 in Theorem 2.10, Bürgisser [Bü90] shows that the lower support functional $\zeta_\theta$ is not in general additive under the direct sum when $\theta_i > 0$ for all $i$. See also [Str91, Comment (iii)]. In particular, this implies that the upper support functional $\zeta^\theta(f)$ and the lower support functional $\zeta_\theta(f)$ are not equal in general, the upper support functional being additive. In fact, to show that the lower support functional is not additive, Bürgisser first shows that (with $F$ algebraically closed) the \textit{typical value} of $\zeta_\theta$ on $F^n \otimes F^n \otimes F^n$ equals $(1 - \min \theta_i) \log_2 n + o(n)$; on the other hand, Tobler [Tob91] shows that the typical value of $\zeta^\theta$ on $F^n \otimes F^n \otimes F^n$ equals $\log_2 n$. (So even generically $\zeta^\theta$ and $\zeta_\theta$ are different on $F^n \otimes F^n \otimes F^n$.) We discuss typical values in Section 4.4.

One verifies directly that statement 1 and 5 of Theorem 2.10 are true. For the proofs of statements 2–4 we refer to [Str91, Section 3].

We finish this section by giving an alternative definition of $F$ in terms of complete flags, in the same spirit as Proposition 2.6 for the upper support functional. Let us first make some general remarks about how (the maximal points in) the support with respect to a basis and the support with respect to a functional. Let us first make some general remarks about how (the maximal points in) the support with respect to a basis and the support with respect to a functional.

\textbf{Definition 2.12.} Let $\Phi \subseteq [n_1] \times \cdots \times [n_k]$ be a subset. We define the \textit{downward closure} of $\Phi$ with respect to the product of the natural orders on $[n_1], \ldots, [n_k]$ as

$$\downarrow \Phi = \{ \alpha \in [n_1] \times \cdots \times [n_k] \mid \exists \beta \in \Phi : \alpha \leq \beta \}.$$ 

To any $C = ((v_{1,\alpha_1}), \ldots, (v_{k,\alpha_k})) \in \mathcal{C}(f)$ we can naturally associate a $k$-tuple of complete flags $F = ((V_{1,\alpha_1}), \ldots, (V_{k,\alpha_k})) \in \mathcal{F}_c(f)$, by defining the subspace $V_{i,\alpha_i}$ as $\text{Span}\{v_{1,\alpha_1}, v_{1,\alpha_{i+1}}, \ldots\}$.

The following key properties of the downward closure and the maximal points follow directly from the definitions.

\textbf{Lemma 2.13.} Let $C \in \mathcal{C}(f)$ and $F \in \mathcal{F}_c(f)$ be associated. Then

$$\downarrow \supp_C f = \supp_F f,$$

$$\max \supp_C f = \max \supp_F f.$$
For $C \in \mathcal{C}(f)$ we have set the notation $M_C f = \max \text{supp}_C f$. For $F \in \mathcal{F}(f)$ we similarly set the notation $M_F f = \max \text{supp}_F f$.

**Proposition 2.14.** $\rho_\theta(f) = \max_{F \in \mathcal{F}_c(f)} H_\theta(M_F f)$ for $\theta \in \mathcal{P}([k])$.

**Proof.** Let $C \in \mathcal{C}(f)$ and $F \in \mathcal{F}_c(f)$ be associated. Then $M_C f = M_F f$ by Lemma 2.13 and hence $H_\theta(M_C f) = H_\theta(M_F f)$. Since any $C$ is associated to some $F$ and vice versa we conclude that

$$\max_{C \in \mathcal{C}(f)} H_\theta(M_C f) = \max_{F \in \mathcal{F}_c(f)} H_\theta(M_F f). \quad \square$$

### 2.3. Comparing the support functionals

Strassen calls his functionals *upper* and *lower* because he proves that the upper support functional is at least the lower support functional [Str91, Corollary 4.3].

**Theorem 2.15.** $\rho^\theta(f) \geq \rho_\theta(f)$.

This is a useful property when doing computations, since $\rho^\theta$ is defined as a minimisation and $\rho_\theta$ is defined as a maximisation.

In fact, in [Str91, Corollary 4.2] Strassen first proves $\zeta^\theta(f \otimes g) \geq \zeta^\theta(f) \zeta^\theta(g)$ and then proves Theorem 2.15 as a corollary. The purpose of this section is to give a more direct proof of Theorem 2.15 for the benefit of the reader. The proof essentially comes down to the following proposition.

**Proposition 2.16.** Let $f : V_1 \otimes \cdots \otimes V_k \to \mathbb{F}$. Let $F, G \in \mathcal{F}_c(f)$. There are permutations $\phi_i : [n_i] \to [n_i]$ ($i \in [k]$) such that $(\phi_1 \times \cdots \times \phi_k) M_G f \subseteq \text{supp}_F f$.

**Proof.** Let $V$ be some $n$-dimensional vector space and let $(V_\alpha)_\alpha$ and $(W_\beta)_\beta$ be complete flags of $V$ such that

$$V_\alpha = \text{Span}\{v_\alpha, v_{\alpha+1}, \ldots\}$$
$$W_\beta = \text{Span}\{w_\beta, w_{\beta+1}, \ldots\}.$$

Define the map

$$\phi : [n] \to [n] : \beta \mapsto \max\{\alpha \in [n] : V_\alpha \cap (w_\beta + W_{\beta+1}) \neq \emptyset\}. \quad (2)$$

The map $\phi$ is injective. Let $\beta, \gamma \in [n]$ with $\beta \leq \gamma$ and suppose $\alpha = \phi(\beta) = \phi(\gamma)$. Then (2) gives

$$\mathbb{P}^x v_\alpha + V_{\alpha+1} \cap (w_\beta + W_{\beta+1}) \neq \emptyset \quad (3)$$
$$V_{\alpha+1} \cap (w_\beta + W_{\beta+1}) = \emptyset \quad (4)$$
$$\mathbb{P}^x v_\alpha + V_{\alpha+1} \cap (w_\gamma + W_{\gamma+1}) \neq \emptyset \quad (5)$$

If $\beta < \gamma$, then using (3) and (5) we may obtain a contradiction to (4). We conclude that $\beta = \gamma$.

We turn to $F$ and $G$. Write

$$F = ((V_{1,\alpha_1}), \ldots, (V_{k,\alpha_k}))$$
$$G = ((W_{1,\beta_1}), \ldots, (W_{k,\beta_k}))$$

13
and for each pair of complete flags \((V_i, \alpha_i), (W_i, \beta_i)\) of \(V_i\) we define the permutation \(\phi_i : [n_i] \rightarrow [n_i]\) in the same way as above. We prove \((\phi_1 \times \cdots \times \phi_k) M_G f \subseteq \text{supp}_F f\). Let \((\beta_1, \ldots, \beta_k)\) be in \(M_G f\). Let \((\alpha_1, \ldots, \alpha_k) = (\phi_1 \times \cdots \times \phi_k)(\beta_1, \ldots, \beta_k)\). Then by definition of the \(\phi_i\) the intersection \(V_i, \alpha_i \cap (w_i, \beta_i + 1)\) is not empty. Choose

\[ \begin{array}{c}
    w_i, \beta_i \in V_i, \alpha_i \cap (w_i, \beta_i + 1).
\end{array} \]

Since \(f\) is multilinear we have for some \(x_{\beta'} \in \mathbb{F}\)

\[ f(w_1, \beta_1, \ldots, w_k, \beta_k) = f(w_1, \beta_1, \ldots, w_k, \beta_k) + \sum_{\beta' > \beta} x_{\beta'} f(w_1, \beta'_1, \ldots, w_k, \beta'_k) \]

with the sum over tuples \(\beta' = (\beta'_1, \ldots, \beta'_k)\) that are strictly larger than \(\beta\) in the product order. Since \(\beta\) is a maximal element in \(\text{supp}_G f\) with respect to the product order, the sum over \(\beta' > \beta\) equals zero. We conclude that \(f(w_1, \beta_1, \ldots, w_k, \beta_k) \neq 0\). Therefore, \(f(V_1, \alpha_1 \times \cdots \times V_k, \alpha_k)\) is not zero and thus \((\alpha_1, \ldots, \alpha_k) \in \text{supp}_F f\).

The above proof of Proposition 2.16 may more naturally be phrased in the language of Schubert cells, when \(\mathbb{F} = \mathbb{C}\), as follows. We use the notation and definitions of [Bri05] with the difference that our complete flags are decreasing instead of increasing so we need to use the group of lower unitriangular matrices \(U^-\) instead of the group of upper unitriangular matrices \(U\).

Proof. Let \(U^-_n \subseteq \text{GL}(\mathbb{C}^n)\) be the unipotent subgroup of lower triangular \(n \times n\) matrices with ones on the diagonal. Let \(Y\) be the variety of complete (decreasing) flags in \(\mathbb{C}^n\). The group \(\text{GL}(\mathbb{C}^n)\) naturally acts on \(Y\). Let the symmetric group \(S_n\) act on \(Y\) via its natural action on the standard basis of \(\mathbb{C}^n\). It is well-known (see e.g. [Bri05, Proposition 1.2.1]) that the variety \(Y\) is the disjoint union of the \(U^-_n\)-orbits \(C^\phi := U^-_n \phi x\) over all permutations \(\phi\) in the symmetric group \(S_n\), where \(x\) is the standard flag \(\text{Span}\{e_1, \ldots, e_n\} \supset \text{Span}\{e_2, \ldots, e_n\} \supset \cdots \supset \text{Span}\{e_n\}\) in \(\mathbb{C}^n\). These orbits are called the Schubert cells.

We turn to the \(k\)-tuples of flags \(F\) and \(G\). Write again

\[ F = ((V_1, \alpha_1), \ldots, (V_k, \alpha_k)) \]
\[ G = ((W_1, \beta_1), \ldots, (W_k, \beta_k)). \]

Without loss of generality we may assume that each \((W_i, \beta_i)\) is the standard flag in \(\mathbb{C}^{n_i}\). By the previous paragraph, for each \(i \in [k]\) there is a permutation \(\phi_i \in S_{n_i}\) such that the flag \((V_i, \phi_i(\beta_i))\) is in the \(U^-_{n_i}\)-orbit of the flag \((W_i, \beta_i)\), that is, there is a group element \(g_i \in U^-_{n_i}\) such that \(V_i, \phi_i(\beta_i) = g_i W_i, \beta_i\) for every \(\beta_i \in [n_i]\). Let \((\beta_1, \ldots, \beta_k)\) be in \(\text{supp}_G f\), so

\[ f(W_1, \beta_1 \times \cdots \times W_k, \beta_k) \neq \{0\}. \]

Suppose \((\beta_1, \ldots, \beta_k)\) is maximal. Then since \(g_i \in U^-_{n_i}\) and since \(f\) is multilinear,

\[ f(W_1, \beta_1 \times \cdots \times W_k, \beta_k) = f(g_1 W_1, \beta_1 \times \cdots \times g_k W_k, \beta_k). \]

By our choice of \(g_i\),

\[ f(g_1 W_1, \beta_1 \times \cdots \times g_k W_k, \beta_k) = f(V_1, \phi_1(\beta_1) \times \cdots \times V_k, \phi_k(\beta_k)). \]

We conclude that \((\phi_1(\beta_1), \ldots, \phi_k(\beta_k))\) is in \(\text{supp}_F f\).
Proof of Theorem 2.15. By Proposition 2.6 and Proposition 2.14 we have

\[
\rho^\theta(f) = \min_{F \in \mathcal{F}_C(f)} H_\theta(\text{supp}_F f), \quad (6)
\]

\[
\rho_\theta(f) = \max_{G \in \mathcal{F}_C(f)} H_\theta(M_G f).
\]

Let \( F \in \mathcal{F}_C(f) \) minimise (6) and let \( G \in \mathcal{F}_C(f) \) maximise (7). It suffices to show

\[
H_\theta(M_G f) \leq H_\theta(\text{supp}_F f).
\]

Let \( P \in \mathcal{P}(M_G f) \) such that \( H_\theta(P) = H_\theta(M_G f) \). Proposition 2.16 gives permutations \( \phi_i : [n_i] \to [n_i] \) such that

\[
(\phi_1 \times \cdots \times \phi_k)M_G f \subseteq \text{supp}_F f.
\]

Let \( Q \in \mathcal{P}(\text{supp}_F f) \) be defined by \( Q(\alpha) = P((\phi_1^{-1} \times \cdots \times \phi_k^{-1})(\alpha)) \). Then \( H_\theta(P) = H_\theta(Q) \). This proves (8). \( \square \)

2.4. Robust and oblique tensors; spectral points

At this point we have the upper support functional \( \zeta^\theta \), which is additive and sub-multiplicative, we have the lower support functional \( \zeta_\theta \), which is super-additive and super-multiplicative, and we know that \( \zeta^\theta(t) \geq \zeta_\theta(t) \) for any \( t \). When the functionals \( \zeta^\theta \) and \( \zeta_\theta \) coincide we have an additive and multiplicative \( \leq \)-monotone, a spectral point. In general (in fact generically for certain formats) the functionals \( \zeta^\theta \) and \( \zeta_\theta \) do not coincide, however (Remark 2.11). To get spectral points from the support functionals we should thus restrict to smaller families of tensors \( X \subseteq \mathcal{T} \). The following definition just gives a name to those tensors for which the support functionals do coincide.

Definition 2.17 (Robust). For \( \theta \in \mathcal{P}([k]) \) we say \( f \) is \( \theta \)-robust if \( \zeta^\theta(f) = \zeta_\theta(f) \). We say \( f \) is robust if \( f \) is \( \theta \)-robust for all \( \theta \in \mathcal{P}([k]) \).

We try to understand what robust tensors look like. Let \( f \) be nonzero. Clearly \( f \) is \( \theta \)-robust if and only if

\[
\zeta^\theta(f) \leq \zeta_\theta(f)
\]

Being \( \theta \)-robust is closed under \( \oplus \) and \( \otimes \), since \( \zeta^\theta(f \oplus g) = \zeta^\theta(f) + \zeta^\theta(g) = \zeta_\theta(f) + \zeta_\theta(g) \leq \zeta_\theta(f \oplus g) \), and \( \zeta^\theta(f \otimes g) \leq \zeta^\theta(f) \zeta^\theta(g) = \zeta_\theta(f) \zeta_\theta(g) \leq \zeta_\theta(f \otimes g) \). Equation (9) means precisely that there exist \( C, D \in \mathcal{C}(f) \) and \( P \in \mathcal{P}(M_D f) \) such that

\[
H_\theta(\text{supp}_C f) \leq H_\theta(P).
\]

In this case we have \( \zeta_\theta(f) = \zeta^\theta(f) = 2^{H_\theta(P)} \). In particular, \( f \) is \( \theta \)-robust if there is a \( C \in \mathcal{C}(f) \) such that the maximisation \( H_\theta(\text{supp}_C f) \) is attained by a probability distribution \( P \in \mathcal{P}(M_C f) \). This criterion is automatically satisfied for all \( \theta \) when in some basis the support of \( f \) equals the maximal points in the support of \( f \). Recall that a subset \( \Phi \) of a partially ordered set is called an antichain if all elements in \( \Phi \) are pairwise incomparable.

Definition 2.18 (Oblique). We say \( f \) is oblique if there exists a \( C \in \mathcal{C}(f) \) such that \( \text{supp}_C f = M_C f \), that is, such that \( \text{supp}_C f \) is an antichain.
Note that the family of oblique tensors is a semiring under $\oplus$ and $\otimes$. Suppose $f$ is oblique. Let $\text{supp}_C f$ be an antichain. Then $f$ is robust and we have simply $\rho_0(f) = \rho^\theta(f) = H_\theta(\text{supp}_C f)$. In the language of asymptotic spectra this means the following.

Theorem 2.19 (Strassen). Let $\mathcal{X} \subseteq \mathcal{T}$ be the family of all oblique $k$-tensors. Let $\theta \in \mathcal{P}([k])$. Given $f \in \mathcal{X}$, let $C_f \in \mathcal{C}(f)$ be any bases in which $\text{supp}_C f$ is an antichain. Then

$$\zeta^\theta(f) = \zeta_\theta(f) = 2^{H_\theta(\text{supp}_C f)}$$

and thus $\zeta^\theta : \mathcal{X} \rightarrow \mathbb{R}_{\geq 0}$ is a restriction monotone, that is normalised on $\langle r \rangle$, additive under direct sum $\oplus$ and multiplicative under tensor product $\otimes$. So $\zeta^\theta$ is an element of the asymptotic spectrum $\Delta(\mathcal{X})$.

Definition 2.20 (Support simplex). For any family of oblique tensors $\mathcal{Y}$, we call the image of $\mathcal{P}([k]) \rightarrow \Delta(\mathcal{Y}) : \theta \rightarrow \zeta_\theta$ the support simplex in the asymptotic spectrum of $\mathcal{Y}$. We denote the support simplex by $\zeta_{\mathcal{P}([k])}(\mathcal{Y})$.

We stress again that the lower support functionals are not additive and can thus not be universal spectral points. The upper support functionals may be universal spectral points, but this can, however, not be shown with the help of the lower support functionals.

3. Quantum functionals

In this section we let the base field $F$ be the complex numbers $\mathbb{C}$. The goal of this section is to construct an explicit family of universal spectral points in the asymptotic spectrum of tensors $\Delta(\mathcal{T})$. In other words, we construct maps $\{\text{complex } k\text{-tensors}\} \rightarrow \mathbb{R}_{\geq 0}$ that are $\geq$-monotone, multiplicative under $\otimes$, additive under $\oplus$ and normalised on the unit tensor $\langle n \rangle$. (In fact our maps will be monotone under degeneration $\geq$, cf. Remark 1.2.)

To achieve this we introduce two families of functionals. The upper quantum functional $F^\theta$ is defined in terms of isotypic projections; the lower quantum functional $F_\theta$ is defined in terms of quantum entropy. Recall that the Strassen support functionals $\zeta^\theta$ and $\zeta_\theta$ are parametrised by the probability distributions $\theta$ on $[k]$, in other words by the probability distributions $\theta$ on bipartitions of $[k]$ of the form $\{\{j\}, [k] \setminus \{j\}\}$. Our functionals $F^\theta$ and $F_\theta$ are parametrised by the probability distributions $\theta$ on all bipartitions $\{S, [k] \setminus S\}$ of $[k]$. We show that the upper functional is always at least the lower functional, $F^\theta(t) \geq F_\theta(t)$. When $\theta$ is supported on bipartitions of the form $\{\{j\}, [k] \setminus \{j\}\}$ we show that $F^\theta(t)$ and $F_\theta(t)$ are equal and are thus universal spectral points, and moreover $\zeta^\theta(t) \geq F^\theta(t)$. Finally, we show that the regularised upper support functional equals the upper quantum functional, $\lim_{n \rightarrow \infty} \zeta^\theta(t^{\otimes n})^{1/n} = F^\theta(t)$, which implies that $F^\theta(t) \geq \zeta_\theta(t)$.

Notation. As always $[k]$ denotes the set $\{1, 2, \ldots, k\}$. For $S \subseteq [k]$ we define the complement $\overline{S} = [k] \setminus S$ and for $j \in [k]$ we define the complement $j = [k] \setminus \{j\}$. As before, $V_1, \ldots, V_k$ are vector spaces over the base field, which is the complex numbers in this section. For a subset $S \subseteq [k]$ we define $V_S = \bigotimes_{i \in S} V_i$.

A partition of $[k]$ is an unordered pair $\{S, \overline{S}\}$ with $S \subseteq [k]$, $\overline{S} \neq \emptyset$ and $S \neq \emptyset$. We say that the bipartitions $\{S, \overline{S}\}$ and $\{T, \overline{T}\}$ are noncrossing if $S \subseteq T$ or $T \subseteq S$ or $S \cap T = \emptyset$. The set of bipartitions of $[k]$ is denoted by $\mathcal{B}$. We will
consider two subsets of the set $\mathcal{P}(B)$ of probability distributions on $B$. Let $\mathcal{P}_s(B)$ be the set of distributions supported on bipartitions of the form $\{\{j\}, \bar{j}\}$. We say a distribution $\theta \in \mathcal{P}(B)$ is noncrossing if $\theta$ is supported on pairwise noncrossing partitions. Let $\mathcal{P}_{nc}(B)$ be the set of noncrossing distributions. Note that the latter is not a convex set in general. Clearly $\mathcal{P}_s(B) \subseteq \mathcal{P}_{nc}(B) \subseteq \mathcal{P}(B)$, with equality if and only if $k \leq 3$.

An integer partition is a sequence $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_d)$ of nonnegative integers satisfying $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_d$. We say that $\lambda$ is a partition of $n$ and write $\lambda \vdash n$ if $\lambda_1 + \cdots + \lambda_d = n$. We denote the number of nonzero parts of $\lambda$ by $\ell(\lambda)$. We write $\lambda \vdash d$ if $\lambda \vdash n$ and $\ell(\lambda) \leq d$. Let $\mathcal{P}_n$ be the set of partitions of $n$. For a partition $\lambda \vdash n$ let $\overline{\lambda}$ be the sequence $(\frac{n}{\lambda_1}, \frac{n}{\lambda_2}, \ldots, \frac{n}{\lambda_d})$. For $e \geq 0$ we may view $\overline{\lambda}$ as a probability distribution on $[e]$, and we can consider the Shannon entropy $H(\overline{\lambda})$ of $\overline{\lambda}$.

### 3.1. Upper quantum functional

To define the upper quantum functional we need concepts from representation theory. Let $V$ be a complex vector space of dimension $d$. Let the symmetric group $S_n$ act on the tensor power $V^\otimes n$ by permuting the tensor legs and let the general linear group $GL(V)$ act on $V^\otimes n$ by acting on the $n$ tensor legs simultaneously. These actions commute, so $V^\otimes n$ is a $GL(V) \times S_n$-module. It decomposes into irreducible modules as

$$V^\otimes n \cong \bigoplus_{\lambda \vdash n} S_\lambda(V) \otimes [\lambda],$$

where $[\lambda]$ is an irreducible $S_n$-module and $S_\lambda(V)$ is an irreducible $GL(V)$-module if $\ell(\lambda) \leq d$ and $0$ otherwise. Decomposition (11) is referred to as Schur–Weyl duality. Let $P^V_\lambda \in \text{End}(V^\otimes n)$ be the equivariant projection onto the subspace isomorphic to $S_\lambda(V) \otimes [\lambda]$. Recall that for $S \subseteq [k]$ we defined $V_S$ as $\bigotimes_{j \in S} V_j$. Generalising the above discussion, for any subset $S \subseteq [k]$, the group $GL(V_S) \times S_n$ acts naturally on $(V_1 \otimes \cdots \otimes V_k)^\otimes n$, which decomposes as

$$(V_1 \otimes \cdots \otimes V_k)^\otimes n \cong \bigoplus_{\lambda \vdash n} S_\lambda(V_S) \otimes [\lambda] \otimes (V_S)^\otimes n$$

with $GL(V_S)$ acting trivially on $(V_S)^\otimes n$. We define $P^V_S \in \text{End}((V_S)^\otimes n)$ to be the equivariant projection onto the subspace isomorphic to $S_\lambda(V_S) \otimes [\lambda] \otimes (V_S)^\otimes n$.

We are interested in powers $t^\otimes n$ so we want to restrict $(V_S)^\otimes n$ to the symmetric subspace $S^{(n)}(V_S)$. The following observation is well-known.

**Lemma 3.1.** Let $S \subseteq [k]$. Let $\lambda \vdash n$.

1. The projections $P^V_{(\mu)}$ and $P^V_{(\lambda)}$ commute.

2. The projections $P^V_{\lambda} P^V_{(\mu)}$ and $P^V_{\mu} P^V_{(\lambda)}$ are equal.

**Proof.** For the first statement, the projector $P^V_{(\mu)}$ is just the symmetriser $\frac{1}{n!} \sum_{\pi \in S_n} \pi$, which commutes with the $S_n$-equivariant projector $P^V_{(\lambda)}$. We now prove the second statement. By Schur–Weyl duality,

$$(V_S)^\otimes n \otimes (V_S)^\otimes n \cong \bigoplus_{\lambda,\mu} S_\lambda(V_S) \otimes [\mu] \otimes S_\mu(V_S) \otimes [\mu].$$
Considering the symmetric subspace gives

\[(V_S \otimes V_S)^{\otimes n})^S_n \cong \bigoplus_{\lambda, \mu} S_\lambda(V_S) \otimes S_\mu(V_S) \otimes ([\lambda] \otimes [\mu])^{S_n}\]

\[\cong \bigoplus_{\lambda} S_\lambda(V_S) \otimes S_\lambda(V_S)\]

since \(([\lambda] \otimes [\mu])^{S_n} \cong ([\lambda]^* \otimes [\mu])^{S_n} \cong \text{Hom}_{S_n}([\lambda], [\mu])\) is one-dimensional if \(\lambda = \mu\) and the zero space otherwise. The image of \(P^{V_S}_{\lambda} P^{V_S}_{(n)}\) and the image of \(P^{V_S}_{\lambda} P^{V_S}_{(n)}\) are both equal to the subspace of \((V_S \otimes V_S)^{\otimes n}\) isomorphic to \(S_\lambda(V_S) \otimes S_\lambda(V_S)\).

Lemma 3.1 says that we can unambiguously label isotypical projections on the symmetric subspace by a pair \((b, \lambda) \in B \times \mathcal{P}_n\). For \(b = \{S, \overline{S}\} \in B\) and a partition \(\lambda \vdash n\), define \(P^b_\lambda := P^{V_S}_{\lambda} P^{V_S}_{(n)}\).

**Lemma 3.2.** If \(b_1, b_2 \in B\) are noncrossing bipartitions and \(\lambda^{(1)}, \lambda^{(2)}\) are partitions, then \(P^{V_{b_1}}_{\lambda^{(1)}} P^{V_{b_2}}_{\lambda^{(2)}}\) commute.

**Proof.** Without loss of generality we may say \(b_1 = \{S_1, \overline{S_1}\}\) and \(b_2 = \{S_2, \overline{S_2}\}\) with \(S_1 \cap S_2 = \emptyset\). Clearly the projectors \(P^{V_{S_1}}_{\lambda^{(1)}}\) and \(P^{V_{S_2}}_{\lambda^{(2)}}\) commute, since the intersection \(S_1 \cap S_2\) is empty. By statement 1 in Lemma 3.1 we obtain

\[P^{V_{b_1}}_{\lambda^{(1)}} P^{V_{b_2}}_{\lambda^{(2)}} = P^{V_{b_1}}_{\lambda^{(1)}} P^{V_{b_2}}_{(n)} P^{V_{b_2}}_{\lambda^{(2)}} P^{V_{b_1}}_{(n)} = P^{V_{b_2}}_{\lambda^{(2)}} P^{V_{b_1}}_{\lambda^{(1)}}. \]

We can now define the upper quantum functional.

**Definition 3.3.** Let \(t \in V_1 \otimes \cdots \otimes V_k\). Let \(\theta \in \mathcal{P}(B)\). Define

\[E^\theta(t) = \sup_{(\lambda^{(b)})_{b \in \text{supp} \theta}} \sum_{b \in \text{supp} \theta} \theta(b) H(\lambda^{(b)})\]

with the supremum over all tuples \((\lambda^{(b)})_{b \in B}\) of partitions \(\lambda^{(b)} \vdash n\) such that

\[\prod_{b \in \text{supp} \theta} P^{V_{b}}_{\lambda^{(b)}} t^{\otimes n} \neq 0.\]

We define \(E^\theta(t) = 2E^\theta(t)\) if \(t \neq 0\) and \(E^\theta(0) = 0\). We call \(E^\theta\) the upper quantum functional and \(E^\theta\) the logarithmic upper quantum functional.

**Remark 3.4.** The definition of \(E^\theta(t)\) in Definition 3.3 depends on the order of the product \(\prod_{b \in \text{supp} \theta} P^{V_{b}}_{\lambda^{(b)}}\), since the projectors in general do not commute. For \(\theta \in \mathcal{P}_{nc}\), however, the projectors do commute and the order does not matter. Since we will mostly focus on \(\theta \in \mathcal{P}_{nc}\) we will leave the order of the elements of \(B\) implicit.

**Theorem 3.5.** Let \(\theta \in \mathcal{P}_{nc}(B)\). Let \(s \in V_1 \otimes \cdots \otimes V_k\) and \(t \in W_1 \otimes \cdots \otimes W_k\).

1. \(F^\theta(r) = r\) for \(r \in \mathbb{N}\).
2. \(F^\theta(s \oplus t) \leq F^\theta(s) + F^\theta(t)\).
3. \(F^\theta(s \otimes t) \leq F^\theta(s) F^\theta(t)\).
4. If \(s \succeq t\), then \(F^\theta(s) \geq F^\theta(t)\).
Statement 4 is a special case of the following lemma.

**Lemma 3.6.** Let $\theta \in \mathcal{P}(B)$. If $s \sqsupset t$, then $F^\theta(s) \geq F^\theta(t)$.

**Proof.** We may assume $s$ and $t$ are not zero. Let $(\lambda^{(b)})_b$ be a tuple of partitions of $n$. It is sufficient to show

$$\prod_{b \in \text{supp } \theta} P_{\lambda^{(b)}} s^n \neq 0 \implies \prod_{b \in \text{supp } \theta} P_{\lambda^{(b)}} t^n \neq 0,$$

since then $E^\theta(s)$ is a supremum over a larger set than $E^\theta(t)$. Suppose that

$$\prod_{b \in \text{supp } \theta} P_{\lambda^{(b)}} s^n = 0,$$

Let $A_i : V_i \to W_i$ (i $\in [k]$) be linear maps. For any $S \subseteq [k]$ and $\lambda \vdash n$

$$P_{\lambda^{(S)}}(A_1 \otimes \cdots \otimes A_k)^n = (A_1 \otimes \cdots \otimes A_k)^n P_{\lambda^{(S)}}.$$

Therefore

$$\prod_{b \in \text{supp } \theta} P_{\lambda^{(b)}}(A_1 \otimes \cdots \otimes A_k)^n s^n = (A_1 \otimes \cdots \otimes A_k)^n \prod_{b \in \text{supp } \theta} P_{\lambda^{(b)}} s^n = 0,$$

and by continuity $\prod_{b \in \text{supp } \theta} P_{\lambda^{(b)}} t^n = 0$. \qed

In the rest of this section we prove statement 2 and 3, for which we need bounds on the dimension of irreducible $\text{GL}(V)$- and $S_n$-representations, and the semigroup property of Kronecker coefficients and Littlewood–Richardson coefficients.

**Remark 3.7 (Dimension bounds).** The number of partitions of $n$ into at most $d$ parts is upper bounded by $(n + 1)^d$. For $\lambda \vdash n$ the dimension of the irreducible $S_n$-module $\mathbb{S}_\lambda$ is given by the hook-length formula

$$\dim \mathbb{S}_\lambda = \frac{n!}{\prod_{(i,j) \in Y(\lambda)} \text{hook}(i,j)}$$

where $Y(\lambda)$ is the Young diagram of shape $\lambda$, and $\text{hook}(i,j)$ equals the number of boxes in $Y(\lambda)$ directly below $(i,j)$ plus the number of boxes directly to the right of $(i,j)$ plus one, and the product is over all coordinates of $Y(\lambda)$. For $\lambda \vdash d$ $n$ the dimension of the irreducible $\text{GL}(V)$-module $\mathbb{S}_\lambda(V)$ is given by the formula

$$\dim \mathbb{S}_\lambda(V) = \prod_{1 \leq i < j \leq d} \frac{\lambda_i - \lambda_j + j - i}{j - i}.$$
Definition 3.8. Let $\mu, \nu \vdash n$ be partitions. Restrict the irreducible $S_n \times S_n$-representation $[\mu] \otimes [\nu]$ to an $S_n$-representation via $S_n \to S_n \times S_n : \pi \mapsto (\pi, \pi)$ and consider the isotypic decomposition,

$$[\mu] \otimes [\nu] \downarrow_{S_n}^{S_n \times S_n} \cong \bigoplus_{\lambda \vdash n} \mathbb{C}^{g_{\lambda, \mu, \nu}} \otimes [\lambda]$$

(14)

where $\mathbb{C}^{g_{\lambda, \mu, \nu}}$ is the multiplicity space for the irreducible representation $[\lambda]$ in its isotypic component. An irreducible number $g_{\lambda, \mu, \nu}$ is called a Kronecker coefficient. (Equivalently, $g_{\lambda, \mu, \nu} = \dim([\lambda] \otimes [\mu] \otimes [\nu])_{S_n}$ which justifies the symmetric notation.) Let $\lambda \vdash_{a+b}$ be a partition. Restrict the irreducible $\text{GL}_a \times \text{GL}_b$-module $S_\lambda(C^{a+b})$ to a $\text{GL}_a \times \text{GL}_b$-module via the block diagonal embedding $\text{GL}_a \times \text{GL}_b \to \text{GL}_{a+b}$, and consider the isotypic decomposition,

$$S_\lambda(C^{a+b}) \downarrow_{\text{GL}_a \times \text{GL}_b}^{\text{GL}_a \times \text{GL}_b} \cong \bigoplus_{\mu, \nu} \mathbb{C}^{c_{\lambda, \mu, \nu}} \otimes S_\mu(C^a) \otimes S_\nu(C^b)$$

(15)

where $\mathbb{C}^{c_{\lambda, \mu, \nu}}$ is the multiplicity space for the irreducible module $S_\mu(C^a) \otimes S_\nu(C^b)$ in its isotypic component. The number $c_{\lambda, \mu, \nu}$ is called a Littlewood–Richardson coefficient.

Remark 3.9 (Semigroup property). If $\lambda$ and $\lambda'$ are partitions, then $\lambda + \lambda'$ is defined by elementwise addition. A fundamental property of the Kronecker coefficients and the Littlewood–Richardson coefficients is the well-known semigroup property (see e.g. [CHM07]):

- if $g_{\lambda, \mu, \nu} > 0$ and $g_{\alpha, \beta, \gamma} > 0$, then $g_{\lambda + \alpha, \mu + \beta, \nu + \gamma} > 0$;
- if $c_{\mu, \nu} > 0$ and $c_{\beta, \gamma} > 0$, then $c_{\mu + \beta, \nu + \gamma} > 0$.

In other words, the triples of partitions for which the Kronecker coefficients are nonzero form a semigroup under elementwise addition, and the same is true for Littlewood–Richardson coefficients.

The semigroup properties can be used to prove the following lemma. Of this lemma, the first statement can be found in [CM06], while we do not know of any source that explicitly states the second statement. For the convenience of the reader we give the proofs of both statements. Let $H(P)$ denote the Shannon entropy and let $h(p)$ denote the binary entropy function.

Lemma 3.10. Let $\lambda, \mu, \nu$ be integer partitions.

1. If $g_{\lambda, \mu, \nu}$ is nonzero, then $H(\lambda) \leq H(\mu) + H(\nu)$.

2. If $c_{\mu, \nu}$ is nonzero, then $H(\lambda) \leq \frac{|\mu|}{|\mu| + |\nu|} H(\mu) + \frac{|\nu|}{|\mu| + |\nu|} H(\nu) + h\left(\frac{|\mu|}{|\mu| + |\nu|}\right)$.

Proof. We begin with the first statement. Suppose $g_{\lambda, \mu, \nu}$ is nonzero. The semigroup property of Kronecker coefficients implies $g_{N\lambda, N\mu, N\nu} \neq 0$ for any positive integer $N$, where $N\lambda$ is the partition with parts $N\lambda_i$. Then the irreducible representation $[N\lambda]$ is isomorphic to a subspace of $[N\mu] \otimes [N\nu]$ by (14). So obviously $\dim[N\lambda] \leq \dim[N\mu] \dim[N\nu]$. We use the dimension bounds (12) to see that $NnH(\lambda) - o(N) \leq NnH(\mu) + NnH(\nu)$ when $N \to \infty$. We thus obtain the required inequality $H(\lambda) \leq H(\mu) + H(\nu)$.
We prove the second statement. By Schur–Weyl duality, as an $S_n \times \text{GL}_{a+b}$-representation $(C^{a+b})^\otimes n$ decomposes as

$$(C^{a+b})^\otimes n \cong \bigoplus_{\lambda^a+b} [\lambda] \otimes S_\lambda(C^{a+b})$$

and for the restriction via $\text{GL}_a \times \text{GL}_b \to \text{GL}_{a+b}$ (denoted by ↓) we obtain

$$(C^{a+b})^\otimes n \downarrow \cong \bigoplus_{\lambda^a+b, \mu^b} \bigoplus_{\nu} [\lambda] \otimes C^{a,\nu}_\lambda \otimes S_\mu(C^a) \otimes S_\nu(C^b).$$

On the other hand, we may first expand $(C^a \oplus C^b)^\otimes n$ and then apply Schur–Weyl duality for $\text{GL}_a$ and $\text{GL}_b$ separately,

$$(C^{a+b})^\otimes n \cong (C^a \oplus C^b)^\otimes n \cong \bigoplus_{k=0}^n C^{a,k} \otimes (C^a)^\otimes k \otimes (C^b)^\otimes n-k \cong \bigoplus_{k=0}^n \bigoplus_{\mu^a+k} \bigoplus_{\nu^b} C^{a,k}_\mu \otimes [\mu] \otimes S_\mu(C^a) \otimes [\nu] \otimes S_\nu(C^b).$$

Suppose $c_{\mu,\nu}^a$ is nonzero. Then the irreducible representation $[\lambda]$ is isomorphic to a subspace of $C^{(\mu)} \otimes [\nu] \otimes [\nu]$. So obviously $\dim [\lambda] \leq \binom{n+\mu}{\mu} \dim [\mu] \dim [\nu]$. Because of the semigroup property we may repeat everything for $N\lambda, N\mu, N\nu$ instead of $\lambda, \mu, \nu$ and obtain $\dim [N\lambda] \leq \binom{Nn+\mu}{\mu} \dim [N\mu] \dim [N\nu]$. We use the dimension bounds (12) to see that $NH(\overline{\lambda}) - o(N) \leq Nh(\frac{m}{n}) + \frac{|\nu|}{n} NH(\overline{\nu}) + \frac{|\nu|}{n} NH(\overline{\nu})$ when $N \to \infty$, where $h(p)$ denotes the binary entropy function.

Let $V$ and $W$ be vector spaces of dimension $d$ and $e$. The tensor power $(V \oplus W)^\otimes n$ is both a $\text{GL}(V) \times \text{GL}(W)$-module and a $\text{GL}(V \oplus W)$-module. Let $\lambda \vdash d+e, n + m, \mu \vdash d, n$ and $\nu \vdash e, m$. The projections $P_{\lambda}^{V \oplus W}$ and $P_{\mu}^{V \oplus W}$ commute, and their product is nonzero if and only if the Littlewood–Richardson coefficient $c_{\mu,\nu}^a$ is nonzero.

**Lemma 3.11.** Let $\theta \in \mathcal{P}_m(B)$. Then $F_{\theta}(s \oplus t) \leq F_{\theta}(s) + F_{\theta}(t)$.

**Proof.** Let $(\lambda^{(b)})_{b \in \text{supp } \theta}$ be a tuple of partitions of $n$ such that

$$\prod_{b \in \text{supp } \theta} P_{\lambda^{(b)}}^{V \oplus W}(s \oplus t)^\otimes n \neq 0. \tag{16}$$

Since $P_{\lambda^{(b)}}^{V \oplus W}(s \oplus t)^\otimes n \cong P_{\lambda^{(b)}}^{V \oplus W}(s \otimes m \otimes t^{(n-m)})$, statement (16) is equivalent to

$$\sum_{m=0}^n \binom{n}{m} \prod_{b \in \text{supp } \theta} P_{\lambda^{(b)}}^{V \oplus W}(s \otimes m \otimes t^{(n-m)}) \neq 0. \tag{17}$$

At least one of the summmands in (17) is nonzero, i.e. for some $m$

$$\prod_{b \in \text{supp } \theta} P_{\lambda^{(b)}}^{V \oplus W}(s \otimes m \otimes t^{(n-m)}) \neq 0.$$
We may write \( s^{\otimes m} \otimes t^{\otimes (n-m)} \) as
\[
s^{\otimes m} \otimes t^{\otimes (n-m)} = \sum_{(\mu^{(b)}), b \in \text{supp } \theta} \prod_{b \in \text{supp } \theta} P_{\lambda^{(b)}_{\mu^{(b)}}} s^{\otimes m} \otimes \prod_{b \in \text{supp } \theta} P_{\nu^{(b)}_{\mu^{(b)}}} t^{\otimes (n-m)}
\]
where the sum is over tuples of partitions \( \mu^{(b)} \vdash m \) and \( \nu^{(b)} \vdash n-m \). There thus exist tuples \((\mu^{(b)})_b, (\nu^{(b)})_b\) such that
\[
P_{\lambda^{(b)}_{\mu^{(b)}}}(P_{\lambda^{(b)}_{\mu^{(b)}}} \otimes P_{\nu^{(b)}_{\mu^{(b)}}}) \neq 0 \quad \text{for all } b \in \text{supp } \theta \quad \tag{18}
\]
\[
\prod_{b \in \text{supp } \theta} P_{\lambda^{(b)}_{\mu^{(b)}}} s^{\otimes m} \neq 0 \quad \tag{19}
\]
\[
\prod_{b \in \text{supp } \theta} P_{\nu^{(b)}_{\mu^{(b)}}} t^{\otimes (n-m)} \neq 0. \quad \tag{20}
\]
Eq. (18) implies that the Littlewood–Richardson coefficient \( c_{\mu^{(b)}, \nu^{(b)}}^{\lambda^{(b)}} \) is nonzero and thus by Lemma 3.10, \( H(\lambda^{(b)}) \leq \frac{m}{n} H(\mu^{(b)}) + (1 - \frac{m}{n}) H(\nu^{(b)}) + h(p) \). Setting \( p = m/n \) we can therefore write
\[
\sum_{b \in \text{supp } \theta} \theta(b) H(\lambda^{(b)}) \leq p \sum_{b \in \text{supp } \theta} \theta(b) H(\mu^{(b)}) + (1 - p) \sum_{b \in \text{supp } \theta} \theta(b) H(\nu^{(b)}) + h(p)
\]
which because of (19) and (20) is at most \( pE^\theta(s) + (1 - p)E^\theta(t) + h(p) \). Taking the supremum over the admissible tuples \((\lambda^{(b)})_b\) gives the inequality \( E^\theta(s \oplus t) \leq pE^\theta(s) + (1 - p)E^\theta(t) + h(p) \). We are done by the following Lemma 3.12.

**Lemma 3.12** (See e.g. [Str91, Eq. 2.13]). Let \( x, y \in \mathbb{R}_{\geq 0} \). Then
\[
\max_{0 \leq p \leq 1} p^x + (1-p)^y + h(p) = 2^x + 2^y.
\]

The tensor power \((V \otimes W)^\otimes n\) is both a \( \text{GL}(V) \times \text{GL}(W) \)-module and a \( \text{GL}(V \otimes W) \)-module. Let \( \lambda \vdash d n, \mu \vdash d n \) and \( \nu \vdash \tau n \). The projections \( P_V^{\otimes W} \) and \( P_V \otimes P_W \) commute, and their product is nonzero if and only if the Kronecker coefficient \( g_{\lambda, \mu, \nu} \) is nonzero.

**Lemma 3.13.** Let \( \theta \in \mathcal{P}_{nc}(B) \). Then \( F^\theta(s \oplus t) \leq F^\theta(s)F^\theta(t) \).

**Proof.** Let \((\lambda^{(b)})_b\) be a tuple of partitions of \( n \) such that
\[
\prod_{b \in \text{supp } \theta} P_{\lambda^{(b)}}(s \otimes t)^\otimes n \neq 0.
\]

We may write
\[
s^{\otimes n} \otimes t^{\otimes n} = \sum_{(\mu^{(b)})_b, (\nu^{(b)})_b \vdash \tau n} \prod_{b \in \text{supp } \theta} P_{\lambda^{(b)}_{\mu^{(b)}}} s^{\otimes n} \otimes \prod_{b \in \text{supp } \theta} P_{\nu^{(b)}_{\mu^{(b)}}} t^{\otimes n}.
\]

There exist tuples \((\mu^{(b)})_b, (\nu^{(b)})_b\) such that
\[
P_{\lambda^{(b)}_{\mu^{(b)}}}(P_{\lambda^{(b)}_{\mu^{(b)}}} \otimes P_{\nu^{(b)}_{\mu^{(b)}}}) \neq 0 \quad \text{for all } b \in \text{supp } \theta \quad \tag{21}
\]
\[
\prod_{b \in \text{supp } \theta} P_{\lambda^{(b)}_{\mu^{(b)}}} s^{\otimes n} \neq 0 \quad \tag{22}
\]
\[
\prod_{b \in \text{supp } \theta} P_{\nu^{(b)}_{\mu^{(b)}}} t^{\otimes n} \neq 0. \quad \tag{23}
\]
By (21) the Kronecker coefficient $g_{\lambda^{(b)},\mu^{(b)},\nu^{(b)}}$ is nonzero and so by Lemma 3.10 we have $H(\lambda^{(b)}) \leq H(\mu^{(b)}) + H(\nu^{(b)})$. Therefore

$$\sum_{b \in \text{supp} \theta} \theta(b) H(\lambda^{(b)}) \leq \sum_{b \in \text{supp} \theta} \theta(b) H(\mu^{(b)}) + \sum_{b \in \text{supp} \theta} \theta(b) H(\nu^{(b)})$$

which because of (22) and (23) is at most $E^\theta(s) + E^\theta(t)$. Taking the supremum over the admissible tuples $(\lambda^{(b)})_b$ gives $E^\theta(s \otimes t) \leq E^\theta(s) + E^\theta(t)$. \qed

### 3.2. Lower quantum functional

To define the lower quantum functional we need concepts from quantum information theory. We work with finite-dimensional Hilbert spaces. A state or density operator on a Hilbert space $H$ is a positive semidefinite linear map $\rho : H \to H$ satisfying $\text{Tr} \rho = 1$. Let $S(H)$ be the set of states on $H$. The von Neumann entropy or quantum entropy of $\rho$ is defined as $H(\rho) = -\text{Tr} \rho \log \rho$. If $H = H_1 \otimes \cdots \otimes H_k$, then the $j$th marginal of a state $\rho \in S(H)$ is $\rho_j = (\text{Tr}_{H_1} \otimes \cdots \otimes \text{Tr}_{H_{j-1}} \otimes \text{Id} \otimes \text{Tr}_{H_{j+1}} \cdots \otimes \text{Tr}_{H_k}) \rho$. More generally, if $S \subseteq [k]$, then $\rho_S = (\bigotimes_{j \in S} \text{Id}_{H_j} \otimes \bigotimes_{j \notin S} \text{Tr}_{H_j}) \rho \in S(H_S)$. A state $\rho$ is pure if $\rho$ has rank one as a linear map. If $\rho \in S(H_1 \otimes \cdots \otimes H_k)$ is pure and $S \subseteq [k]$, then $\rho_S$ and $\rho_T$ are unitarily equivalent up to enlarging the underlying Hilbert spaces $H_1, \ldots, H_k$, and therefore the entropy of $\rho_S$ equals the entropy of $\rho_T$.

#### Remark 3.14

A probability distribution $P \in \mathcal{P}(X)$ gives rise to a state $\rho \in S(\mathbb{C}^X)$ as $\rho = \sum_{x \in X} P(x) |x \rangle \langle x|$. This gives a bijection between probability distributions on $X$ and those states on $\mathbb{C}^X$ which are diagonal in the standard basis. In this way $\mathcal{P}(X)$ can be identified with a subset of $S(\mathbb{C}^X)$, and this identification is compatible with the notations of marginal distribution/marginal state and Shannon entropy/von Neumann entropy.

#### Definition 3.15

Let $H = H_1 \otimes \cdots \otimes H_k$ be a Hilbert space. Given $\theta \in \mathcal{P}(B)$ and a nonzero vector $\psi \in H$, define

$$H_\theta(\psi) := \sum_{S \in B} \theta(S) \left( \frac{1}{\langle \psi | \psi \rangle} |\psi \rangle \langle \psi| \right)_S.$$

#### Definition 3.16

Let $t \in V_1 \otimes \cdots \otimes V_k$. Choose inner products on each $V_j$ so that they become inner product spaces. For any $\theta \in \mathcal{P}(B)$ we define

$$E_\theta(t) = \sup_{A_1, \ldots, A_k} H_\theta((A_1 \otimes \cdots \otimes A_k)t)$$

where the supremum is over invertible linear maps $A_j \in \text{GL}(V_j)$. For $t = 0$ we set $E_\theta(t) = -\infty$. We also define $F_\theta(t) = 2^{E_\theta(t)}$ if $t \neq 0$ and $F_\theta(0) = 0$. We call $E_\theta$ the logarithmic lower quantum functional and we call $F_\theta$ the lower quantum functional.

#### Remark 3.17

The supremum in the definition of $E_\theta$ is independent of the chosen inner products, since different inner products on a vector space are related by invertible linear maps. We could equivalently define $E_\theta(t)$ as a supremum of $H_\theta(t)$ over the choice of local inner products. We may as well restrict to those $k$-tuples of inner products for which $(t|t)$ equals one.
The definition of $E_\theta$ is not sensitive to embedding each $V_j$ in some larger vector space. In fact, we get another equivalent definition of $E_\theta$ by choosing large enough inner product spaces $H_1, \ldots, H_k$ and taking the supremum of $H_\theta((A_1 \otimes \cdots \otimes A_k)t)$ over all injective linear maps $A_j : V_j \rightarrow H_j$.

**Example 3.18.** $F_\theta(\langle r \rangle) = r$. The proof is as follows. Let $\psi = \langle r \rangle$. If we make the local bases orthonormal, then $|\psi\rangle \langle \psi|_S$ is a projector of rank $r$, and the entropy of the normalised spectrum equals $\log_2(r)$, for any $\emptyset \neq S \subseteq [k]$. This gives the lower bound. On the other hand, the quantum entropy of a density matrix $\rho$ is at most the matrix rank of $\rho$. Therefore, $H(|\psi\rangle \langle \psi|_S) \leq R(|\psi\rangle \langle \psi|_S) \leq R(\text{flatten}_S(\psi))$. This gives the upper bound.

The key properties of the lower quantum functional are as follows.

**Theorem 3.19.** Let $\theta \in P(B)$. Let $s \in V_1 \otimes \cdots \otimes V_k$, $t \in W_1 \otimes \cdots \otimes W_k$.

1. $F_\theta(\langle r \rangle) = r$ for any $r \in \mathbb{N}$.
2. $F_\theta(s \oplus t) \geq F_\theta(s) + F_\theta(t)$.
3. $F_\theta(s \otimes t) \geq F_\theta(s)F_\theta(t)$.
4. If $s \geq t$, then $F_\theta(s) \geq F_\theta(t)$.
5. $0 \leq F_\theta(s) \leq \prod \{\dim(V_S) \cdot \dim(V_S')\}^{\theta(|S, S'|)}$.

Statement 1 is Example 3.18. Statement 5 is easy to prove. In the rest of this section we prove statement 2–4.

**Lemma 3.20.** If $s \geq t$, then $F_\theta(s) \geq F_\theta(t)$.

**Proof.** We may without loss of generality assume that $s$ and $t$ are elements of the same space $V_1 \otimes \cdots \otimes V_k$ (Remark 3.17). Let $G = \text{GL}(V_1) \times \cdots \times \text{GL}(V_k)$. By assumption, there is a sequence $(A_1, \ldots, A_{k, i}) \in G$ such that $\lim_{i \rightarrow \infty} (A_1, \ldots, A_{k, i})s = t$. For any $k$-tuple $(B_1, \ldots, B_k) \in G$ we have therefore

$$E_\theta(s) = \sup_{C_1, \ldots, C_k} H_\theta((C_1 \otimes \cdots \otimes C_k)s)$$

$$\geq \lim_{i \rightarrow \infty} H_\theta((B_1 A_1, \ldots, B_k A_{k, i})t)$$

$$= H_\theta((B_1 \otimes \cdots \otimes B_k)t)$$

by continuity of $H$. Now take the supremum over all $(B_1, \ldots, B_k) \in G$ to obtain $E_\theta(s) \geq E_\theta(t)$ and thus $F_\theta(s) \geq F_\theta(t)$.

**Lemma 3.21** (Recursion property of quantum entropy, e.g. [BZ06, Eq. 12.19]). Let $\rho_i$ be density matrices with support in orthogonal subspaces $H_i$ of a Hilbert space $H = \bigoplus_{i=1}^M H_i$. Then the density matrix $\rho = \sum_i p_i \rho_i$ has quantum entropy $H(\rho) = H(p) + \sum_{i=1}^M p_i H(\rho_i)$.

**Lemma 3.22.** $F_\theta(s \oplus t) \geq F_\theta(s) + F_\theta(t)$.

**Proof.** Choose inner products $\langle \cdot, \cdot \rangle_j$ on $V_j$ and $\langle \cdot, \cdot \rangle'_j$ on $W'_j$ in such a way that $s$ and $t$ have norm one, and consider $H_\theta(s)$ and $H_\theta(t)$ with respect to these inner products (cf. Remark 3.17). Equip $V'_j \oplus W'_j$ with the direct sum of the inner
products. Let \( u = \sqrt{p} \otimes \sqrt{1-p} t \in (V_1 \otimes W_1) \otimes \cdots \otimes (V_k \otimes W_k) \). For any \( p \in [0,1] \) the vector \( u \) has norm one and \( u \cong s \otimes t \). For any subset \( S \subseteq [k] \) the reduced states are related as \( |u|_{S} = p |s|_{S} \oplus (1-p) |t|_{S} \). Therefore, by recursivity of quantum entropy (Lemma 3.21),

\[
H(|u|_S) = p H(|s|_S) + (1-p) H(|t|_S) + h(p).
\]

Taking the \( \theta \)-weighted average of both sides gives

\[
H_{\theta}(u) = p H_{\theta}(s) + (1-p) H_{\theta}(t) + h(p).
\]

Since \( E_{\theta}(s \otimes t) \geq H_{\theta}(u) \) we may take the supremum over the choices of inner products to get

\[
E_{\theta}(s \otimes t) \geq p E_{\theta}(s) + (1-p) E_{\theta}(t) + h(p).
\]

Apply Lemma 3.12 to see that \( F_{\theta}(s \otimes t) \geq F_{\theta}(s) + F_{\theta}(t) \).

**Lemma 3.23.** \( F_{\theta}(s \otimes t) \geq F_{\theta}(s)F_{\theta}(t) \).

**Proof.** Choose inner products \( \langle \cdot, \cdot \rangle_j \) on \( V_j \) and \( \langle \cdot, \cdot \rangle'_j \) on \( W_j \) in such a way that \( s \) and \( t \) have norm one, and consider \( H_{\theta}(s) \) and \( H_{\theta}(t) \) with respect to these inner products. Equip \( V_j \otimes W_j \) with the tensor product of the inner products. Then \( s \otimes t \in (V_1 \otimes W_1) \otimes \cdots \otimes (V_k \otimes W_k) \) has norm one. For any subset \( S \subseteq [k] \) the reduced states are related as \( |s \otimes t|_{S} = |s|_{S} \otimes |t|_{S} \). Therefore \( H(|s \otimes t|_{S}) = H(|s|_{S}) + H(|t|_{S}) \). Taking the \( \theta \)-weighted average of both sides gives \( H_{\theta}(s \otimes t) = H_{\theta}(s) + H_{\theta}(t) \). Since \( E_{\theta}(s \otimes t) \geq H_{\theta}(s \otimes t) \), taking the supremum over the choice of inner products gives \( E_{\theta}(s \otimes t) \geq E_{\theta}(s) + E_{\theta}(t) \) and thus \( F_{\theta}(s \otimes t) \geq F_{\theta}(s)F_{\theta}(t) \) \( \Box \).

### 3.3. Comparing the quantum functionals

In this section we show that in the general regime \( \theta \in \mathcal{P}(B) \) the upper quantum functional is at least the lower quantum functional, justifying the names.

**Theorem 3.24.** Let \( \theta \in \mathcal{P}(B) \). Then

\[
E^{\theta}(t) \geq E_{\theta}(t).
\]

To prove Theorem 3.24 we use the gentle measurement lemma and the spectrum estimation theorem. A proof of the gentle measurement lemma can be found in [Win99], see also [ON02, Lemma 5] for the specific bound that we use. Keyl and Werner proved the spectrum estimation theorem [KW01], see also [HM02] and [CM06, Theorem 1] for a more succinct proof. Let \( \mathcal{S}_{\leq}(\mathcal{H}) \) be the set of positive semidefinite operators \( \rho : \mathcal{H} \to \mathcal{H} \) satisfying \( \text{Tr} \rho \leq 1 \).

**Lemma 3.25** (Gentle measurement). Let \( \rho \in \mathcal{S}_{\leq}(\mathcal{H}) \). Let \( X : \mathcal{H} \to \mathcal{H} \) be a linear map such that \( 0 \leq X \leq I \). Then

\[
\| X \rho X - \rho \|_1 \leq 2\sqrt{1 - \text{Tr}(X \rho X)}.
\]

**Corollary 3.26.** Let \( \rho_0 \in \mathcal{S}(\mathcal{H}) \) and \( 0 \leq X_1, \ldots, X_n \leq I \). Suppose that for each \( i \in [n] \) the inequality \( \text{Tr}(X_i \rho_0 X_i) \geq 1 - \varepsilon \) holds. Then

\[
\| X_n \cdots X_2 X_1 \rho_0 X_1 X_2 \cdots X_n - \rho_0 \|_1 \leq 2n \sqrt{\varepsilon}.
\]

In particular,

\[
\text{Tr}(X_n \cdots X_2 X_1 \rho_0 X_1 X_2 \cdots X_n) \geq 1 - 2n \sqrt{\varepsilon}.
\]
Proof. Let $\rho_i = X_i\rho_{i-1}X_i$ for $i \in [n]$. By the Hölder inequality and the gentle measurement lemma (Lemma 3.25), we have

$$\|\rho_i - \rho_0\|_1 \leq \|X_i\rho_{i-1} - \rho_0\|_1 + \|X_i\rho_0X_i - \rho_0\|_1$$

$$\leq \|X_i\rho_{i-1} - \rho_0\|_1 \cdot \|X_i\rho_0X_i - \rho_0\|_1 + 2\sqrt{1 - \text{Tr}(X_i\rho_0X_i)}$$

$$\leq \|\rho_{i-1} - \rho_0\|_1 + 2\sqrt{\varepsilon}.$$  

The statement follows by induction on $i$. \hfill \Box

**Theorem 3.27** (Spectrum estimation). Let $\mathcal{H}$ have dimension $d$. Let $\rho \in \mathcal{S}(\mathcal{H})$ be a density matrix. Let $r \in \mathbb{R}^d$ be the sequence of eigenvalues of $\rho$ ordered nonincreasingly. Let $\lambda \vdash n$. Then

$$\text{Tr}(P^H_{\lambda} \rho^\otimes n) \leq (n + 1)^{(d-1)/2} 2^{-nD(\rho\|r)}$$

where $D(\rho\|q) = \sum_i p_i \log_2 p_i - \log_2 q_i$ is the Kullback–Leibler divergence of probability distributions $p$ and $q$.

**Corollary 3.28.** Let $\rho \in \mathcal{S}(\mathcal{H})$. Let $\varepsilon > 0$. Let $X = \sum P^H_{\lambda}$ where the sum is over partitions $\lambda \vdash n$ such that $H(X) \geq H(\rho) - \varepsilon$. Then $\lim_{n \to \infty} \text{Tr}(\rho^\otimes n X) = 1$.

Proof. We refer to Corollary 2 in [CM06]. \hfill \Box

**Proof of Theorem 3.24.** Let $\ell = |\text{supp} \theta|$. Let $\varepsilon > 0$ be arbitrary. Choose inner products on $V_1, \ldots, V_k$ such that $t$ is a unit vector and such that the $\theta$-weighted average of the marginal quantum entropies of $t$ is at least $E_\theta(t) - \varepsilon$. Let $(r_b)_{b \in B}$ be the tuple of marginal quantum entropies of $t$. By Corollary 3.28 there is an $n_0$ such that for every $n \geq n_0$ and $i \in [\ell]$ the inequality $\text{Tr}(\langle t | i^n X_i \rangle) \geq 1 - \varepsilon$ holds, where

$$X_i = \sum_{\lambda \vdash n : H(X) \geq r_{\pi(i)} - \varepsilon} P^H_{\lambda}.\n$$

According to Corollary 3.26 this in turn implies

$$\text{Tr}(X_1 \cdots X_n | t^n X_1 X_2 \cdots X_n \rangle) \geq 1 - 2\ell \sqrt{\varepsilon}.$$  

In particular, $X_1 \cdots X_n t^n \neq 0$ for $\varepsilon$ small enough and $n$ large enough. By the definition of the operators $X_i$ this implies that there is a tuple $(\lambda^{(i)})_{i \in [\text{supp} \theta]}$ of partitions $\lambda^{(i)} \vdash n$ such that

$$P^V_{\lambda(\pi(i))} \cdots P^V_{\lambda(\pi(2))} P^V_{\lambda(\pi(1))} t^n \neq 0$$

and such that $H(\Gamma^{(i)}) \geq r_{\pi(i)} - \varepsilon$ for each $i \in [\ell]$. Therefore,

$$E_\theta(t) \geq \sum_{b \in \text{supp} \theta} \theta(b) r_b - \varepsilon \geq E_\theta(t) - 2\varepsilon.$$  

This holds for any $\varepsilon > 0$, so $E_\theta(t) \geq E_\theta(t)$. \hfill \Box
3.4. Universal spectral points

We have the upper quantum functional $F^θ$ which is sub-additive and sub-multiplicative, we have the lower quantum functional $F_θ$ which is super-additive and super-multiplicative, both are ≤-monotone and normalised on $⟨r⟩$, and we know that $F^θ(t) ≥ F_θ(t)$. (See Theorem 3.5, Theorem 3.19 and Theorem 3.24.)

In this section we show that $F^θ(t)$ and $F_θ(t)$ in fact coincide for the singleton regime $θ ∈ P_δ(B)$ and for any tensor $t$, which shows that $F^θ = F_θ$ is a universal spectral point. The ingredient for the proof is an object called the entanglement polytope and two characterisations for this object.

Let $H = H_1 ⊗ · · · ⊗ H_k$ be a Hilbert space. Let $G = GL(H_1) × · · · × GL(H_k)$. Let $|ψ⟩ ∈ H$ be a unit vector. Then $|ψ⟩⟨ψ|$ ∈ $S(H)$ and therefore we can consider the marginals $|ψ⟩⟨ψ|_j$ of $|ψ⟩⟨ψ|$ for $j ∈ [k]$. For each $j ∈ [k]$, let $r_j(ψ)$ be the nonincreasingly ordered tuple of eigenvalues of the $j$-th marginal. The entanglement polytope of a nonzero tensor $t ∈ H_1 ⊗ · · · ⊗ H_k$ is the subset $∆_t ⊆ R^{dim H_1 + · · · + dim H_k}$ defined as

$$∆_t = \{ (r_1(ψ),...,r_k(ψ)) \mid |ψ⟩ ∈ G^{-1}t \text{ and } ∥ψ∥ = 1 \}. $$

Note that $∆_t$ does not depend on the choice of inner products on $H_j$. To relate $E^θ$ to $F_θ$ we crucially use the following equivalent characterisation of $∆_t$, which is due to [WDGC13, Theorem 1] and [SOK14] and based on [Bri87].

**Theorem 3.29.** The entanglement polytope $∆_t$ is the Euclidean closure of the set

$$\{ (λ^{(1)},...,λ^{(k)}) \mid (P^{H_1}_{λ^{(1)}} ⊗ · · · ⊗ P^{H_k}_{λ^{(k)}})t^⊗n \neq 0 \}. $$

We now return to the quantum functionals.

**Theorem 3.30.** Let $θ ∈ P_δ(B)$. Then $E^θ(t) = E_θ(t)$.

**Proof.** Since $θ ∈ P_δ(B)$, the value of $E_θ(t)$ is related to the entanglement polytope $∆_t$ as follows

$$E_θ(t) = \max \{ θ(1)H(r_1) + · · · + θ(k)H(r_k) \mid (r_1,...,r_k) ∈ ∆_t \}. \quad (25)$$

Using the characterisation in Theorem 3.29 we can write (25) as

$$\sup \{ θ(1)H(λ^{(1)}) + · · · + θ(k)H(λ^{(k)}) \mid (P^{V_1}_{λ^{(1)}} ⊗ · · · ⊗ P^{V_k}_{λ^{(k)}})t^⊗n \neq 0 \}$$

which equals $E^θ(t)$. \hfill \Box

By the discussion at the beginning of this section the following is immediate.

**Corollary 3.31.** Let $θ ∈ P_δ(B)$. Then $F^θ = F_θ$ is a point in the asymptotic spectrum $Δ(T)$, a universal spectral point.

**Definition 3.32** (Quantum simplex). Analogous to the support simplex (Definition 2.20), we define for any family $X$ the quantum simplex $F_{P(|k|)}(X)$ as the image of the map $P(|k|) → Δ(X): θ → F^θ$.

We relate the Strassen upper support functional $ρ^θ$ to the new upper quantum functional $E^θ$. Recall that $ρ^θ$ is defined for $θ ∈ P_δ(B) ⊆ P(B)$, that is, for probability distributions $θ$ supported on bipartitions of the form $\{(j,j)\}$. Let $t ∈ V_1 ⊗ · · · ⊗ V_k$. 

27
Remark 3.33 (Majorisation property). Let $v_1, \ldots, v_d$ be a basis of $V$. Let $T \subseteq \text{GL}(V)$ be the subgroup of linear maps having diagonal matrix in this basis, a maximal torus for $\text{GL}(V)$. For a tuple $i = (i_1, \ldots, i_n) \in [d]^n$ let the vector $v_i \in V^\otimes n$ be defined as

$$v_i := v_{i_1} \otimes \cdots \otimes v_{i_n}.$$ 

Let $i \in [d]^n$ and let $Q$ be the type of $i$. Then the vector $v_i$ is a weight vector for $T$ with weight $nQ$, meaning that for any $t = \text{diag}(t_1, \ldots, t_d) \in T$ we have $t \cdot v_i = \prod_{j=1}^d (t_j)^{nQ(j)} v_i$. The module $S_\lambda(V)$ has so-called highest weight $\lambda$, and thus $B_\lambda v_i \neq 0$ implies that $nQ \preceq \lambda$, where $\preceq$ is the dominance order or majorisation order on partitions. We call this the majorisation property. In particular, $P^\lambda v_i \neq 0$ implies the inequality of Shannon entropies $H(Q) \geq H(\lambda)$, since the Shannon entropy is Schur-concave.

Theorem 3.34. Let $\theta \in P_\kappa(B)$. Then $\rho^\theta(t) \geq E^\theta(t)$. 

Proof. Choose a $k$-tuple of bases $C = ((v_{1,\alpha_1}), \ldots, (v_{k,\alpha_k}))$ in $C(t)$ such that $\rho^\theta(t) = H_\theta(\text{supp}_C t)$. Suppose $(P^V_{\lambda(1)} \otimes \cdots \otimes P^V_{\lambda(k)}) t \otimes v \neq 0$ for some $k$-tuple of partitions $(\lambda^{(1)}, \ldots, \lambda^{(k)})$ and $N \in \mathbb{N}$. Given an $N$-type $Q \in P_N([n_1] \times \cdots \times [n_k])$, define the vector $v_Q \in V_1^\otimes N \otimes \cdots \otimes V_k^\otimes N$ as

$$v_Q := \sum_{x} \bigotimes_{m=1}^N v_{1,x,m} \otimes v_{2,x,m,2} \otimes \cdots \otimes v_{k,x,m,k}$$

where the sum is over all $N$-tuples $x = (x_1, \ldots, x_N) \in T^N_q$ of type $Q$, and define the element $t_Q \in C$ as $t_Q := \prod_{x \in \text{supp} Q} (t_{1,x_1,\cdots,x_N})^{NQ(x)}$. Then the tensor power $t^\otimes N$ can be written as

$$t^\otimes N = \sum_Q t_Q v_Q$$

where the sum is over $Q \in P_N(\text{supp}_C t)$. Then $(P^V_{\lambda^{(1)}} \otimes \cdots \otimes P^V_{\lambda^{(k)}}) v_Q \neq 0$ for some $Q \in P_N(\text{supp}_C t)$. By the majorization property this implies for all $j \in [k]$ that $NQ_j \preceq \lambda_j^{(j)}$ and thus $H(Q_j) \geq H(\lambda_j^{(j)})$. This shows that

$$\rho^\theta(t) = H_\theta(\text{supp}_C t) \geq H_\theta(Q) = \sum_{j=1}^k \theta(j) H(Q_j) \geq \sum_{j=1}^k \theta(j) H(\lambda^{(j)}_j).$$

Take the supremum of the right hand side over the possible $(\lambda^{(1)}, \ldots, \lambda^{(k)})$ to get the required inequality $\rho^\theta(t) \geq E^\theta(t)$. \qed

By Theorem 3.34, for $\theta \in P([k])$, $E^\theta(t) \leq \rho^\theta(t) \leq n_1^{\theta(1)} \cdots n_k^{\theta(k)}$. Therefore, if $E^\theta(t) = n_1^{\theta(1)} \cdots n_k^{\theta(k)}$, then $\rho^\theta(t) = n_1^{\theta(1)} \cdots n_k^{\theta(k)}$. We will show the converse of this in Theorem 4.33, that is, when $\theta_i > 0$ for all $i$, if $E^\theta(t) \leq n_1^{\theta(1)} \cdots n_k^{\theta(k)}$, then $\rho^\theta(t) < n_1^{\theta(1)} \cdots n_k^{\theta(k)}$.

The final result of this section is that the regularisation of the upper support functional equals the upper quantum functional.

Theorem 3.35. Let $t \in V_1 \otimes \cdots \otimes V_k$, $\theta \in P([k])$. Then

$$\lim_{n \to \infty} \frac{1}{n} \rho^\theta(t^\otimes n) = E^\theta(t)$$
Proof. By Theorem 3.34, \( E^\theta(t) \leq \rho^\theta(t) \), so \( \lim_{n \to \infty} \frac{1}{n} E^\theta(t^{\otimes n}) \leq \lim_{n \to \infty} \frac{1}{n} \rho^\theta(t^{\otimes n}) \). We know \( E^\theta(t) = \lim_{n \to \infty} \frac{1}{n} E^\theta(t^{\otimes n}) \) by Corollary 3.31. Therefore

\[
\lim_{n \to \infty} \frac{1}{n} \rho^\theta(t^{\otimes n}) \geq E^\theta(t).
\]

We now prove \( \leq \). Let \( n \in \mathbb{N} \) and consider the decomposition

\[
V_t^{\otimes n} = \bigoplus_{\lambda \vdash n} S_\lambda(V_t) \otimes [\lambda].
\]

In each direct summand on the right hand side choose a basis with (disjoint) index sets \( I_{1,\lambda}, \lambda \vdash n \). Together they form a basis of \( V_t^{\otimes n} \). Let \( C \in \mathcal{C}(t^{\otimes n}) \) be this \( k \)-tuple of bases. Define the set

\[
\Phi = \bigcup \left\{ I_{1,\lambda(1)} \times \cdots \times I_{k,\lambda(k)} \mid \lambda^{(1)}, \ldots, \lambda^{(k)} : \sum_{i=1}^k \theta(i) H(\overline{\lambda^{(i)}}) \leq E^\theta(t) \right\}.
\]

By definition, \( \text{supp}_C t^{\otimes n} \subseteq \Phi \), and therefore

\[
\rho^\theta(t^{\otimes n}) \leq H_\theta(\text{supp}_C t^{\otimes n}) \leq H_\theta(\Phi).
\]

To estimate \( H_\theta(\Phi) \), note that by concavity of the entropy and invariance of \( \Phi \) under permutations within each \( I_{1,\lambda} \), the maximum is attained at a distribution which is uniform within each block \( I_{1,\lambda(1)} \times \cdots \times I_{k,\lambda(k)} \). Let \( P \) be such a distribution with \( H_\theta(\Phi) = H_\theta(P) \), and let the total probability within block \( I_{1,\lambda(1)} \times \cdots \times I_{k,\lambda(k)} \) be \( Q(\lambda^{(1)}, \ldots, \lambda^{(k)}) \). Thus \( Q \) describes a “coarse” structure of \( P \), which is uniform conditioned on \( Q \). The marginals of \( P \) inherit a similar structure: \( P_i \) is uniform within each block \( I_{1,\lambda(i)} \) and the total probability of such a block is \( Q_i(\lambda^{(i)}) \). The recursion property allows us to express \( H(P_i) \) as

\[
H(P_i) = H(Q_i) + \sum_{\lambda^{(i)}} Q_i(\lambda^{(i)}) \log_2 |I_{1,\lambda^{(i)}}|
\]

\[
= H(Q_i) + \sum_{\lambda^{(i)}, \ldots, \lambda^{(k)}} Q(\lambda^{(1)}, \ldots, \lambda^{(k)}) \log_2 |I_{1,\lambda^{(i)}}|,
\]

and therefore

\[
H_\theta(P) = H_\theta(Q) + \sum_{\lambda^{(1)}, \ldots, \lambda^{(k)}} Q(\lambda^{(1)}, \ldots, \lambda^{(k)}) \sum_{i=1}^k \theta(i) \log_2 |I_{1,\lambda^{(i)}}|. \tag{27}
\]

We use the estimates

\[
|I_{1,\lambda}| = \dim(S_\lambda(V_t) \otimes [\lambda]) \leq (n+1)^{d_1(d_1-1)/2} \left( \binom{n}{\lambda_1, \ldots, \lambda_{d_1}} \right) \leq 2^{nH(\overline{\lambda}) + o(n)}
\]

and

\[
H_\theta(Q) \leq H(Q) \leq d_1 d_2 \cdots d_k \log_2 (n+1) = o(n).
\]

In the sum over \( \lambda^{(1)}, \ldots, \lambda^{(k)} \) in (27) only those terms can have nonzero probability which satisfy

\[
\sum_{i=1}^k \theta(i) H(\overline{\lambda^{(i)}}) \leq E^\theta(t).
\]
For these terms we have the upper bound
\[
\sum_{i=1}^{k} \theta(i) \log_2 |I_{i,\lambda(i)}| \leq n \sum_{i=1}^{k} \theta(i) H(\lambda(i)) + o(n) \leq nE^\theta(t) + o(n).
\]
The sum of all \(Q(\lambda^{(1)}, \ldots, \lambda^{(k)})\) equals 1, and therefore
\[
\rho^\theta(t) \leq H(\lambda) \leq nE^\theta(t) + o(n),
\]
which implies \(\lim_{n \to \infty} \frac{1}{n} \rho^\theta(t) \leq E^\theta(t).\)

**Remark 3.36.** In the language of comment 7.(ii) in [Str91], our Theorem 3.35 shows that the set \(Z^\theta\) is a single point.

Using Theorem 3.35 we can see that the quantum functionals are at least the lower support functional.

**Corollary 3.37.** Let \(t \in V_1 \otimes \cdots \otimes V_k, \theta \in \mathcal{P}_s(B)\). Then \(E^\theta(t) \geq \rho^\theta(t)\).

**Proof.** By Theorem 3.35,
\[
E^\theta(t) = \lim_{n \to \infty} \zeta^\theta(t \otimes n)^{1/n}.
\]
We know \(\zeta^\theta(t) \geq \zeta_\theta(t)\) by Theorem 2.15 and thus
\[
\lim_{n \to \infty} \zeta^\theta(t \otimes n)^{1/n} \geq \lim_{n \to \infty} \zeta_\theta(t \otimes n)^{1/n}.
\]
The lower support functional \(\zeta_\theta\) is super-multiplicative under \(\otimes\) (Theorem 2.10), so
\[
\lim_{n \to \infty} \zeta_\theta(t \otimes n)^{1/n} \geq \zeta_\theta(t)
\]
Combining these statements proves the theorem.

4. Several families of tensors

With the definitions and key properties of the Strassen support functionals and the quantum functionals in place, we will now focus on various classes of tensors. We begin with the class of tight tensors that was introduced by Strassen. Tight tensors are oblique and Strassen showed that for tight tensors the support functionals are sufficient to compute the asymptotic subrank.

Next we discuss reduced polynomial multiplication and the cap set problem. Tightness is useful in this context.

Then we introduce free tensors. Oblique tensors are free. For complex free tensors we show that the upper support functional and the quantum functionals coincide in the singleton regime.

Finally we compute generic values of the quantum functionals and compare this to the known results on the generic value of the Strassen support functionals. Related to this we apply the Hilbert–Mumford criterion to see that if the upper quantum functional is not maximal, then the upper support functional is not maximal.
4.1. Tight tensors

Let \( \mathbb{F} \) be an arbitrary field. Strassen defined the family of tight tensors as follows.

**Definition 4.1** (Tight). Let \( \Phi \subseteq I_1 \times \cdots \times I_k \). We say the set \( \Phi \) is tight if there are injective maps \( u_i : I_i \to \mathbb{Z} \) (\( i \in [k] \)) such that

\[
u_1(\alpha_1) + \cdots + u_k(\alpha_k) = 0 \quad \text{for each} \quad (\alpha_1, \ldots, \alpha_k) \in \Phi.
\]

We say a tensor \( f \) is tight if there is a \( C \in C(f) \) such that the set \( \text{supp}_C f \) is tight.

**Remark 4.2.** Clearly, tight tensors are oblique. To summarise the families of tensors that we have defined up to now, we have

\[
\{\text{tight}\} \subseteq \{\text{oblique}\} \subseteq \{\text{robust}\} \subseteq \{\theta\text{-robust}\}.
\]

Recall that the families of oblique, robust and \( \theta \)-robust tensors each form a semiring under \( \otimes \) and \( \oplus \). Tight tensors have the same property [Str91, Section 5]. Another property is that any subset of a tight set is tight.

**Example 4.3.** Let \( k \geq 3 \) be fixed. For any integer \( n \geq 1 \) and \( c \in [n] \) the following set is clearly tight:

\[
\Phi_n(c) = \{\alpha \in \{0, \ldots, n-1\}^k \mid \alpha_1 + \cdots + \alpha_k = c\}.
\]

For any integer \( n \geq 2 \) and any \( c \in [n] \) the following set is not tight (cf. Exercise 15.20 in [BCS97]):

\[
\Psi_n(c) = \{\alpha \in \{0, \ldots, n-1\}^k \mid \alpha_1 + \cdots + \alpha_k = c \text{ mod } n\}.
\]

The tensor \( f_n = \sum_{\alpha \in \Psi_n(n-1)} e_{\alpha_1} \otimes \cdots \otimes e_{\alpha_k} \in (\mathbb{F}^n)^{\otimes k} \) is tight, however, when the ground field \( \mathbb{F} \) contains a primitive \( n \)th root of unity \( \zeta \). Namely, the elements \( v_j = \sum_{\alpha} \zeta^{\alpha_j} e_j \) (\( j \in [n] \)) form a basis of \( \mathbb{F}^n \). Let \( C \in C(f_n) \) be the corresponding product basis. We have \( f_n = \sum_{j=1}^n v_j \otimes \cdots \otimes v_j \), and thus \( \text{supp}_C f_n = \{\alpha \in [n]^k \mid \alpha_1 = \cdots = \alpha_k\} \) which is clearly tight. (See also [BCS97, Exercise 15.25].) When the characteristic of \( \mathbb{F} \) equals \( n \) the tensor \( f_n \) is also tight, as we will see in Section 4.2.2.

We care about tight tensors because of the following remarkable theorem for tight 3-tensors, proved by Strassen in [Str91, Lemma 5.1] using the method of Coppersmith and Winograd [CW90]. We need the concept of asymptotic subrank of a set (cf. [Str91, Section 5]). We say \( D \subseteq I_1 \times \cdots \times I_k \) is a diagonal when any two distinct \( \alpha, \beta \in D \) are distinct in all \( k \) coordinates. In other words, for elements in \( D \), the value at one coordinate uniquely determines the value at the other \( k-1 \) coordinates. Let \( \Phi \subseteq I_1 \times \cdots \times I_k \). We say a diagonal \( D \subseteq I_1 \times \cdots \times I_k \) is free for \( \Phi \) or simply \( D \subseteq \Phi \) is a free diagonal if \( D = \Phi \cap (D_1 \times \cdots \times D_k) \), where \( D_i = \{x_i \mid (x_1, \ldots, x_k) \in D\} \). Define the subrank \( Q(\Phi) \) as the size of the largest free diagonal \( D \subseteq \Phi \). For \( \Phi \subseteq I_1 \times \cdots \times I_k \) and \( \Psi \subseteq J_1 \times \cdots \times J_k \) we naturally define the product \( \Phi \times \Psi \subseteq (I_1 \times J_1) \times \cdots \times (I_k \times J_k) \). Define the asymptotic subrank \( \tilde{Q}(\Phi) = \lim_{n \to \infty} Q(\Phi^{\times n})^{1/n} \). Let \( t \in \mathbb{F}^{n_1} \otimes \cdots \otimes \mathbb{F}^{n_v} \) and let \( \Phi \) be the support of \( t \) in the standard basis. Then \( Q(\Phi) \leq Q(t) \) and \( Q(\Phi) \leq Q(t) \). The number \( Q(\Phi) \) may be interpreted as the largest number \( n \) such that \( \langle n \rangle \) can be
Let $\Phi \subseteq [n_1] \times \cdots \times [n_k]$ and let $t \in \mathbb{F}^{n_1} \otimes \cdots \otimes \mathbb{F}^{n_k}$ be any tensor with support equal to $\Phi$. Then

$$Q(\Phi) \leq Q(t), \quad \text{and} \quad \tilde{Q}(\Phi) \leq \tilde{Q}(t).$$

So we may upper bound $\tilde{Q}(\Phi)$ by choosing the field $\mathbb{F}$ and the tensor $t$ cleverly and upper bounding $\tilde{Q}(t)$.

**Theorem 4.4.** Let $\Phi \subseteq I_1 \times I_2 \times I_3$ be tight. Then

$$\tilde{Q}(\Phi) = \max_{P \in \mathcal{P}(\Phi)} \min \{2^H(P_1), 2^H(P_2), 2^H(P_3)\}. \quad (29)$$

The amazing consequence of Theorem 4.4 is that the Strassen support functionals are sufficiently powerful to compute the asymptotic subrank of tight 3-tensors.

**Corollary 4.5** ([Str91, Proposition 5.4]). Let $f \in V_1 \otimes V_2 \otimes V_3$ be tight. Then

$$\tilde{Q}(f) = \min_{\theta \in \mathcal{P}(f)} \zeta^\theta(f).$$

Moreover, if $\Phi$ is a tight support for $f$, then $\tilde{Q}(f) = Q(\Phi)$.

**Remark 4.6.** Strassen conjectured in [Str94, Conjecture 5.3] that for the family $\mathcal{Y}$ of tight 3-tensors the support functionals give all spectral points in the asymptotic spectrum $\Delta(\mathcal{Y})$, i.e. the support simplex and the asymptotic spectrum coincide for $\mathcal{Y}$. In [Str91] numerous examples are given of subfamilies of $\mathcal{Y}$ for which this is the case. We focus on one such example in the next section.

**Remark 4.7.** The assumption $\Phi \subseteq I_1 \times I_2 \times I_3$ in Theorem 4.4 is crucial. Namely, statement (29) becomes false when instead we let $\Phi \subseteq I_1 \times \cdots \times I_k$ with $k \geq 4$ and we let the right-hand side of the equation be $\max_{P \in \mathcal{P}(\Phi)} \min_i 2^H(P_i)$, see [CVZ16, Example 1.1.38]. In [CVZ16] the construction of Theorem 4.4 is extended to obtain a lower bound for $\tilde{Q}(\Phi)$ when $k \geq 4$. This lower bound is not known to be tight in general. We give an illustration in Example 4.8 of a case in which the lower bound is tight.

**Example 4.8** (Type tensors). Let $\lambda = (\lambda_1, \ldots, \lambda_n)$ be a partition of $k$. Define the set $\Phi_\lambda$ and the type tensor $D_\lambda$ as

$$\Phi_\lambda := \{ \alpha \in [n]^k \mid \text{type}(\alpha) = \lambda \}$$

$$D_\lambda := \sum_{\alpha \in \Phi_\lambda} e_{\alpha_1} \otimes \cdots \otimes e_{\alpha_k} \in (\mathbb{F}^n)^{\otimes k}$$

where $\text{type}(\alpha) = \lambda$ means that the $k$-tuple $\alpha$ is a permutation of the $k$-tuple consisting of $\lambda_1$ ones, $\lambda_2$ twos, etc. In quantum information theory (let $\mathbb{F} = \mathbb{C}$) each type tensor $D_\lambda$ encodes a so-called *Dicke state of type* $\lambda$. The type tensor $D_{(k-1,1)}$ encodes what is called the *generalised* $W$ state. Clearly $D_\lambda$ is oblique, since the support in the standard basis $\Phi_\lambda$ is an antichain (in fact, $\Phi_\lambda$ is clearly tight), and therefore

$$\rho^\theta(D_\lambda) = \rho^\theta(D_\lambda) = H_\theta(\Phi_\lambda)$$
(and over $\mathbb{C}$ the quantum functionals $E^\theta(D_\lambda)$ and $E_\theta(D_\lambda)$ take the same value as $\rho^\theta(D_\lambda)$). One verifies that $H_\theta(\Phi_\lambda) \geq H((\frac{n}{1}, \ldots, \frac{n}{k}))$, with equality when $\theta$ is uniform. Of course $H_\theta(\Phi_\lambda) \leq n$, with equality when $\theta$ equals one of $(1,0,\ldots,0)$, $(0,1,0,\ldots,0)$, etc. It is shown in [VC15] and [CVZ16], and independently in [HX15], that for all $k \geq 3$ we have $Q(D_{(1,k-1)}) = Q(\Phi_{(1,k-1)}) = 2^{h(k-1)}$, where $h(p)$ denotes the binary entropy function. In [CVZ16] it is shown that $Q(D_{(2,2)}) = Q(\Phi_{(2,2)}) = 2$ which in [AVZ] is extended to, for all even $k \geq 4$, $Q(D_{(k/2,k/2)}) = Q(\Phi_{(k/2,k/2)}) = 2$.

We are going to extend Corollary 4.5. Suppose $\Psi \subseteq I_1 \times I_2 \times I_3$ is not tight, but has a tight subset $\Phi \subseteq \Psi$. In the rest of this section we focus on obtaining a lower bound on $Q(\Psi)$ via $\Phi$. This has an application in the context of tri-colored sum-free sets (Section 4.2.2) for example. We begin with the following standard notion.

**Definition 4.9.** Let $\Phi \subseteq \Psi \subseteq I_1 \times \cdots \times I_k$. We say that $\Phi$ is a *combinatorial degeneration* of $\Psi$, and write $\Psi \triangleright \Phi$, if there are maps $u_i : I_i \to \mathbb{Z}$ ($i \in [k]$) such that for all $\alpha \in I_1 \times \cdots \times I_k$, if $\alpha \in \Psi \setminus \Phi$, then $\sum_{i=1}^k u_i(\alpha_i) > 0$, and if $\alpha \in \Phi$, then $\sum_{i=1}^k u_i(\alpha_i) = 0$. Note that the maps $u_i$ need not be injective.

Combinatorial degeneration gets its name from the following standard proposition, see e.g. [BCS97, Proposition 15.30].

**Proposition 4.10.** Let $f \in V_1 \otimes \cdots \otimes V_k$. Let $\Psi = \supp_C f$. Let $\Phi \subseteq \Psi$ such that $\Psi \triangleright \Phi$. Then $f \triangleright f|\Phi$.

Proposition 4.10 brings us only slightly closer to our goal. Namely, given $f \in V_1 \otimes \cdots \otimes V_k$ with $\Psi = \supp_C f$, and given $\Phi \subseteq \Psi$ such that $\Psi \triangleright \Phi$, it follows directly from Proposition 4.10 that $f \triangleright f|\Phi$ and thus $Q(f) \geq Q(f|\Phi)$. This, however, does not give us a lower bound on $Q(\Psi)$. The following theorem does. Our theorem extends the result in [KSS16].

**Theorem 4.11.** Let $\Phi \subseteq \Psi \subseteq I_1 \times \cdots \times I_k$. If $\Psi \triangleright \Phi$, then

$$Q(\Psi) \geq Q(\Phi).$$

**Lemma 4.12.** Let $S \subseteq T \subseteq I_1 \times \cdots \times I_k$, and $W_i \subseteq I_i$. Let $n \in \mathbb{N}$.

- If $T \supseteq S$, then $T^\times n \supseteq S^\times n$.
- If $S$ is a free diagonal in $T$, then $S^\times n$ is a free diagonal in $T^\times n$.
- If $S$ is a free diagonal in $T$, then $S \cap (W_1 \times \cdots \times W_k)$ is a free diagonal in $T \cap (W_1 \times \cdots \times W_k)$.
- If $S$ is a free diagonal in $T \cap (W_1 \times \cdots \times W_k)$, then $S$ is a free diagonal in $T$.

The proof of Lemma 4.12 is straightforward and left to the reader. First we lower bound $Q(\Psi)$ by the (nonsymptotic) subrank $Q(\Phi)$.

**Lemma 4.13.** Let $\Phi \subseteq \Psi \subseteq I_1 \times \cdots \times I_k$. If $\Psi \triangleright \Phi$, then $Q(\Psi) \geq Q(\Phi)$.  

33
Proof. Pick maps \( u_i : I_i \rightarrow \mathbb{Z} \) such that
\[
\sum_{i=1}^{k} u_i(\alpha_i) = 0 \quad \text{for} \quad \alpha \in \Phi
\]
\[
\sum_{i=1}^{k} u_i(\alpha_i) > 0 \quad \text{for} \quad \alpha \in \Psi \setminus \Phi.
\]
Let \( D \) be a free diagonal in \( \Phi \) with \( |D| = Q(\Phi) \) and let
\[
w_i = \sum_{x \in D_i} u_i(x).
\]
Let \( n \in \mathbb{N} \) and define
\[
W_i = \left\{ (x_1, \ldots, x_{n|D|}) \in I_i^{n|D|} \bigg| \sum_{j=1}^{n|D|} u_i(x_j) = nw_i \right\}.
\]
Then
\[
\Psi^{\times n|D|} \cap (W_1 \times \cdots \times W_k) = \Phi^{\times n|D|} \cap (W_1 \times \cdots \times W_k).
\]
The inclusion \( \supseteq \) is clear. To show \( \subseteq \), let \( (x_1, \ldots, x_k) \in \Psi^{\times n|D|} \cap (W_1 \times \cdots \times W_k) \). Write \( x_i = (x_{i,1}, x_{i,2}, \ldots, x_{i,n|D|}) \) and consider the \( n|D| \times k \) matrix of evaluations
\[
\begin{array}{cccc}
  u_1(x_{1,1}) & u_2(x_{2,1}) & \cdots & u_k(x_{k,1}) \\
  u_1(x_{1,2}) & u_2(x_{2,2}) & \cdots & u_k(x_{k,2}) \\
  \vdots & \vdots & \ddots & \vdots \\
  u_1(x_{1,n|D|}) & u_2(x_{2,n|D|}) & \cdots & u_k(x_{k,n|D|})
\end{array}
\]
The sum of the \( i \)th column is \( nw_i \) by definition of \( W_i \), and \( \sum_{i=1}^{k} nw_i = 0 \). The row sums are nonnegative by definition of the maps \( u_1, \ldots, u_k \). We conclude that the row sums are zero. Therefore \( (x_1, \ldots, x_k) \) is an element of \( \Phi^{\times n|D|} \).

Since \( D \) is a free diagonal in \( \Phi \), \( D^{\times n|D|} \) is a free diagonal in \( \Phi^{\times n|D|} \), and also \( D^{\times n|D|} \cap (W_1 \times \cdots \times W_k) \) is a free diagonal in \( \Phi^{\times n|D|} \cap (W_1 \times \cdots \times W_k) \), which in turn is equal to \( \Psi^{\times n|D|} \cap (W_1 \times \cdots \times W_k) \). Therefore \( D^{\times n|D|} \cap (W_1 \times \cdots \times W_k) \) is also a free diagonal in \( \Psi^{\times n|D|} \), i.e.
\[
Q(\Psi^{\times n|D|}) \geq \left| D^{\times n|D|} \cap (W_1 \times \cdots \times W_k) \right|.
\]
In the set \( D^{\times n|D|} \) consider the strings with uniform type, i.e. where all \( |D| \) elements of \( D \) occur exactly \( n \) times. These are clearly in \( W_1 \times \cdots \times W_k \), and their number is \( \binom{n|D|}{n, \ldots, n} \). Therefore
\[
Q(\Psi^{\times n|D|}) \geq \binom{n|D|}{n, \ldots, n} = |D|^{n|D| - o(n)},
\]
which implies \( Q(\Psi) = \lim_{n \to \infty} Q(\Psi^{\times n|D|})^{1/n} \geq |D| \).

Proof of Theorem 4.11. Clearly, we may write \( Q(\Psi) = \lim_{n \to \infty} Q(\Psi^{\times n})^{1/n} \). By Lemma 4.13,
\[
\lim_{n \to \infty} Q(\Psi^{\times n})^{1/n} \geq \lim_{n \to \infty} Q(\Phi^{\times n})^{1/n}.
\]
The right-hand side equals \( Q(\Phi) \) by definition.
4.2. The cap set problem

In this section we go on a combinatorial excursion. A subset $A \subseteq (\mathbb{Z}/3\mathbb{Z})^n$ is called a cap set if any line in $A$ is a point, a line being a triple of points of the form $(u, u + v, u + 2v)$. The cap set problem is to decide whether the maximal size of a cap set in $(\mathbb{Z}/3\mathbb{Z})^n$ grows like $3^n - o(n)$ or like $c^n + o(n)$ for some $c < 3$. Gijswijt and Ellenberg in [EG17], inspired by the work of Croot, Lev and Pach in [CLP17], settled this problem, showing that $c \leq 3(207 + 33\sqrt{33})^{1/3}/8 \approx 2.755$.

Tao realised in [Tao16] that the cap set problem may naturally be phrased as the problem of computing the size of the largest main diagonal in powers of the cap set tensor

$$
\sum_\alpha e_{\alpha_1} \otimes e_{\alpha_2} \otimes e_{\alpha_3}
$$

where the sum is over $\alpha_1, \alpha_2, \alpha_3 \in F_3$ with $\alpha_1 + \alpha_2 + \alpha_3 = 0$. Here main diagonal refers to a subset $A$ of the basis elements such that restricting the cap set tensor to $A \times A \times A$ gives the tensor $\sum_{v \in A} v \otimes v \otimes v$. We show (in hindsight!) that the cap set tensor is in the $\text{GL}_3(F_3)^{\times 3}$-orbit of the reduced polynomial multiplication tensor, the prime example in [Str91], and we show how all recent results follow from this connection, using Theorem 4.11. We first state Strassen’s result on reduced polynomial multiplication and then the cap set problem.

4.2.1. Reduced polynomial multiplication

Let $F \in \mathbb{F}[x]$ be a polynomial of positive degree $n$. Let $\mathbb{F}[x]/(F)$ be the algebra of univariate polynomials modulo the ideal generated by the polynomial $F$. Reduced polynomial multiplication refers to multiplying in this algebra. For any algebra $A$ with multiplication map $m : A \times A \to A$ the structure tensor is defined as $\sum a_3(m(a_1, a_2)) a_1 \otimes a_2 \otimes a_3$, where the sum goes over elements $a_1, a_2, a_3$ in some basis of $A$. Let $\mathbb{F}[x]/(F)$ denote the structure tensor of the algebra $\mathbb{F}[x]/(F)$.

Let us take $F = x^n$. For example, $\mathbb{F}[x]/(x^2)$ is the tensor

$$
e_0 \otimes e_0 \otimes e_0 + e_0 \otimes e_1 \otimes e_1 + e_1 \otimes e_0 \otimes e_1$$

and $\mathbb{F}[x]/(x^3)$ is the tensor

$$
e_0 \otimes e_0 \otimes e_0 + e_0 \otimes e_1 \otimes e_1 + e_1 \otimes e_0 \otimes e_1 + e_0 \otimes e_2 \otimes e_2 + e_2 \otimes e_0 \otimes e_2 + e_1 \otimes e_1 \otimes e_2$$

and $\mathbb{F}[x]/(x^n)$ is the tensor $\sum_\alpha e_{\alpha_1} \otimes e_{\alpha_2} \otimes e_{\alpha_3}$ where the sum is over $(\alpha_1, \alpha_2, \alpha_3)$ in $\{0, 1, \ldots, n-1\}^3$ such that $\alpha_1 + \alpha_2 = \alpha_3$. The support of $\mathbb{F}[x]/(x^n)$ equals

$$\{(\alpha_1, \alpha_2, \alpha_3) \in \{0, 1, \ldots, n-1\}^3 \mid \alpha_1 + \alpha_2 = \alpha_3\},$$

which via $\alpha_3 \mapsto n - 1 - \alpha_3$ we may identify with the set

$$\Phi_n = \{(\alpha_1, \alpha_2, \alpha_3) \in \{0, 1, \ldots, n-1\}^3 \mid \alpha_1 + \alpha_2 + \alpha_3 = n - 1\}. \quad (30)$$

The support $\Phi_n$ is tight (cf. Example 4.3). We thus already know that the support simplex $\zeta^\Phi([3])(\mathbb{F}[x]/(x^n)) = \{\zeta^\theta(\mathbb{F}[x]/(x^n)) \mid \theta \in \mathcal{P}([3])\}$ is a subset of $\Delta(\mathbb{F}[x]/(x^n))$, and we know that the minimum value of $\zeta^\theta(\mathbb{F}[x]/(x^n))$ over $\theta \in \mathcal{P}([3])$ equals the asymptotic subrank $Q(\Phi_n)$ (Corollary 4.5). Strassen proves
in [Str91, Theorem 6.7] the remarkable result that for $F[x]/(x^n)$ the asymptotic spectrum and the support simplex in fact coincide and are equal to the interval

$$\Delta(F[x]/(x^n)) = \zeta^{P(3)}(F[x]/(x^n)) = [z(n), n]$$

where $z(n)$ is defined as

$$z(n) := \frac{\gamma^n - 1}{\gamma - 1} - 2^{(n-1)/3}$$

with $\gamma$ equal to the unique real positive solution of the equation

$$\frac{1}{\gamma - 1} - \frac{n}{\gamma^{n-1}} = \frac{n-1}{3}.$$ 

In particular, $Q(\Phi_n) = z(n)$. We collect some (rounded) values of $z(n)$ for small $n$ in the following table. For $z(2)$ and $z(3)$ one can write down the exact value. See also [Str91, Table 1].

<table>
<thead>
<tr>
<th>$n$</th>
<th>$z(n)$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>rounded</td>
</tr>
<tr>
<td>2</td>
<td>1.88988   3/2^{2/3} = 2^{h(1/3)}</td>
</tr>
<tr>
<td>3</td>
<td>2.75510   3(207 + 33\sqrt{33})^{1/3}/8</td>
</tr>
<tr>
<td>4</td>
<td>3.61072</td>
</tr>
<tr>
<td>5</td>
<td>4.46158</td>
</tr>
<tr>
<td>6</td>
<td>5.30973</td>
</tr>
<tr>
<td>7</td>
<td>6.15620</td>
</tr>
<tr>
<td>8</td>
<td>7.00155</td>
</tr>
<tr>
<td>9</td>
<td>7.84612</td>
</tr>
<tr>
<td>10</td>
<td>8.69012</td>
</tr>
</tbody>
</table>

Strassen in [Str91, Theorem 6.7] in fact computes the asymptotic spectrum of $F[x]/(F)$ for arbitrary $F$. The result is as follows. Let $F = \beta \prod_{i=1}^{r}(x - \alpha_i)^{n_i}$ be the prime factorisation of $F$ over the algebraic closure of $F$ such that the $\alpha_i$ are pairwise distinct (i.e. the $n_i$ are the multiplicities of the roots of $F$). Then the asymptotic spectrum and support simplex of $F[x]/(F)$ coincide and equal the following interval,

$$\Delta(F[x]/(F)) = \zeta^{P(3)}(F[x]/(F)) = \left[ \sum_{i=1}^{r} z(n_i), n \right].$$

where $z(n_i)$ is the function defined in (31).

Franz Mauch in [Mau98] computes the support simplex for a natural higher-order generalisation of the support $\Phi_n$ and the tensor $F[x]/(x^n)$ which he calls a Hang.

### 4.2.2. The cap set problem

With the asymptotic spectrum of reduced polynomial multiplication at our disposal, we turn to the cap set problem. We begin by reintroducing the problem in the terminology that is commonly used in the literature.

**Definition 4.14.** A three-term progression-free set is a set $A \subseteq (\mathbb{Z}/m\mathbb{Z})^n$ satisfying the following. For all $(x_1, x_2, x_3) \in A \times 3$: there are $u, v \in (\mathbb{Z}/m\mathbb{Z})^n$ such that $(x_1, x_2, x_3) = (u, u+v, u+2v)$ if and only if $x_1 = x_2 = x_3$. Let $r_3((\mathbb{Z}/m\mathbb{Z})^n)$ be the size of the largest three-term progression-free set in $(\mathbb{Z}/m\mathbb{Z})^n$ and define the regularisation $r_3((\mathbb{Z}/m\mathbb{Z})^n) = \lim_{n \to \infty} r_3((\mathbb{Z}/m\mathbb{Z})^n)^{1/n}$. 

36
A three-term progression-free set in $(\mathbb{Z}/3\mathbb{Z})^n$ is called a cap or cap set. The cap set problem is to decide whether $r_3(\mathbb{Z}/3\mathbb{Z})$ equals 3 or is strictly smaller than 3. We next introduce an asymmetric variation on three-term progression free sets, called tri-colored sum-free sets, which are potentially larger. They are interesting since all known upper bound techniques for the size of three-term progression-free set turn out to be upper bounds on the size of tri-colored sum-free sets.

Definition 4.15. Let $G$ be an abelian group. Let $\Gamma \subseteq G \times G \times G$. For $i \in [3]$ we define the marginal sets $\Gamma_i = \{x \in G \mid \exists \alpha \in \Gamma : \alpha_i = x\}$. We say $\Gamma$ is tricolored sum-free if the following holds. The set $\Gamma$ is a diagonal, and for any $\alpha \in \Gamma_1 \times \Gamma_2 \times \Gamma_3$: $\alpha_1 + \alpha_2 + \alpha_3 = 0$ if and only if $\alpha \in \Gamma$. (Recall that $\Gamma \subseteq I_1 \times I_2 \times I_3$ is a diagonal when any two distinct $\alpha, \beta \in \Gamma$ are distinct in all coordinates.) Let $s_3(G)$ be the size of the largest tricolored sum-free set in $G \times G \times G$ and define the regularisation $s_3(\Gamma) = \lim_{n \to \infty} s_3(G^{\times n})^{1/n}$. Equivalently, $\Gamma \subseteq G \times G \times G$ is a tricolored sum-free set if and only if $\Gamma$ is a free diagonal in $\{\alpha \in G \times G \times G \mid \alpha_1 + \alpha_2 + \alpha_3 = 0\}$.

Clearly, if the set $A \subseteq G = (\mathbb{Z}/m\mathbb{Z})^n$ is three-term progression-free, then the set $\Gamma = \{(a, a, -2a) : a \in A\} \subseteq G \times G \times G$ is tri-colored sum-free. Therefore, we have $r_3(\mathbb{Z}/m\mathbb{Z}) \leq s_3(\mathbb{Z}/m\mathbb{Z})$.

Let us briefly summarise the recent history on the cap set problem. For clarity we focus on $m = 3$; we refer the reader to the references for the general results. Edel in [Ede04] proved the lower bound $2.21739 \leq r_3(\mathbb{Z}/3\mathbb{Z})$. Ellenberg and Gijswijt in [EG17] proved the upper bound $r_3(\mathbb{Z}/3\mathbb{Z}) \leq 3(207 + 33\sqrt{33})^{1/3}/8 \approx 2.755$, settling the cap set problem, but leaving open the problem of computing $r_3(\mathbb{Z}/3\mathbb{Z})$. Blasiak et al. [BCC17] proved that in fact $s_3(\mathbb{Z}/3\mathbb{Z}) \leq 3(207 + 33\sqrt{33})^{1/3}/8$.

This upper bound was shown to be an equality in the three papers [KSS16, Nor16, Peb16].

Theorem 4.16. $s_3(\mathbb{Z}/3\mathbb{Z}) = 3(207 + 33\sqrt{33})^{1/3}/8$.

We reprove Theorem 4.16 by proving that $s_3(\mathbb{Z}/m\mathbb{Z})$ equals the asymptotic subrank $z(m)$ of $F_m[x]/(x^m)$ discussed in Section 4.2.1, when $m$ is a prime power. We emphasise that the significance of our proof lies in the explicit connection to the framework of asymptotic spectra and not in the obtained value, which also for prime powers $m$ was already computed in [BCC17, KSS16, Nor16, Peb16].

Proof. We will prove $s_3(\mathbb{Z}/m\mathbb{Z}) = z(m)$ when $m$ is a prime power. By definition $s_3(\mathbb{Z}/m\mathbb{Z})$ is equal to the asymptotic subrank of the set

$$\{\alpha \in \{0, \ldots, m-1\}^3 \mid \alpha_1 + \alpha_2 + \alpha_3 = 0 \text{ mod } m\}$$

which via $\alpha_3 \mapsto \alpha_3 - (m-1)$ we may identify with the set

$$\Psi_m = \{\alpha \in \{0, \ldots, m-1\}^3 \mid \alpha_1 + \alpha_2 + \alpha_3 = m-1 \text{ mod } m\}$$
and so $\mathcal{L}(\mathbb{Z}/m\mathbb{Z}) = Q(\Psi_m)$. Define the tight set (cf. Example 4.3)

$$\Phi_m = \{\alpha \in \{0, \ldots, m-1\}^3 \mid \alpha_1 + \alpha_2 + \alpha_3 = m-1\}.$$ 

We know $Q(\Phi_m) = \mathcal{Z}(m)$ (Section 4.2.1). We will show that $Q(\Phi_m) = Q(\Psi_m)$ when $m$ is a prime power. This proves the theorem.

We first prove $Q(\Phi_m) \leq Q(\Psi_m)$. There is a combinatorial degeneration $\Psi_m \triangleright \Phi_m$. Indeed, let $u_i : \{0, \ldots, m-1\} \rightarrow \{0, \ldots, m-1\}$ be the identity map. If $\alpha \in \Phi_m$, then $\sum_{i=1}^3 u_i(\alpha_i) = m-1$, and if $\alpha \in \Psi_m \setminus \Phi_m$, then $\sum_{i=1}^3 u_i(\alpha_i)$ equals $m-1$ plus a positive multiple of $m$. This means Theorem 4.11 applies, and we thus obtain $Q(\Phi_m) \leq Q(\Psi_m)$. This proves the claim.

We show $Q(\Psi_m) \leq Q(\Phi_m)$ when $m$ is a power of the prime $p$. Let $F = F_p$. Let $f_m \in \mathbb{F}^m \otimes \mathbb{F}^m \otimes \mathbb{F}^m$ have support $\Psi_m$ with all nonzero coefficients equal to 1. Obviously, $Q(\Psi_m) \leq Q(f_m)$. To compute $Q(f_m)$ we show that there is a basis in which the support of $f_m$ equals the tight set $\Phi_m$. Then $Q(f_m) = Q(\Phi_m)$ (Corollary 4.5). This implies the claim. We prepare to give the basis (which is the same basis as used in [BCC+17]). First observe that the rule $x \mapsto \binom{x}{a}$ gives a well-defined map $\mathbb{Z}/m\mathbb{Z} \rightarrow \mathbb{Z}/p\mathbb{Z}$, since for $a \in \{0, 1, \ldots, m-1\}$, if $x \equiv y \pmod{m}$ then $\binom{x}{a} = \binom{y}{a} \pmod{p}$ by Lucas’ theorem. Let $\{e_a\}$ be the standard basis of $\mathbb{F}^m$. The elements $((\sum_{z \in \mathbb{Z}/m\mathbb{Z}} \binom{z}{a})e_x)_{a \in \mathbb{Z}/m\mathbb{Z}}$ form a basis of $\mathbb{F}^m$ since the matrix $((\binom{z}{a}))_{a \in \mathbb{Z}/m\mathbb{Z}}$ is upper triangular with ones on the diagonal. We will now rewrite $f_m$ in the basis $((\sum_{z \in \mathbb{Z}/m\mathbb{Z}} \binom{z}{a})e_x)_{a \in \mathbb{Z}/m\mathbb{Z}}$. We first prove that $Q(\Phi_m)$ equals 1 if and only if $x$ equals $m-1$, and hence

$$f_m = \sum_{x, y, z \in \mathbb{Z}/m\mathbb{Z} : x + y + z = m-1} c_x \otimes c_y \otimes c_z = \sum_{x, y, z \in \mathbb{Z}/m\mathbb{Z}} \binom{x + y + z}{m-1} e_x \otimes e_y \otimes e_z.$$

The identity $\binom{x + y + z}{m-1} = \sum \binom{x}{a} \binom{y}{b} \binom{z}{c}$ with sum over $a, b, c \in \{0, 1, \ldots, m-1\}$ such that $a + b + c = w$ is true and thus

$$\sum_{x, y, z \in \mathbb{Z}/m\mathbb{Z}} \binom{x + y + z}{m-1} e_x \otimes e_y \otimes e_z = \sum_{x, y, z \in \mathbb{Z}/m\mathbb{Z} : a + b + c = m-1} \binom{x}{a} \binom{y}{b} \binom{z}{c} e_x \otimes e_y \otimes e_z. \quad (32)$$

We may simply rewrite (32) as

$$\sum_{a, b, c \in \{0, 1, \ldots, m-1\} : a + b + c = m-1} \binom{x}{a} e_x \otimes \sum_{y \in \mathbb{Z}/m\mathbb{Z}} \binom{y}{b} e_y \otimes \sum_{z \in \mathbb{Z}/m\mathbb{Z}} \binom{z}{c} e_z.$$

Therefore, with respect to the basis $((\sum_{x \in \mathbb{Z}/m\mathbb{Z}} \binom{x}{a})e_x)_{a \in \{0, 1, \ldots, m-1\}}$, the support of $f_m$ equals the tight set $\Phi_m$. (And even stronger, $f_m$ is isomorphic to the tensor $F[x]/(x^m)$ of Section 4.2.1.)

4.3. Free tensors

Let $F$ be the complex numbers $\mathbb{C}$. We consider a family of tensors called free tensors. These were introduced by Franz in [Fra02]. Recall that for any $\theta \in \mathcal{F}_2(B)$
and any tensor $t$

$$E^\theta(t) = E_\theta(t) \quad \text{(Theorem 3.30)}$$

and

$$\rho^\theta(t) \geq E^\theta(t) \quad \text{(Theorem 3.34)}.$$ 

We care about free tensors, because for any free tensor $t$ we can show that the three values $\rho^\theta(t)$, $E^\theta(t)$, $E_\theta(t)$ coincide for $\theta \in \mathcal{P}_s(B)$.

**Definition 4.17** (Free). Let $\Phi \subseteq I_1 \times \cdots \times I_k$. We say $\Phi$ is free if the following holds. For any $x, y \in \Phi$, if $x \neq y$, then the tuples $x$ and $y$ differ in at least two positions. Let $t \in V_1 \otimes \cdots \otimes V_k$. We say $t$ is free if $\text{supp}_C t$ is free for some choice of bases $C \in \mathcal{C}(t)$.

**Remark 4.18.** Clearly, oblique tensors are free. We thus have \{tight\} $\subseteq$ \{oblique\} $\subseteq$ \{free\}. Free tensors from a semigroup under $\otimes$ and $\oplus$, like the tight tensors and oblique tensors.

**Remark 4.19.** We prove there exist tensors that are not free in $\mathbb{C}^n \otimes \mathbb{C}^n \otimes \mathbb{C}^n$ for $n \geq 5$. First we upper bound the maximal size of a free support. Let $\Phi \subseteq [n] \times [n] \times [n]$ be free. Then $|\Phi| = |\{(\alpha_1, \alpha_2) : \alpha \in \Phi\}| \leq n^2$. Second we apply a lemma of Bürigger [Bü90], which is as follows. Let

$$Z_n = \{t \in \mathbb{C}^n \otimes \mathbb{C}^n \otimes \mathbb{C}^n : \exists C \in \mathcal{C}(t) \ |\text{supp}_C t| < n^3 - 3n^2\}.$$ 

Let $Y_n = \mathbb{C}^n \otimes \mathbb{C}^n \otimes \mathbb{C}^n \setminus Z_n$. Then the set $Y_n$ is Zariski open and nonempty. Now let $n \geq 5$ and let $t \in Y_n$. Then $\forall C \in \mathcal{C}(t) \ |\text{supp}_C t| \geq n^3 - 3n^2 > n^2$. We conclude $t$ is not free.

**Theorem 4.20.** Let $t$ be free. Let $\theta \in \mathcal{P}_s(B)$. Then $\rho^\theta(t) = E^\theta(t) = E_\theta(t)$.

**Proof.** By Theorem 3.34 and Theorem 3.30 we have $\rho^\theta(t) \geq E^\theta(t) = E_\theta(t)$. We prove $\rho^\theta(t) \leq E_\theta(t)$. Let $C \in \mathcal{C}(t)$ such that $\text{supp}_C t$ is free. It is sufficient to prove the following claim. For any $P \in \mathcal{P}(\text{supp}_C t)$ the tuple of ordered marginals of $P$ is in the entanglement polytope $\Delta_t$. This result can be found in [Spa98, Fra02, Smi04], see also [Wer13]. Then $\rho^\theta(t) \leq \max_{P \in \mathcal{P}(\text{supp}_C t)} H_\theta(P) \leq E_\theta(t)$. We give a proof of the above claim. Write $C = ((c_{a_1}), \ldots, (c_{a_k}))$. Let $T \subseteq G = S \in \mathbb{N}_1 \times \cdots \times S \in \mathbb{N}_k$ be the corresponding torus. Let $\mu_G$ map any density matrix $\rho$ to its reduced density matrices,

$$\mu_G : \rho \mapsto (\rho_{1}, \ldots, \rho_{k}).$$

Let $\mu_T$ map any density matrix to its reduced density matrices projected onto diagonal matrices by setting the off-diagonal entries to zero,

$$\mu_T : \rho \mapsto (P_{\text{diag}} \rho_{1}, \ldots, P_{\text{diag}} \rho_{k}).$$

The maps $\mu_G$ and $\mu_T$ can be identified with the moment map for $K = S \in \mathbb{N}_1 \times \cdots \times S \in \mathbb{N}_k$ and abelian moment map for $T \subseteq K$ in the sense of [Bri87, Smi04]. We write $\mu_G(\psi)$ for $\mu_G(|\psi\rangle\langle\psi|)$ and we write $\mu_T(\psi)$ for $\mu_T(|\psi\rangle\langle\psi|)$. Let $D_1$ be the set of $k$-tuples of diagonal marginals with ordered diagonal entries. It is a nontrivial fact that $\mu_T(T \cdot t)$ is a convex polytope, and that $\mu_G(G \cdot t) \cap D_1$ is a convex polytope, namely the entanglement polytope $\Delta_t$ [Bri87]. The elements
of $\mu_T(T \cdot t)$ and $\mu_G(G \cdot t) \cap D_4$ are tuples of diagonal matrices whose diagonal entries are a probability distribution. We will identify these elements with these tuples of probability distributions. Let $\Phi = \text{supp}_C t$ be the support of $t$. Let $e_\alpha \in \Phi$. Then

$$e_{\alpha_1} \otimes \cdots \otimes e_{\alpha_k} \in T \cdot t.$$ 

Therefore, the $k$-tuple of point distributions $(\delta_{\alpha_1}, \ldots, \delta_{\alpha_k})$ is in $\mu_T(T \cdot t)$. Since the set $\mu_T(T \cdot t)$ is convex, the convex hull $\text{conv}\{ (\delta_{\alpha_1}, \ldots, \delta_{\alpha_k}) \mid \alpha \in \Phi \}$ is a subset of $\mu_T(T \cdot t)$. Observe that $\text{conv}\{ (\delta_{\alpha_1}, \ldots, \delta_{\alpha_k}) \mid \alpha \in \Phi \}$ equals the set $M = \{(P_1, \ldots, P_k) \mid P \in \mathcal{P}(\text{supp}_C t)\}$. So

$$M \subseteq \mu_T(T \cdot t).$$

Let $M \subseteq M$ consist of the tuples of ordered marginal distributions, then

$$M \subseteq \mu_T(T \cdot t) \cap D_4.$$ 

Now suppose that $\text{supp}_C t$ is free. The freeness implies that the marginal density matrices of $t$ are diagonal matrices and thus $\mu_G(G \cdot t)$ consists of diagonal matrices, i.e. $\mu_T(T \cdot t) = \mu_G(G \cdot t)$. Clearly, $\mu_G(G \cdot t) \subseteq \mu_G(G \cdot t)$. Intersecting with $D_4$ gives

$$M \subseteq \mu_T(T \cdot t) \cap D_4 \subseteq \mu_G(G \cdot t) \cap D_4 = \Delta_t.$$ 

We conclude that for any $P \in \mathcal{P}(\text{supp}_C t)$ the tuple of ordered marginals of $P$ is in the entanglement polytope $\Delta_t$. 

**Remark 4.21.** In the special case of $t$ being oblique, the statement in our Theorem 4.20 corresponds to Corollary 12 in the paper [Str05] of Strassen.

**Example 4.22.** Coppersmith and Winograd in [CW90] obtained the upper bound $\omega \leq 2.41$ on the matrix multiplication exponent $\omega$ by analysing the easy Coppersmith–Winograd tensor, which for $q \geq 1$, is defined as

$$\text{CW}_q = \frac{1}{\sqrt{3q}} \left( \sum_{i=1}^{q} |0\rangle i + |i0\rangle + |ii0\rangle \right) \in \mathbb{C}^{q+1} \otimes \mathbb{C}^{q+1} \otimes \mathbb{C}^{q+1}.$$ 

This tensor is free, so $E_\theta(\text{CW}_q)$ equals $\rho^q(\text{CW}_q)$. We compute for any $\theta \in \mathcal{P}([3])$ the lower bound

$$E_\theta(\text{CW}_q) \geq H \left( \frac{q}{3q} |0\rangle \langle 0 | + \frac{2}{3q} \sum_{j=1}^{q} |j\rangle \langle j | \right)$$

$$= H \left( \frac{1}{3}, \frac{2}{3q}, \ldots, \frac{2}{3q} \right)$$

$$= \frac{3}{4} \log_2 q + h(\frac{1}{3}).$$ 

On the other hand, let $\theta = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ be uniform. Then using the KKT-conditions (see [Str91, Proposition 2.1]) and the recursion property of Shannon entropy we compute the upper bound

$$\rho \left( \frac{1}{3}, \frac{1}{3}, \frac{1}{3} \right)(\text{CW}_q) \leq H \left( \frac{1}{3}, \frac{1}{3}, \frac{1}{3} \right) \left( \{(0, i, i), (i, 0, i), (i, i, 0) : i \in [q]\} \right)$$

$$= H \left( \frac{1}{3}, \frac{2}{3q}, \ldots, \frac{2}{3q} \right)$$

$$= \frac{3}{4} \log_2 q + h(\frac{1}{3}).$$

40
We conclude that \( \rho^\theta(CW_q) \geq \frac{3}{2} \log_2 q + h(\frac{1}{2}) \) with equality when \( \theta \) is uniform. Obviously, \( \rho^\theta(CW_q) \leq \log_2(q+1) \) with equality when \( \theta \) equals \((1,0,0)\) or \((0,1,0)\) or \((0,0,1)\).

Let \( q = 1 \). Then the tensor \( CW_1 \) is tight and therefore oblique, which implies that \( \rho^\theta(CW_1) = \rho^\theta(CW_1) \geq h(\frac{1}{2}) \) with equality when \( \theta \) is uniform. In fact \( CW_1 \) equals the tensor \( D_{(2,1)} \) (the W state) from Example 4.8.

**Remark 4.23.** As shown in Example 4.22, the support and quantum functionals give for all \( q \geq 2 \) the lower bound \( R(CW_q) \geq q + 1 \). If the asymptotic rank \( R(CW_2) \) equals 3, then the exponent of matrix multiplication \( \omega \) equals 2 (see e.g. [BCS97, Remark 15.44] and [Blä13, Section 9]). Bläser and Lysikov in [BL16] obtained the border rank lower bound \( R(CW_q^{\otimes n}) \geq (q + 1)^n + 2^n - 1 \) which asymptotically also implies \( R(CW_q) \geq q + 1 \), and thus leaves open the possibility that \( R(CW_q) = q + 1 \).

**Remark 4.24.** We summarise the known relationships among the support and quantum functionals.

1. \( F^\theta(t) \geq F_\theta(t) \) if \( \theta \in \mathcal{P}(B) \) \hspace{1cm} (Theorem 3.24)
2. \( \zeta^\theta(t) \geq \lim_{n \to \infty} \zeta^\theta(t^{\otimes n})^{1/n} = F^\theta(t) = F_\theta(t) \geq \lim_{n \to \infty} \zeta_\theta(t^{\otimes n})^{1/n} \geq \zeta_\theta(t) \)
   if \( \theta \in \mathcal{P}_s(B) \) \hspace{1cm} (Theorems 3.30, 3.34 and 3.35)
3. \( \zeta^\theta(t) = \zeta_\theta(t) \) if \( \theta \in \mathcal{P}_s(B) \) and \( t \) is oblique \hspace{1cm} (Theorem 2.19)
4. \( \zeta^\theta(t) = F^\theta(t) = F_\theta(t) \geq \zeta_\theta(t) \) if \( \theta \in \mathcal{P}_s(B) \) and \( t \) is free \hspace{1cm} (Theorem 4.20)

### 4.4. Generic tensors

The quantum functionals over \( \mathbb{C} \) and the support functionals over algebraically closed fields, when restricted to tensors of a specific format, have a generic value, i.e. there is a Zariski open set on which the value of the functional is constant. Recall that in an irreducible affine variety (e.g. \( \mathbb{C}^{n_1 \times n_2 \times n_3} \)) the intersection of any two nonempty Zariski open sets is nonempty, and therefore in an irreducible affine variety any nonempty Zariski open set is Zariski dense.

#### 4.4.1. Generic value of the Strassen support functionals

Let \( F \) be algebraically closed. For any format \( (n_1, n_2, n_3) \) there is a nonempty Zariski open subset of \( \mathbb{F}^{n_1 \times n_2 \times n_3} \) on which \( \zeta^\theta \) and \( \zeta_\theta \) each have a constant value. (Indeed, a constructible set \( X \) in an affine variety \( Y \) contains a subset that is open and dense in the closure of \( X \) in \( Y \), see e.g. [Bor91]. There are only finitely many possible supports for tensors in \( V = \mathbb{F}^{n_1 \times n_2 \times n_3} \). The map \( \zeta^\theta \) therefore attains only finitely many values on \( V \) and thus has finitely many fibres, each of which is a constructible set. At least one of these fibres is dense in \( V \) by irreducibility of \( V \). This fibre contains a subset that is open and dense in \( V \). The same argument holds for \( \zeta_\theta \).)

**Definition 4.25.** The above value of \( \zeta^\theta \) and \( \zeta_\theta \) is called **typical** or **generic**, and is denoted by \( \zeta^\theta(n_1, n_2, n_3) \) and \( \zeta_\theta(n_1, n_2, n_3) \) respectively.
Obviously, $0 \leq \zeta_\theta(n_1, n_2, n_3) \leq \zeta_\theta(n_1, n_2, n_3) \leq n_1^{\theta(1)} n_2^{\theta(2)} n_3^{\theta(3)}$.

**Definition 4.26.** Let $\Psi \subseteq [n_1] \times [n_2] \times [n_3]$. We call $\Psi$ *comfortable* if there is a subset $\Phi \subseteq \Psi$ and a probability distribution $P \in \mathcal{P}(\Phi)$ such that $\Phi$ is an antichain and each marginal $P_1$ is uniform on $[n_1]$. We call a format $(n_1, n_2, n_3)$ comfortable if $[n_1] \times [n_2] \times [n_3]$ itself is comfortable. (In $[\text{Tob91}]$ comfortable is called *bequem*.)

**Example 4.27.** The format $(n, n, n)$ is comfortable. Namely, let $\Phi$ be the set $\{(\alpha, \alpha, n + 1 - \alpha) \mid \alpha \in [n]\}$ and let $P \in \mathcal{P}(\Phi)$ be the uniform probability distribution. Clearly $\Phi$ is an antichain and each marginal $P_1$ is uniform on $[n_1]$. The format $(ab, bc, ca)$ is comfortable. Namely let $\Phi$ be the support of the matrix multiplication tensor $(a, b, c)$ and let $P \in \mathcal{P}(\Phi)$ be the uniform probability distribution. The set $\Phi$ is tight and hence an antichain with respect to some product order, and $P_1$ is uniform on $[ab]$, $P_2$ is uniform on $[bc]$ and $P_3$ is uniform on $[ca]$.

**Theorem 4.28 ([$\text{Tob91, Bü90}$]).** Let $\theta \in \mathcal{P}([3])$.

1. $\zeta^\theta_\theta(n_1, n_2, n_3) = n_1^{\theta(1)} n_2^{\theta(2)} n_3^{\theta(3)}$ when $(n_1, n_2, n_3)$ is comfortable.

2. $\zeta^\theta_\theta(n, n, n) = n^{1-\text{min}_i(\theta(i)+\alpha(1))}$ when $n \to \infty$.

Recall that the upper support functional is additive. An interesting byproduct of Bürgisser’s proof of statement 2 (see $[\text{Bü90, Theorem 3.1}]$) is that the lower support functional is not additive.

**Proposition 4.29.** Let $\theta \in \mathcal{P}([3])$ with $\theta_i > 0$ for all $i$. Then $\zeta^\theta_\theta$ is not additive.
**Definition 4.30.** The above value of $F_\theta$ is called typical or generic, and is denoted by $F_\theta(n_1, \ldots, n_k)$.

Obviously, $0 \leq F_\theta(n_1, \ldots, n_k) \leq n_1^{\theta(1)} \cdots n_k^{\theta(k)}$.

We say that $|\psi\rangle \in \mathbb{C}^n_1 \otimes \cdots \otimes \mathbb{C}^n_k$ has completely mixed marginals if for each $i \in [k]$ the marginal quantum entropy $H(|\psi\rangle\langle\psi|)$ equals the maximal value $\log_2 n_i$. Bryan et al. in [BRVR17] give a characterisation of the formats $(n_1, \ldots, n_k)$ for which a state with completely mixed marginals exists. (The earlier characterisation by Klyachko in [Kly02, Corollary 2.5.2.2] appears to be incorrect.)

**Theorem 4.31.** Let $\theta \in \mathcal{P}([k])$ such that $\theta_i > 0$ for all $i$. We have

$$F_\theta(n_1, \ldots, n_k) = n_1^{\theta(1)} \cdots n_k^{\theta(k)}$$

if and only if $\mathbb{C}^n_1 \otimes \cdots \otimes \mathbb{C}^n_k$ contains a pure quantum state with completely mixed marginals.

**Proof.** Let $(n_1, \ldots, n_k)$ be a format such that $\mathbb{C}^n_1 \otimes \cdots \otimes \mathbb{C}^n_k$ contains a pure quantum state with completely mixed marginals. Let $U$ be the Zariski open set defined above. Let $t \in U$. Then the entanglement polytope $\Delta_t$ is the maximal entanglement polytope and thus the maximisation

$$E_\theta(t) = E^\theta(t) = \max \{ \theta(1) H(r_1) + \cdots + \theta(k) H(r_k) \mid (r_1, \ldots, r_k) \in \Delta_t \}$$

achieves the value $\theta(1) \log_2 n_1 + \cdots + \theta(k) \log_2 n_k$. We conclude that $F_\theta(n_1, \ldots, n_k)$ equals the upper bound $n_1^{\theta(1)} \cdots n_k^{\theta(k)}$. On the other hand, if $\mathbb{C}^n_1 \otimes \cdots \otimes \mathbb{C}^n_k$ does not contain a pure quantum state with completely mixed marginals, then $((\frac{1}{n_1}, \ldots, \frac{1}{n_k}), (\frac{1}{n_1}, \ldots, \frac{1}{n_k}))$ is not in the maximal entanglement polytope and thus $\max \{ \theta(1) H(r_1) + \cdots + \theta(k) H(r_k) \mid (r_1, \ldots, r_k) \in \Delta_t \}$ is strictly smaller than $\theta(1) \log_2 n_1 + \cdots + \theta(k) \log_2 n_k$ for $t \in U$ (and in fact for any $t$).

**Remark 4.32.** In particular, $F_\theta(n_1, \ldots, n_k) = n_1^{\theta(1)} \cdots n_k^{\theta(k)}$ for any comfortable format $(n_1, \ldots, n_k)$. Indeed, if $(n_1, \ldots, n_k)$ is comfortable, then by definition there is a set $\Phi \subseteq [n_1] \times \cdots \times [n_k]$ and a probability distribution $P \in \mathcal{P}(\Phi)$ such that $\Phi$ is an antichain and each marginal $P_i$ is uniform on $[n_i]$. Define the tensor

$$t = \sum_{\alpha \in \Phi} \sqrt{P(\alpha)} e_{\alpha_1} \otimes \cdots \otimes e_{\alpha_k} \in \mathbb{C}^n_1 \otimes \cdots \otimes \mathbb{C}^n_k.$$  

Since $\Phi$ is an antichain and thus free, $t$ has completely mixed marginals. Following the proof of Theorem 4.31 the value of $F_\theta(n_1, \ldots, n_k)$ equals $n_1^{\theta(1)} \cdots n_k^{\theta(k)}$.

### 4.4.3. Hilbert–Mumford criterion and instability

We have seen that generically for certain formats $E^\theta$ and $\rho^\theta$ are maximal. We give a related application of the Hilbert–Mumford criterion. Let $t \in \mathbb{C}^n_1 \otimes \cdots \otimes \mathbb{C}^n_k$. Let $\theta \in \mathcal{P}([k])$. We know that

$$E^\theta(t) \leq \rho^\theta(t) \leq n_1^{\theta(1)} \cdots n_k^{\theta(k)}.$$  

Therefore,

$$E^\theta(t) = n_1^{\theta(1)} \cdots n_k^{\theta(k)} \implies \rho^\theta(t) = n_1^{\theta(1)} \cdots n_k^{\theta(k)}.$$  

We show the following converse.
**Theorem 4.33.** Let \( t \in \mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_k} \). Let \( \theta \in \mathcal{P}([k]) \) such that \( \theta_i > 0 \) for all \( i \). If \( E^0(t) < n_1^\theta(1) \cdots n_k^\theta(k) \), then \( \rho^0(t) < n_1^\theta(1) \cdots n_k^\theta(k) \).

We will use terminology and results from [Nes84]. Let the point in the entanglement polytopes corresponding to quantum states that are locally maximally mixed will be called the origin and denoted by \( O \).

**Lemma 4.34.** \( O \notin \mu_G(G \cdot t) \) if and only if \( t \) is \( G \)-unstable i.e. \( 0 \in G \cdot t \).

**Proof.** According to [Nes84, Lemma 2.3], if \( t \) is \( G \)-semistable, then \( G \cdot t \) contains a closed orbit. This implies according to [Nes84, Theorem 2.2(ii)] and [Nes84, Definition 2.1(iii)] that \( O \in \mu_G(G \cdot t) \). If \( t \) is \( G \)-unstable, then all \( G \)-invariants vanish, which implies \( O \notin \mu_G(G \cdot t) \). \( \square \)

**Proof of Theorem 4.33.** Let \( G = \text{SL}(\mathbb{C}^{n_1}) \times \cdots \times \text{SL}(\mathbb{C}^{n_k}) \) act naturally on \( V = \mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_k} \). Let \( t \in V \). If \( E^0(t) < n_1 \theta(1) \cdots n_k \theta(k) \), then \( O \notin \Delta_t \). This means \( t \) is \( G \)-unstable (Lemma 4.34) i.e. \( 0 \in G \cdot t \). By the Hilbert–Mumford criterion [Nes84, Theorem 3.1] this implies that there exists a one-parameter subgroup \( \lambda \) such that \( t \) is \( \lambda \)-unstable. This implies that there exists a maximal torus \( T \subseteq G \) such that \( t \) is \( T \)-unstable. By Lemma 4.34 this implies that \( O \notin \mu_T(T \cdot t) \). Let \( C \subseteq \mathcal{C}(t) \) be a choice of bases compatible with \( T \). As shown in the proof of Theorem 4.20,

\[
M = \{(P_1, \ldots, P_k) \mid P \in \mathcal{P}(\text{supp}_C t)\} \subseteq \mu_T(T \cdot t).
\]

Therefore \( O \notin M \). Noting that \( H_\theta(\text{supp}_C t) \) equals \( n_1 \theta(1) \cdots n_k \theta(k) \) if and only if \( O \in M \), we see that \( H_\theta(\text{supp}_C t) < n_1 \theta(1) \cdots n_k \theta(k) \). Therefore we conclude \( \rho^0(t) \leq \max_{P \in \text{supp}_C t} H_\theta(P) < n_1 \theta(1) \cdots n_k \theta(k) \). \( \square \)

We prove a quantitative version of Theorem 4.33 in terms of instability, which moreover holds over all fields. Our result improves a result of Blasiak et al. in [BCC+17] when the tensor format \((n_1, \ldots, n_k)\) is nonuniform enough.

Instability is a standard notion in geometric invariant theory. Let \( \mathbb{F} \) be algebraically closed. Let \( t \in \mathbb{F}^{n_1} \otimes \cdots \otimes \mathbb{F}^{n_k} \). Let \( G = \text{SL}_{n_1}(\mathbb{F}) \times \cdots \times \text{SL}_{n_k}(\mathbb{F}) \). The tensor \( t \) is called unstable if 0 is in the orbit closure \( G \cdot t \), with the closure taken in the Zariski topology, and otherwise \( t \) is called semistable.

We use the quantitative notion of instability for a tensor \( t \in \mathbb{F}^{n_1} \otimes \cdots \otimes \mathbb{F}^{n_k} \) defined in [BCC+17], with \( \mathbb{F} \) an arbitrary field,

\[
\text{instab}(t) = \max_{C \subseteq \mathcal{C}(t)} \max_{w_1, \ldots, w_k} \sup_{a \in \text{supp}_C t} \left\{ \varepsilon \geq 0 \mid \forall a \in \text{supp}_C t : \sum_{i=1}^k w_i(a_i) \leq \sum_{i=1}^k \left( \frac{1}{|I_i|} \sum_{x \in I_i} w_i(x) - \varepsilon \max_x w_i(x) \right) \right\}
\]

where the first maximum is over the choice of bases of \( \mathbb{F}^{n_1}, \ldots, \mathbb{F}^{n_k} \), with index sets \( I_1, \ldots, I_k \), and the second maximum is over weight functions \( w_i : I_i \to \mathbb{R}_{\geq 0} \) that are not identically zero. When \( \mathbb{F} \) is algebraically closed, the Hilbert–Mumford criterion says that \( t \) is unstable if and only if \( \text{instab}(t) > 0 \).

For probability distributions \( P, Q \in \mathcal{P}(X) \), define the total variation distance \( V(P, Q) = \frac{1}{2} \sum_{x \in X} |P(x) - Q(x)| \) and define the Kullback-Leibler divergence
\(D(P||Q) = \sum_{x \in X} P(x) \log_2(P(x)/Q(x))\). We will use Pinsker’s inequality, which says
\[
V(P, Q) \leq \sqrt{\frac{\ln 2}{2} D(P||Q)},
\] (34)
see e.g. [Tsy09, Lemma 2.5].

**Theorem 4.35.** For \(t \in \mathbb{F}^n_1 \otimes \cdots \otimes \mathbb{F}^n_k\) and \(\theta \in \mathcal{P}([k])\),
\[
\rho(\theta)(t) \leq \sum_{i=1}^k \theta(i) \log_2 n_i - \left(\frac{2}{\min_i \theta(i)}\right) \text{instab}(t)^2.
\] (35)

**Proof.** The two sides of (35) are continuous in \(\theta\), so we may without loss of generality assume that \(\theta(i) > 0\) for all \(i \in [k]\). Choose bases \(C \in \mathcal{C}(t)\), weights \(w_1, \ldots, w_k\), and \(\varepsilon \geq 0\), such that for every \(a \in \text{supp}_C t\)
\[
\sum_{i=1}^k w_i(a_i) \leq \sum_{i=1}^k \left(\frac{1}{|I_i|} \sum_{x \in I_i} w_i(x) - \varepsilon \max_x w_i(x)\right)
\]
i.e. for every \(a \in \text{supp}_C t\)
\[
\varepsilon \sum_{i=1}^k \max_x w_i(x) \leq \sum_{i=1}^k \frac{1}{|I_i|} \sum_{x \in I_i} w_i(x) - \sum_{i=1}^k w_i(a_i).
\]
Let \(P \in \mathcal{P}(\text{supp}_C t)\) be a probability distribution on \(\text{supp}_C t\). Then
\[
\varepsilon \sum_{i=1}^k \max_x w_i(x) \leq \sum_{i=1}^k \frac{1}{|I_i|} \sum_{x \in I_i} w_i(x) - \sum_{a \in \text{supp}_C t} P(a) \sum_{i=1}^k w_i(a_i)
\]
\[
= \sum_{i=1}^k \sum_{x \in I_i} \left(\frac{1}{|I_i|} - P_i(x)\right) w_i(x)
\]
\[
\leq \frac{1}{\varepsilon} \sum_{i=1}^k \sum_{x \in I_i} \left|\frac{1}{|I_i|} - P_i(x)\right| \max_x w_i(x).
\]
Let \(U_i\) be the uniform probability distribution on \(I_i\). We apply Pinsker’s inequality (34) to get
\[
\frac{1}{2} \sum_{i=1}^k \sum_{x \in I_i} \left|\frac{1}{|I_i|} - P_i(x)\right| \max_x w_i(x) \leq \sum_{i=1}^k \sqrt{\frac{\ln 2}{2} D(P_i||U_i)} \max_x w_i(x)
\]
Next we apply the Cauchy-Schwarz inequality to obtain
\[
\sum_{i=1}^{k} \sqrt{\frac{\ln 2}{2} D(P_i||U_i)} \max_x w_i(x)
\leq \sqrt{\frac{\ln 2}{2} \left( \sum_{i=1}^{k} D(P_i||U_i) \right)^{1/2} \left( \sum_{i=1}^{k} \max_x w_i(x)^2 \right)^{1/2}}
\leq \frac{\ln 2}{\min_i \theta(i)} \left( \sum_{i=1}^{k} \theta(i) D(P_i||U_i) \right)^{1/2} \sum_{i=1}^{k} \max_x w_i(x)
= \frac{\ln 2}{\min_i \theta(i)} \left( \sum_{i=1}^{k} \theta(i) |I_i| - \theta(i) H(P_i) \right)^{1/2} \sum_{i=1}^{k} \max_x w_i(x).
\]
We conclude
\[
\epsilon^2 \leq \frac{\ln 2}{2} \frac{1}{\min_i \theta(i)} \left( \sum_{i=1}^{k} \theta(i) |I_i| - \theta(i) H(P_i) \right)
\]
and thus
\[
\sum_{i=1}^{k} \theta(i) H(P_i) \leq \sum_{i=1}^{k} \theta(i) |I_i| - \frac{2}{\ln 2} \min_i \theta(i) \epsilon^2.
\]
Now take the supremum over $P$, and the infimum over $\epsilon$, $w_1$, and $C$. \hfill \Box

5. Subrank, slice rank and multi-slice rank

This section is about two variations on tensor rank, namely slice rank, which was introduced by Tao [Tao16], and multi-slice rank, which was introduced by Naslund [Nas17]. We know that subrank is asymptotically upper bounded by the quantum functionals. It is known that subrank is at most slice rank and multi-slice rank. The result of this section is that, remarkably, even slice rank and multi-slice rank are asymptotically upper bounded by the quantum functionals. In fact, we show that the asymptotic slice rank $\lim_{n \to \infty} \text{slicerank}(t^\otimes n)^{1/n}$ exists and equals the minimum $\min_{\theta \in \mathcal{P}([k])} F^{\theta}(t)$.

5.1. Definition

Recall that a simple tensor is a tensor of the form $v_1 \otimes \cdots \otimes v_k \in V_1 \otimes \cdots \otimes V_k$ with $v_i \in V_i$ for $i \in [k]$, and that the rank $R(t)$ of a tensor $t \in V_1 \otimes \cdots \otimes V_k$ is the smallest number $r$ such that $t$ can be written as a sum of $r$ simple tensors.

**Definition 5.1.** A *slice* is a tensor of the form $v \otimes w \in V_1 \otimes \cdots \otimes V_k$ where $v \in V_j$ and $w \in V_j$ for some $j \in [k]$. A *multi-slice* is a tensor of the form $v \otimes w \in V_1 \otimes \cdots \otimes V_k$ where $v \in V_S$ and $w \in V_S$ for some subset $S \subseteq [k]$. Let $t \in V_1 \otimes \cdots \otimes V_k$. The *slice rank* of $t$, denoted $\text{slicerank}(t)$, is the smallest number $r$ such that $t$ can be written as a sum of $r$ slices. The *multi-slice rank* of $t$, denoted $\text{ multislicerank}(t)$, is the smallest number $r$ such that $t$ can be written as a sum of $r$ multi-slices.
5.2. Key properties

Slice rank and multi-slice rank are clearly $\leq$-monotones. The slice rank and multi-slice rank of $(\tau)$ equals $r$ [Tao16, Nas17]. It follows that subrank is at most multi-slice rank which is at most slice rank,

$$Q(t) \leq \text{multislicerank}(t) \leq \text{slicerank}(t).$$

Computing upper bounds on subrank $Q(t)$ and asymptotic subrank $Q(t)$ was the main motivation for introducing slice rank in [Tao16].

Example 5.2. While tensor rank is easily seen to be sub-multiplicative with respect to tensor products, and subrank to be super-multiplicative, slice rank and multi-slice rank are neither sub-multiplicative nor super-multiplicative. For example, the tensors $\sum_{i=1}^{n} e_i \otimes e_i \otimes 1$, $\sum_{i=1}^{n} e_i \otimes 1 \otimes e_i$, $\sum_{i=1}^{n} 1 \otimes e_i \otimes e_i$ have slice rank one, while their tensor product equals the matrix multiplication tensor $(n, n, n)$ which has slice rank $n^2$ (see [BCC+17, Remark 4.6]). This shows slice rank is not sub-multiplicative. To see slice rank is not super-multiplicative, take for example $W$ to be the tensor $e_1 \otimes e_1 \otimes e_2 + e_1 \otimes e_2 \otimes e_1 + e_2 \otimes e_1 \otimes e_1$. The slice rank of $W$ equals two. The value of the logarithmic upper support functional $\rho^{(1/3, 1/3, 1/3)}(W)$ equals $h(1/3) \approx 0.918296$ (see Example 4.8). In Theorem 5.9 we will show that for any $\theta \in \mathcal{P}(|k|)$ the value of slicerank($t^\otimes n$) is at most $2^{\rho^\theta(t)n+o(n)}$ when $n \to \infty$. We thus obtain the upper bound slicerank($W^\otimes n$) $\leq 2^{h(1/3)n+o(n)}$, which proves the claim. Now it is also clear that slice rank is not equal to subrank or border subrank in general, since subrank and border subrank are super-multiplicative.

Since slice rank and multi-slice rank are not sub-multiplicative and not super-multiplicative, the limit $\lim_{n \to \infty} \text{slicerank}(t^\otimes n)^{1/n}$ and the analogous limit for multi-slice rank might not exist. We will show that $\lim_{n \to \infty} \text{slicerank}(t^\otimes n)^{1/n}$ in fact does exist. For now define

$$\text{SR}^\sim = \lim_{n \to \infty} \sup \text{slicerank}(t^\otimes n)^{1/n},$$

$$\text{MSR}^\sim(t) = \lim_{n \to \infty} \sup \text{multislicerank}(t^\otimes n)^{1/n}.$$

We have

$$Q(t) \leq \text{MSR}^\sim(t) \leq \text{SR}^\sim(t).$$

5.3. Upper bound by support and quantum functionals

We show that slice rank and multi-slice rank are asymptotically upper bounded by the quantum functionals.

Theorem 5.3. Let $\mathbb{F}$ be the complex numbers $\mathbb{C}$. Let $t \in V_1 \otimes \cdots \otimes V_k$. If $\theta \in \mathcal{P}(B)$, then $\text{MSR}^\sim(t) \leq F^\theta(t)$. If $\theta \in \mathcal{P}_B(B)$, then $\text{SR}^\sim(t) \leq F^\theta(t)$.

Proof. Let $c = E^\theta(t)$. Let $\ell = |\text{supp} \theta|$. Let $n \in \mathbb{N}$. Let $t_0 = t^\otimes n$. Define a sequence of tensors $t_1, \ldots, t_\ell$ by

$$t_i = \left(\text{Id} - \sum_{H(\lambda^{(i)}) \leq c} P_{\lambda^{(i)}}\right) t_{i-1}.$$
Let
\[ s_i = \sum P_{\lambda^{(i)}} t_i - 1. \]

We can write \( t_\ell \) in two ways. On the one hand,
\[ t_\ell = t^\otimes n - s_1 - s_2 - \cdots - s_\ell. \]

On the other hand,
\[
\begin{align*}
t_\ell &= \left( \text{Id} - \sum_{\lambda^{(i)} \vdash n: \mu} P_{\lambda^{(i)}} \right) \cdots \left( \text{Id} - \sum_{\lambda^{(i)} \vdash n: \mu} P_{\lambda^{(i)}} \right) t^\otimes n \\
&= \sum_{\lambda^{(i)} \vdash n: \mu} \prod_{1 \leq i \leq \ell} P_{\lambda^{(i)}} t^\otimes n
\end{align*}
\]

which is 0 by definition of \( E^\theta \). Therefore \( t^\otimes n = \sum_{i=1}^\ell s_i \).

The multi-slice rank of an element in the image of \( P_{\lambda} \) is at most \( 2^n H(\lambda) \). Each \( s_i \) is in the image of \( \sum P_{\lambda^{(i)}} \) where the sum is over partitions \( \lambda \vdash n \) with \( H(\lambda) \leq c \) and with at most \( d_1 d_2 \cdots d_k \) parts. There are at most \( (n+1)^{d_1 d_2 \cdots d_k} \) such partitions. Therefore \( \text{multislicerank}(s_i) \leq (n+1)^{d_1 d_2 \cdots d_k} 2^{nc} \). This implies
\[
\mathbf{MSR}^{-}(t) \leq \limsup_{n \to \infty} \left( \ell(n+1)^{d_1 d_2 \cdots d_k} 2^{nc} \right)^{1/n} = F^\theta(t).
\]

This proves the first statement. The proof of the second statement is the same, except that it uses the upper bound \( \text{slicerank}(s_i) \leq (n+1)^{d_1 d_2 \cdots d_k} 2^{nc} \). \( \square \)

Let us focus on the singleton regime \( \theta \in \mathcal{P}([k]) \). We show that the asymptotic slice rank exists and equals the minimum value over \( F^\theta \) with \( \theta \in \mathcal{P}([k]) \).

Lemma 5.4. Let \( t \in \mathbb{C}^{d_1} \otimes \cdots \otimes \mathbb{C}^{d_k} \) be nonzero. For any \( \varepsilon > 0 \) there is an \( n_0 \in \mathbb{N} \) such that for all \( n \geq n_0 \) there are partitions \( \lambda^{(1)}, \ldots, \lambda^{(k)} \vdash n \) satisfying \( (P_{\lambda^{(1)}} \otimes \cdots \otimes P_{\lambda^{(k)}}) t^\otimes n \neq 0 \) and
\[
\min\{H(\frac{1}{n} \lambda^{(1)}), \ldots, H(\frac{1}{n} \lambda^{(k)})\} \geq \min_{\theta \in \mathcal{P}([k])} E^\theta(t) - \varepsilon.
\]

Proof. Let \( \Delta_t \) be the moment polytope of \( t \). By the minimax theorem
\[
\min_{\theta \in \mathcal{P}([k])} E^\theta(t) = \min_{\theta \in \mathcal{P}([k])} \max_{(r_1, \ldots, r_k) \in \Delta_t} \sum_{j=1}^k \theta(j) H(r_j)
\]
\[
= \max_{(r_1, \ldots, r_k) \in \Delta_t} \min_{\theta \in \mathcal{P}([k])} \sum_{j=1}^k \theta(j) H(r_j)
\]
\[
= \max_{(r_1, \ldots, r_k) \in \Delta_t} \min_{j \in [k]} H(r_j).
\]

Let \( \mu^{(1)}, \ldots, \mu^{(k)} \vdash m \) be partitions such that
\[
\min_{j \in [k]} H(\frac{1}{m} \mu^{(j)}) \geq \min_{\theta \in \mathcal{P}([k])} E^\theta(t) - \varepsilon/2
\]

48
and \((P_{\lambda(1)} \otimes \cdots \otimes P_{\lambda(k)}) t^{\otimes n} \neq 0\). This is possible since \(\Delta_t\) is the closure of such rescaled partitions and the entropy is continuous. We use that \((P_{\lambda(1)} \otimes \cdots \otimes P_{\lambda(k)}) t = t \neq 0\) and that the tuples of partitions \(\lambda^{(1)}, \ldots, \lambda^{(k)}\) satisfying \((P_{\lambda(1)} \otimes \cdots \otimes P_{\lambda(k)}) t^{\otimes n} \neq 0\) form a semigroup.

Any \(n \in \mathbb{N}\) can be written as \(n = mq + r\) where \(q, r \in \mathbb{N}, 0 \leq r < m\). Let \(\lambda^{(j)} = q \mu^{(j)} + (r)\). By the semigroup property, \((P_{\lambda^{(1)}} \otimes \cdots \otimes P_{\lambda^{(k)}}) t^{\otimes n} \neq 0\). The entropies can be estimated using the concavity of entropy as

\[
H\left(\frac{1}{n} (q \mu^{(j)} + (r))\right) = H\left(\frac{2m}{n} \frac{1}{m} \mu^{(j)} + \frac{r}{n} \left(\frac{1}{n}\right)\right) \geq \frac{2m}{n} H\left(\frac{1}{m} \mu^{(j)}\right) \geq \frac{1}{n} H\left(\frac{1}{m} \mu^{(j)}\right).
\]

When \(n\) is large enough, this is greater than \(H\left(\frac{1}{m} \mu^{(j)}\right) - \varepsilon/2\). Choose \(n_0\) such that this is true for all \(j \in [k]\).

**Lemma 5.5.** Let \(t \in \mathbb{C}^{d_1} \otimes \cdots \otimes \mathbb{C}^{d_k}\) and suppose that \((P_{\lambda^{(1)}} \otimes \cdots \otimes P_{\lambda^{(k)}}) t^{\otimes n} \neq 0\) for some partitions \(\lambda^{(1)}, \ldots, \lambda^{(k)} \vdash n\). Then

\[
slicerank(t^{\otimes n}) \geq \min\{\dim[\lambda_1], \ldots, \dim[\lambda_k]\}.
\]

**Proof.** \(t^{\otimes n}\) restricts to \((P_{\lambda^{(1)}} \otimes \cdots \otimes P_{\lambda^{(k)}}) t^{\otimes n}\), which by assumption is nonzero. Choose rank one projections \(A_j\) in the vector spaces \(S_{\lambda^{(j)}}(\mathbb{C}^{d_j})\) such that

\[
\left(\left(\text{Id}_{\lambda^{(1)}} \otimes A_1\right) \otimes \cdots \otimes \left(\text{Id}_{\lambda^{(k)}} \otimes A_k\right)\right)(P_{\lambda^{(1)}} \otimes \cdots \otimes P_{\lambda^{(k)}}) t^{\otimes n} \neq 0.
\]

The resulting tensor is invariant under the diagonal \(S_n\)-action, therefore the marginals are maximally mixed. In the GIT language, this means they are unstable, therefore the slice rank is the smallest local dimension by [BCC+17, Theorem 4.6].

**Theorem 5.6.** For any tensor \(t \in \mathbb{C}^{d_1} \otimes \cdots \otimes \mathbb{C}^{d_k}\),

\[
\lim_{n \to \infty} \text{slicerank}(t^{\otimes n})^{1/n} \geq \min_{\theta \in P([k])} 2^{E^\theta(t)}.
\]

**Proof.** Choose \(\varepsilon > 0\) and for all large enough \(n\) choose \(\lambda^{(1)}, \ldots, \lambda^{(k)} \vdash n\) as in Lemma 5.4. By Lemma 5.5,

\[
slicerank(t^{\otimes n}) \geq \min\{\dim[\lambda^{(1)}], \ldots, \dim[\lambda^{(k)}]\} \geq \min\{2^n H\left(\frac{1}{n} \lambda^{(1)}\right), \ldots, 2^n H\left(\frac{1}{n} \lambda^{(k)}\right)\} 2^{-o(n)} \geq 2^{n \min_{\theta \in P([k])} E^\theta(t) - \varepsilon} 2^{-o(n)},
\]

therefore

\[
\liminf_{n \to \infty} \text{slicerank}(t^{\otimes n})^{1/n} \geq 2^{\min_{\theta \in P([k])} E^\theta(t) - \varepsilon}.
\]

This is true for any \(\varepsilon > 0\), therefore

\[
\liminf_{n \to \infty} \text{slicerank}(t^{\otimes n})^{1/n} \geq 2^{\min_{\theta \in P([k])} E^\theta(t)}.
\]
Corollary 5.7. For complex tensors, asymptotic slice rank exists and equals the minimum value of the support functionals over $\theta \in P([k])$,

$$SR(t) := \lim_{n \to \infty} \operatorname{slicerank}(t^\otimes n)^{1/n} = \min_{\theta \in P([k])} 2^{E_{\theta}(t)},$$

Remark 5.8. Connections between asymptotic slice rank and moment polytopes have also been observed in [BGO+17, Corollary 6.5] and [CGN+].

For tensors over arbitrary fields we have the following result. (See also [TS16] for a similar statement.)

**Theorem 5.9.** Let $t \in V_1 \otimes \cdots \otimes V_k$. Let $\theta \in P([k])$. Then

$$\operatorname{SR}^\sim(t) \leq \zeta^\theta(t).$$

**Proof.** Choose a $k$-tuple of bases $C = ([v_1,1, \ldots, v_1,d_1], \ldots, [v_k,1, \ldots, v_k,d_k])$ in $\mathcal{C}(f)$. Let $n \in \mathbb{N}$. As in the proof of Theorem 3.34, given a distribution $Q \in P_n([d_1] \times \cdots \times [d_k])$, define the vector $v_Q \in V_1^\otimes n \otimes \cdots \otimes V_k^\otimes n$ as

$$v_Q := \sum_{m=1}^{n} \bigotimes_{n} v_{1,x,m,1} \otimes v_{2,x,m,2} \otimes \cdots \otimes v_{k,x,m,k}$$

where the sum is over all $n$-tuples $x = (x_1, \ldots, x_n) \in T_Q^n$, and define the element $t_Q \in \mathcal{C}$ as $t_Q := \prod_{i \in \supp Q} (t_{i_1, \ldots, i_k})^{nQ(i)}$. Then the tensor power $t^\otimes n$ can be written as

$$t^\otimes n = \sum_{Q} t_Q v_Q$$

(36)

where the sum is over $Q \in P_n(\supp_C t)$. Let $W_1 \otimes \cdots \otimes W_k \subseteq V_1^\otimes n \otimes \cdots \otimes V_k^\otimes n$ be the subspace where $W_j$ is spanned by all vectors $v_{j,i_1} \otimes \cdots \otimes v_{j,i_k}$ such that the empirical distribution of $(i_1, \ldots, i_k) \in [d_j]^n$ is equal to the marginal distribution $P_{ij}$. Then $v_Q \in W_1 \otimes \cdots \otimes W_k$. This implies that

$$\operatorname{slicerank}(v_Q) \leq \min_{k \in [k]} 2^{nH(Q,j)} \leq 2^{n \sum_{j=1}^{k} \theta(j)H(Q,j)}.$$

The number of terms in (36) is at most $(n+1)^{|\supp_C t|}$. Therefore,

$$\log_2 \operatorname{SR}^\sim(t) \leq \limsup_{n \to \infty} \frac{1}{n} \left( |\supp_C t| \log_2(n+1) + n \max_{P} \sum_{j=1}^{k} \log(p) H(P_j) \right)$$

$$= \max_{P} \sum_{j=1}^{k} \log(p) H(P_j)$$

with both maximisations over $P \in \mathcal{P}(\supp_C t)$. This inequality holds for any basis choice $C \in \mathcal{C}(f)$ and therefore we may minimise the right side over all basis choices. This proves the theorem.

**Corollary 5.10.** Let $t$ be a tight 3-tensor. Then $\operatorname{SR}(t) = \lim_{n \to \infty} \operatorname{slicerank}(t^\otimes n)^{1/n}$ exists and equals the asymptotic subrank $Q(t) = \lim_{n \to \infty} Q(t^\otimes n)^{1/n}$.

**Proof.** We have

$$Q(t) \leq \operatorname{SR}^\sim(t) \leq \min_{\theta} \zeta^\theta(t)$$

and $\min_{\theta} \zeta^\theta(t)$ equals $Q(t)$ since $t$ is a tight 3-tensor (Corollary 4.5).
Remark 5.11. We summarise the relationships among the functionals and asymptotic subrank, rank and slice-rank.

1. \( \tilde{\text{SR}}(t) := \lim_{n \to \infty} \text{slicerank}(t \otimes^n)^{1/n} = \min_{\theta \in \mathcal{P}_s(B)} F^\theta(t) \) \quad (Corollary 5.7)

2. \( \tilde{Q}(t) \leq \text{MSR}^{-}(t) \leq \text{SR}(t) \leq F^\theta(t) \leq \zeta^\theta(t) \) if \( \theta \in \mathcal{P}_s(B) \) (Theorems 5.3 and 5.9)

3. \( \tilde{Q}(t) = \text{MSR}^{-}(t) = \text{SR}(t) = \min_{\theta \in \mathcal{P}_s(B)} F^\theta(t) = \min_{\theta \in \mathcal{P}_s(B)} \zeta^\theta(t) \) if \( t \) tight order-3 (Corollary 5.10)

4. \( \tilde{Q}(t) = \text{MSR}^{-}(t) \leq F^\theta(t) \) if \( \theta \in \mathcal{P}(B) \) (Theorem 5.3)

5. \( \zeta^\theta(t) \leq \tilde{R}(t) \) if \( \theta \in \mathcal{P}_s(B) \)

6. \( F^\theta(t) \leq \tilde{R}(t) \) if \( \theta \in \mathcal{P}_{nc}(B) \).

Acknowledgements. The authors thank the members of the QMATH Tensor reading group and Peter Bürgisser for much valuable discussion. Part of this work was carried out while Jeroen Zuiddam and Péter Vrana were visiting QMATH. We acknowledge financial support from the European Research Council (ERC Grant Agreement no. 337603), the Danish Council for Independent Research (Sapere Aude), and VILLUM FONDEN via the QMATH Centre of Excellence (Grant no. 10059). JZ is supported by NWO (617.023.116).

References


51


Matthias Christandl
Department of Mathematical Sciences, University of Copenhagen,
Universitetsparken 5, 2100 Copenhagen Ø, Denmark.
Email: christandl@math.ku.dk

Péter Vrana
Department of Geometry, Budapest University of Technology and Economics,
Egry József u. 1., 1111 Budapest, Hungary.
Email: vranap@math.bme.hu

Jeroen Zuiddam
Centrum Wiskunde & Informatica and Institute for Logic, Language and Computation, University of Amsterdam,
Science Park 123, 1098 XG Amsterdam, Netherlands.
Email: j.zuiddam@cwi.nl