Algebraic complexity, asymptotic spectra and entanglement polytopes

Jeroen Zuiddam

This is an updated version of my PhD dissertation at the University of Amsterdam which was originally published as ILLC Dissertation Series DS-2018-13 with ISBN 978-94-028-1175-9. This document was compiled on April 6, 2020.

Contents

Preface

This is an updated version of my PhD dissertation. I thank Ronald de Wolf, Sven Polak, Yinan Li, Farrokh Labib, Monique Laurent and Péter Vrana for their helpful suggestions, which have been incorporated in this version.

Péter Vrana, Martijn Zuiddam and Māris Ozols I thank again for proofreading the draft of the dissertation. I thank Jop Briët, Dion Gijswijt, Monique Laurent, Lex Schrijver, Péter Vrana, Matthias Christandl, Maris Ōzols, Michael Walter and Bart Sevenster for helpful discussions regarding the results in Chapter 2 and Chapter 3 of this dissertation. I again thank all my coauthors for the very fruitful collaboration that lead to this dissertation: Harry Buhrman, Matthias Christandl, Péter Vrana, Jop Briët, Chris Perry, Asger Jensen, Markus Bläser, Christian Ikenmeyer, and Karl Bringmann.

Princeton — Jeroen Zuiddam
September, 2018.

Publications

This dissertation is primarily based on the work in the following four papers.

[BIZ17] Karl Bringmann, Christian Ikenmeyer, and Jeroen Zuiddam. On algebraic branching programs of small width.
In Ryan O'Donnell, editor, 32nd Computational Complexity Conference (CCC), 2017.
https://doi.org/10.4230/LIPIcs.CCC.2017.20
https://arxiv.org/abs/1702.05328

Journal of the ACM, volume 65, number 5, article 32, 2018.
https://doi.org/10.1145/3209663

[CVZ16] Matthias Christandl, Péter Vrana, and Jeroen Zuiddam. Asymptotic tensor rank of graph tensors: beyond matrix multiplication.
Computational complexity, 2018.
http://dx.doi.org/10.1007/s00037-018-0172-8
https://arxiv.org/abs/1609.07476

[CVZ18] Matthias Christandl, Péter Vrana, and Jeroen Zuiddam. Universal Points in the Asymptotic Spectrum of Tensors: Extended Abstract.
In Proceedings of 50th Annual ACM SIGACT Symposium on the Theory of Computing (STOC), pages 289–296, 2018.
https://doi.org/10.1145/3188745.3188766
https://arxiv.org/abs/1709.07851

[Zui18] Jeroen Zuiddam. The asymptotic spectrum of graphs and the Shannon capacity.
Combinatorica, to appear.
http://arxiv.org/abs/1807.00169

Chapter 1

Introduction

Volker Strassen published in 1969 his famous algorithm for multiplying any two $n \times n$ matrices using only $\mathcal{O}(n^{2.81})$ rather than $\mathcal{O}(n^3)$ arithmetical operations [Str69]. His discovery marked the beginning of a still ongoing line of research in the field of algebraic complexity theory; a line of research that by now touches several fields of mathematics, including algebraic geometry, representation theory, (quantum) information theory and combinatorics. This dissertation is inspired by and contributes to this line of research.

No further progress followed for almost 10 years after Strassen's discovery, despite the fact that "many scientists understood that discovery as a signal to attack the problem and to push the exponent further down" [Pan84]. Then in 1978 Pan improved the exponent from 2.81 to 2.79 [Pan78, Pan80]. One year later, Bini, Capovani, Lotti and Romani improved the exponent to 2.78 by constructing fast "approximative" algorithms for matrix multiplication and making these algorithms exact via the method of interpolation [BCRL79, Bin80]. Cast in the language of tensors, the result of Bini et al. corresponds to what we now call a "border rank" upper bound. The idea of studying approximative complexity or border complexity for algebraic problems has nowadays become an important theme in algebraic complexity theory.

Schönhage then obtained the exponent 2.55 by constructing a fast algorithm for computing many "disjoint" small matrix multiplications and transforming this into an algorithm for one large matrix multiplication [Sch81]. The upper bound was improved shortly after by works of Pan [Pan81], Romani [Rom82], and Coppersmith and Winograd [CW82], resulting in the exponent 2.50. Then in 1987 Strassen published the laser method with which he obtained the exponent 2.48 [Str87]. The laser method was used in the same year by Coppersmith and Winograd to obtain the exponent 2.38 [CW87]. To do this they invented a method for constructing certain large combinatorial structures. This method, or actually the extended version that Strassen published in [Str91], we now call the Coppersmith–Winograd method. All further improvements on upper bounding the exponent essentially follow the framework of Coppersmith and Winograd, and the improvements do not affect the first two digits after the decimal point [CW90, Sto10, Wil12, LG14].

Define ω to be the smallest possible exponent of n in the cost of any matrix multiplication algorithm. (The precise definition will be given in Section 1.1.) We call ω the exponent of matrix multiplication. To summarise the above historical account on upper bounds: $\omega < 2.38$. On the other hand, the only lower bound we currently have is the trivial lower bound $2 \leq \omega$.

The history of upper bounds on the matrix multiplication exponent ω, which began with Strassen's algorithm and ended with the Strassen laser method and Coppersmith–Winograd method, is well-known and well-documented, see e.g. [BCS97, Section 15.13]. However, there is remarkable work of Strassen on a *theory of lower bounds* for ω and similar types of exponents, and this work has received almost no attention in the literature. This theory of lower bounds is the *theory of asymptotic spectra of tensors* and is the topic of a series of papers by Strassen [Str86, Str87, Str88, Str91, Str05].

In the foregoing, the word tensor has popped up twice—namely, when we mentioned border rank and just now when we mentioned asymptotic spectra of tensors—but we have not discussed at all why tensors should be relevant for understanding the complexity of matrix multiplication. First, we give a mini course on tensors. A k-tensor $t = (t_{i_1,\ldots,i_k})_{i_1,\ldots,i_k}$ is a k-dimensional array of numbers from some field, say the complex numbers $\mathbb{C}$. Thus, a 2-tensor is simply a matrix. A k-tensor is called *simple* if there exist k vectors $v_1, \ldots, v_k$ such that the entries of t are given by the products $t_{i_1,\ldots,i_k} = (v_1)_{i_1} \cdots (v_k)_{i_k}$ for all indices i_j. The *tensor rank* of t is the smallest number r such that t can be written as a sum of r simple tensors. Thus the tensor rank of a 2-tensor is simply its matrix rank. Returning to the problem of finding the complexity of matrix multiplication, there is a special 3-tensor, called the matrix multiplication tensor, that encodes the computational problem of multiplying two 2×2 matrices. This 3-tensor is commonly denoted by $\langle 2, 2, 2\rangle$. It turns out that the matrix multiplication exponent ω is exactly the asymptotic rate of growth of the tensor rank of the "Kronecker powers" of the tensor $\langle 2, 2, 2\rangle$. This important observation follows from the fundamental fact that the computational problem of multiplying matrices is "self-reducible". Namely, we can multiply two matrices by viewing them as block matrices and then performing matrix multiplication at the level of the blocks.

We wrap up this introductory story. To understand the computational complexity of matrix multiplication, one should understand the asymptotic rate of growth of the tensor rank of a certain family of tensors, a family that is obtained by taking powers of a fixed tensor. The theory of asymptotic spectra is the theory of bounds on such asymptotic parameters of tensors.

The main story line of this dissertation concerns the theory of asymptotic spectra. In Section 1.1 of this introduction we discuss in more detail the computational problem of multiplying matrices. In Section 1.2 we discuss the asymptotic spectrum of tensors and discuss a new result: an explicit description of infinitely many elements in the asymptotic spectrum of tensors. In Section 1.3 we consider a new higher-order Coppersmith–Winograd method.

The theory of asymptotic spectra of tensors is a special case of an abstract theory of asymptotic spectra of preordered semirings, which we discuss in Section 1.4. In Section 1.5 we apply this abstract theory to a new setting, namely to graphs. By doing this we obtain a new dual characterisation of the Shannon capacity of graphs.

The second story line of this dissertation is about degeneration, an algebraic kind of approximation related to the concept of border rank of Bini et al. We discuss degeneration in the context of tensors in Section 1.6. There is a combinatorial version of tensor degeneration which we call combinatorial degeneration. We discuss a new result regarding combinatorial degeneration in Section 1.7. Finally, Section 1.8 is about a new result concerning degeneration for algebraic branching programs, an algebraic model of computation.

We finish in Section 1.9 with a discussion of the organisation of this dissertation into chapters.

1.1 Matrix multiplication

In this section we discuss in more detail the computational problem of multiplying two matrices.

Algebraic complexity theory studies algebraic algorithms for algebraic problems. Roughly speaking, algebraic algorithms are algorithms that use only the basic arithmetical operations $+$ and $\times$ over some field, say $\mathbb{R}$ or $\mathbb{C}$. For an overview of the field of algebraic complexity theory the reader should consult [BCS97] and [Sap16]. A fundamental example of an algebraic problem is matrix multiplication.

If we multiply two $n \times n$ matrices by computing the inner products between any row of the first matrix and any column of the second matrix, one by one, we need roughly $2 \cdot n^3$ arithmetical operations ($+$ and $\times$). For example, we can multiply two 2×2 matrices with 12 arithmetical operations, namely 8 multiplications and 4 additions:

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{pmatrix}.$$

Since matrix multiplication is a basic operation in linear algebra, it is worthwhile to see if we can do better than $2 \cdot n^3$. In 1969 Strassen [Str69] published a better algorithm. The base routine of Strassen's algorithm is an algorithm for multiplying two 2×2 matrices with 7 multiplications, 18 additions and certain sign changes:

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} x_1 + x_4 - x_5 + x_7 & x_3 + x_5 \\ x_2 + x_4 & x_1 + x_3 - x_2 + x_6 \end{pmatrix}$$

with

$$
\begin{aligned}
x_1 &= (a_{11} + a_{22})(b_{11} + b_{22}) \\
x_2 &= (a_{21} + a_{22})b_{11} \\
x_3 &= a_{11}(b_{12} - b_{22}) \\
x_4 &= a_{22}(-b_{11} + b_{21}) \\
x_5 &= (a_{11} + a_{12})b_{22} \\
x_6 &= (-a_{11} + a_{21})(b_{11} + b_{12}) \\
x_7 &= (a_{12} - a_{22})(b_{21} + b_{22}).
\end{aligned}
$$

The general routine of Strassen's algorithm multiplies two $n \times n$ matrices by recursively dividing the matrices into four blocks and applying the base routine to multiply the blocks (this is the self-reducibility of matrix multiplication that we mentioned earlier). The base routine does not assume commutativity of the variables for correctness, so indeed we can take the variables to be matrices. After expanding the recurrence we see that Strassen's algorithm uses $4.7{\cdot}n^{\log_2 7} \approx 4.7{\cdot}n^{2.81}$ arithmetical operations. Over the years, Strassen's algorithm was improved by many researchers. The best algorithm known today uses $C \cdot n^{2.38}$ arithmetical operations where C is some constant [CW90, Sto10, Wil12, LG14]. The *exponent of matrix multiplication* ω is defined as the infimum over all real numbers β such that for some constant C_β we can multiply, for any $n \in \mathbb{N}$, any two $n \times n$ matrices with at most $C_\beta \cdot n^\beta$ arithmetical operations. From the above it follows that $\omega \leq 2.38$. From a simple flattening argument it follows that $2 \leq \omega$. We are left with the following well-known open problem: what is the value of the matrix multiplication exponent ω?

The constant C for the currently best algorithm is impractically large (for a discussion of this issue see e.g. [Pan18]). For a practical fast algorithm one should either improve C or find a balance between C and the exponent of n. We will ignore the size of C in this dissertation and focus on the exponent ω.

1.2 The asymptotic spectrum of tensors

We now discuss the theory of asymptotic spectra for tensors.

Let s and t be k-tensors over a field $\mathbb{F}$, $s \in \mathbb{F}^{n_1} \otimes \cdots \otimes \mathbb{F}^{n_k}$, $t \in \mathbb{F}^{m_1} \otimes \cdots \otimes \mathbb{F}^{m_k}$. We say s restricts to t and write $s \geqslant t$ if there are linear maps $A_i : \mathbb{F}^{n_i} \to \mathbb{F}^{m_i}$ such that $(A_1 \otimes \cdots \otimes A_k)(s) = t$. Let $[n] \coloneqq \{1, \ldots, n\}$ for $n \in \mathbb{N}$. We define the product $s \otimes t \in \mathbb{F}^{n_1 m_1} \otimes \cdots \otimes \mathbb{F}^{n_k m_k}$ by $(s \otimes t)_{(i_1,j_1),\ldots,(i_k,j_k)} = s_{i_1,\ldots,i_k} t_{j_1,\ldots,j_k}$ for $i \in [n_1] \times \cdots [n_k]$ and $j \in [m_1] \times \cdots \times [m_k]$. This product generalizes the well-known Kronecker product of matrices. We refer to this product as the tensor (Kronecker) product. We define the direct sum $s \oplus t \in \mathbb{F}^{n_1+m_1} \otimes \cdots \otimes \mathbb{F}^{n_k+m_k}$ by $(s \oplus t)_{\ell_1,\ldots,\ell_k} = s_{\ell_1,\ldots,\ell_k}$ if $\ell \in [n_1] \times \cdots \times [n_k]$, $(s \oplus t)_{n_1+\ell_1,\ldots,n_k+\ell_k} = t_{\ell_1,\ldots,\ell_k}$ if $\ell \in [m_1] \times \cdots \times [m_k]$ and $(s \oplus t)_{\ell_1,\ldots,\ell_k} = 0$ for the remaining indices.

The asymptotic restriction problem asks to compute the infimum of all real numbers $\beta \geq 0$ such that for all $n \in \mathbb{N}$

$$s^{\otimes \beta n + o(n)} \geqslant t^{\otimes n}.$$

We may think of the asymptotic restriction problem as having two directions, namely to find

1. obstructions, "certificates" that prohibit $s^{\otimes \beta n + o(n)} \geqslant t^{\otimes n}$, or
2. constructions, linear maps that carry out $s^{\otimes \beta n + o(n)} \geqslant t^{\otimes n}$.

Ideally, we would like to find matching obstructions and constructions so that we indeed learn the value of β.

What do obstructions look like? We set β equal to one; it turns out that it is sufficient to understand this case. We say s restricts asymptotically to t and write $s \gtrsim t$ if

$$s^{\otimes n + o(n)} \geqslant t^{\otimes n}.$$

What do obstructions look like for asymptotic restriction $\gtrsim$? More precisely: what do obstructions look like for $\gtrsim$ restricted to a subset $S \subseteq \{k\text{-tensors over } \mathbb{F}\}$? Let us assume S is closed under direct sum and tensor product and contains the diagonal tensors $\langle n \rangle \coloneqq \sum_{i=1}^{n} e_i \otimes \cdots \otimes e_i$ for $n \in \mathbb{N}$, where $e_1, \ldots, e_n$ is the standard basis of $\mathbb{F}^n$. Let $\mathbf{X}(S)$ be the set of all maps $\phi : S \to \mathbb{R}_{\geq 0}$ that are

(a) monotone under restriction $\geqslant$,

(b) multiplicative under the tensor Kronecker product $\otimes$,

(c) additive under the direct sum $\oplus$,

(d) normalised to $\phi(\langle n \rangle) = n$ at the diagonal tensor $\langle n \rangle$.

The elements $\phi \in \mathbf{X}(S)$ are called spectral points of S. The set $\mathbf{X}(S)$ is called the asymptotic spectrum of S.

Spectral points $\phi \in \mathbf{X}(S)$ are obstructions! Let $s, t \in S$. If $s \gtrsim t$, then by definition we have a restriction $s^{\otimes n + o(n)} \geqslant t^{\otimes n}$. Then (a) and (b) imply the inequality $\phi(s)^{n+o(n)} = \phi(s^{\otimes n + o(n)}) \geqslant \phi(t^{\otimes n}) = \phi(t)^n$. This implies $\phi(s) \geqslant \phi(t)$. We negate that statement: if $\phi(s) < \phi(t)$ then not $s \gtrsim t$. In that case ϕ is an obstruction to $s \gtrsim t$.

The remarkable fact is that $\mathbf{X}(S)$ is a complete set of obstructions for $\gtrsim$. Namely, for $s, t \in S$ the asymptotic restriction $s \gtrsim t$ holds *if and only if* we have $\phi(s) \geqslant \phi(t)$ for all spectral points $\phi \in \mathbf{X}(S)$. This was proven by Volker Strassen in [Str86, Str88]. His proof uses a theorem of Becker and Schwarz [BS83] which is commonly referred to as the Kadison–Dubois theorem (for historical reasons) or the real representation theorem. (We will say more about this completeness result in Section 1.4.)

Let us introduce tensor rank and subrank, and their asymptotic versions. The tensor rank of t is the size of the smallest diagonal tensor that restricts to t, $\mathrm{R}(t) \coloneqq \min\{r \in \mathbb{N} : t \leqslant \langle r\rangle\}$, and the subrank of t is the size of the largest diagonal tensor to which t restricts, $\mathrm{Q}(t) = \max\{r \in \mathbb{N} : \langle r\rangle \leqslant t\}$. Asymptotic rank is defined as

$$\underset{\sim}{\mathrm{R}}(t) = \lim_{n\to\infty} \mathrm{R}(t^{\otimes n})^{1/n}$$

and asymptotic subrank is defined as

$$\underset{\sim}{\mathrm{Q}}(t) = \lim_{n\to\infty} \mathrm{Q}(t^{\otimes n})^{1/n}.$$

From Fekete's lemma it follows that $\widetilde{\mathrm{Q}}(t) = \sup_n \mathrm{Q}(t^{\otimes n})^{1/n}$ and $\underset{\sim}{\mathrm{R}}(t) = \inf_n \mathrm{R}(t^{\otimes n})^{1/n}$. One easily verifies that every spectral point $\phi \in \mathbf{X}(S)$ is an upper bound on asymptotic subrank and a lower bound on asymptotic rank for any tensor $t \in S$,

$$\underset{\sim}{\mathrm{Q}}(t) \leq \phi(t) \leq \underset{\sim}{\mathrm{R}}(t).$$

Strassen used the completeness of $\mathbf{X}(S)$ for $\lesssim$ to prove $\widetilde{\mathrm{Q}}(t) = \min_{\phi\in\mathbf{X}(S)} \phi(t)$ and $\underset{\sim}{\mathrm{R}}(t) = \max_{\phi\in\mathbf{X}(S)} \phi(t)$. One should think of these expressions as being dual to the defining expressions for $\widetilde{\mathrm{Q}}$ and $\underset{\sim}{\mathrm{R}}$.

We mentioned that Strassen was motivated to study the asymptotic spectrum of tensors by the study of the complexity of matrix multiplication. The precise connection with matrix multiplication is as follows. The matrix multiplication exponent ω is characterised by the asymptotic rank $\underset{\sim}{\mathrm{R}}(\langle 2,2,2\rangle)$ of the matrix multiplication tensor

$$\langle 2,2,2\rangle \coloneqq \sum_{i,j,k\in[2]} e_{ij} \otimes e_{jk} \otimes e_{ki} \in \mathbb{F}^4 \otimes \mathbb{F}^4 \otimes \mathbb{F}^4$$

via $\underset{\sim}{\mathrm{R}}(\langle 2,2,2\rangle) = 2^\omega$. We know the trivial lower bound $2 \leq \omega$, see Section 4.3. We know the (nontrivial) upper bound $\omega \leq 2.3728639$, which is by Coppersmith and Winograd [CW90] and improvements by Stothers [Sto10], Williams [Wil12] and Le Gall [LG14]. It may seem that for the study of matrix multiplication only the asymptotic rank $\underset{\sim}{\mathrm{R}}$ is of interest, and that the asymptotic subrank $\widetilde{\mathrm{Q}}$ is just a toy parameter. Asymptotic subrank, however, plays an important role in the currently best matrix multiplication algorithms. We will discuss this idea in the context of the asymptotic subrank of so-called complete graph tensors in Section 5.5.

The important message is: once we understand the asymptotic spectrum of tensors $\mathbf{X}(S)$, we understand asymptotic restriction $\lesssim$, the asymptotic subrank $\widetilde{\mathrm{Q}}$ and the asymptotic rank $\underset{\sim}{\mathrm{R}}$ of tensors. Of course we should now find an explicit description of $\mathbf{X}(S)$.

Our main result regarding the asymptotic spectrum of tensors is the explicit description of an infinite family of elements in the asymptotic spectrum of all complex tensors $\mathbf{X}(\{\text{complex } k\text{-tensors}\})$, which we call the *quantum functionals* (Chapter 6). Finding such an infinite family has been an open problem since the work of Strassen. Moment polytopes (studied under the name entanglement polytopes in quantum information theory [WDGC13]) play a key role here. To each tensor t is associated a convex polytope $\mathbf{P}(t)$ collecting representation-theoretic information about t, called the moment polytope of t. (See e.g. [Nes84, Bri87, WDGC13, SOK14].) The moment polytope has two important equivalent descriptions.

Quantum marginal spectra description. We begin with the description of $\mathbf{P}(t)$ in terms of quantum marginal spectra.

Let V be a (finite-dimensional) Hilbert space. In quantum information theory, a positive semidefinite hermitian operator $\rho : V \to V$ with trace one is called a density operator. The sequence of eigenvalues of a density operator ρ is a probability vector. We let $\mathrm{spec}(\rho) = (p_1, \ldots, p_n)$ be the sequence of eigenvalues of ρ, ordered non-increasingly, $p_1 \geq \cdots \geq p_n$. Let V_1 and V_2 be Hilbert spaces. Given a density operator ρ on $V_1 \otimes V_2$, the reduced density operator $\rho_1 = \mathrm{tr}_2\, \rho$ is uniquely defined by the property that $\mathrm{tr}(\rho_1 X_1) = \mathrm{tr}(\rho(X_1 \otimes \mathrm{Id}_{V_2}))$ for all operators X_1 on V_1. The operator ρ_1 is again a density operator. The operation tr_2 is called the partial trace over V_2. In an explicit form, ρ_1 is given by $\langle e_i, \rho_1(e_j)\rangle = \sum_\ell \langle e_i \otimes f_\ell, \rho(e_j \otimes f_\ell)\rangle$, where the e_i form a basis of V_1 and the f_i form an orthonormal basis of V_2 (the statement is independent of basis choice).

Let V_i be a Hilbert space and consider the tensor product $V_1 \otimes V_2 \otimes V_3$. Associate with $t \in V_1 \otimes V_2 \otimes V_3$ the dual element $t^* \coloneqq \langle t, \cdot\rangle \in (V_1 \otimes V_2 \otimes V_3)^*$. Then $\rho^t \coloneqq tt^*/\langle t, t\rangle = t\langle t, \cdot\rangle/\langle t, t\rangle$ is a density operator on $V_1 \otimes V_2 \otimes V_3$. Viewing ρ^t as a density operator on the regrouped space $V_1 \otimes (V_2 \otimes V_3)$ we may take the partial trace of ρ^t over $V_2 \otimes V_3$ as described above. We denote the resulting density operator by $\rho_1^t \coloneqq \mathrm{tr}_{23}\, \rho^t$. We similarly define ρ_2^t and ρ_3^t.

Let $V = V_1 \otimes V_2 \otimes V_3$. Let $G = \mathrm{GL}(V_1) \times \mathrm{GL}(V_2) \times \mathrm{GL}(V_3)$ act naturally on V. Let $t \in V \setminus 0$. The moment polytope of t is

$$\mathbf{P}(t) \coloneqq \mathbf{P}(\overline{G \cdot t}) \coloneqq \{(\mathrm{spec}(\rho_1^u), \mathrm{spec}(\rho_2^u), \mathrm{spec}(\rho_3^u)) : u \in \overline{G \cdot t} \setminus 0\}.$$

Here $\overline{G \cdot t}$ denotes the Zariski closure or, equivalently, the Euclidean closure in V of the orbit $G \cdot t = \{g \cdot t : g \in G\}$.

Representation-theoretic description. On the other hand, there is a description of $\mathbf{P}(t)$ in terms of non-vanishing of representation-theoretic multiplicities. We do not state this description here, but stress that it is crucial for our proofs.

Quantum functionals. For any probability vector $\theta \in \mathbb{R}^k$ (i.e. $\sum_{i=1}^k \theta(i) = 1$ and $\theta(i) \geq 0$ for all $i \in [k]$) we define the quantum functional F^θ as an optimisation over the moment polytope:

$$F^\theta(t) \coloneqq \max\bigl\{2^{\sum_{i=1}^k \theta(i)\, H(x^{(i)})} : (x^{(1)}, \ldots, x^{(k)}) \in \mathbf{P}(t)\bigr\}.$$

Here $H(y)$ denotes Shannon entropy of the probability vector y. We prove that F^θ satisfies properties (a), (b), (c) and (d) for all complex k-tensors:

Theorem (Theorem 6.11)**.** $F^\theta \in \mathbf{X}(\{\text{complex } k\text{-tensors}\})$.

To put our result into context: Strassen in [Str91] constructed elements in the asymptotic spectrum of $S = \{\text{oblique } k\text{-tensors over } \mathbb{F}\}$ with the preorder $\leqslant|_S$. The set S is a strict and non-generic subset of all k-tensors over $\mathbb{F}$. These elements we call the (Strassen) support functionals. On oblique tensors over $\mathbb{C}$, the quantum functionals and the support functionals coincide. An advantage of the support functionals over the quantum functionals is that they are defined over any field. In fact, the support functionals are "powerful enough" to reprove the result of Ellenberg and Gijswijt on cap sets [EG17]. We discuss the support functionals in Section 4.4.

1.3 Higher-order CW method

Recall that in the asymptotic restriction problem we have an obstruction direction and a construction direction. The quantum functionals and the support functionals provide obstructions. Now we look at the construction direction. Constructions are asymptotic transformations $s^{\otimes \beta n + o(n)} \geqslant t^{\otimes n}$. We restrict attention to the case that t is a diagonal tensor $\langle r \rangle$. Constructions in this case essentially correspond to lower bounds on the asymptotic subrank $\underset{\sim}{\mathrm{Q}}(s)$. The goal is now to construct good lower bounds on $\underset{\sim}{\mathrm{Q}}(s)$.

Strassen solved the problem of computing the asymptotic subrank for so-called tight 3-tensors with the Coppersmith–Winograd (CW) method and the support functionals [CW90, Str91]. The CW method is combinatorial. Let us introduce the combinatorial viewpoint. Let $I_1, \ldots, I_k$ be finite sets. We call a set $D \subseteq I_1 \times \cdots \times I_k$ a *diagonal* if any two distinct elements $a, b \in D$ differ in all k coordinates. Let $\Phi \subseteq I_1 \times \cdots \times I_k$. We call a diagonal $D \subseteq \Phi$ *free* if $D = \Phi \cap (D_1 \times \cdots \times D_k)$. Here $D_i = \{a_i : a \in D\}$ is the projection of D onto the ith coordinate. The *subrank* $\mathrm{Q}(\Phi)$ of Φ is the size of the largest free diagonal $D \subseteq \Phi$. For two sets $\Phi \subseteq I_1 \times \cdots \times I_k$ and $\Psi \subseteq J_1 \times \cdots \times J_k$ we define the product $\Phi \times \Psi \subseteq (I_1 \times J_1) \times \cdots \times (I_k \times J_k)$ by $\Phi \times \Psi \coloneqq \{((a_1, b_1), \ldots, (a_k, b_k)) : a \in \Phi,\, b \in \Psi\}$. The asymptotic subrank is defined as $\underset{\sim}{\mathrm{Q}}(\Phi) \coloneqq \lim_{n\to\infty} \mathrm{Q}(\Phi^{\times n})^{1/n}$. One may think of Φ as a k-partite hypergraph and of a free diagonal in Φ as an induced k-partite matching.

How does this combinatorial version of subrank relate to the tensor version of subrank that we defined earlier? Let $t \in \mathbb{F}^{n_1} \otimes \cdots \otimes \mathbb{F}^{n_k}$. Expand t in the standard basis, $t = \sum_{i \in [n_1] \times \cdots \times [n_k]} t_i \, e_{i_1} \otimes \cdots \otimes e_{i_k}$. Let $\mathrm{supp}(t)$ be the support of t in the standard basis, $\mathrm{supp}(t) \coloneqq \{i \in [n_1] \times \cdots \times [n_k] : t_i \neq 0\}$. Then $\mathrm{Q}(\mathrm{supp}(t)) \leq \mathrm{Q}(t)$.

We want to construct large free diagonals. Let $\Phi \subseteq I_1 \times \cdots \times I_k$. We call Φ *tight* if there are injective maps $\alpha_i : I_i \to \mathbb{Z}$ such that: if $a \in \Phi$, then $\sum_{i=1}^k \alpha_i(a_i) = 0$. For a set X let $\mathcal{P}(X)$ be the set of probability distributions on X. For $\theta \in \mathcal{P}([k])$ let

$H_\theta(\Phi) \coloneqq \max_{P\in\mathcal{P}(\Phi)} \sum_{i=1}^k \theta(i) H(P_i)$, where $H(P_i)$ denotes the Shannon entropy of the ith marginal distribution of P. In [Str91] Strassen used the CW method and the support functionals to characterise the asymptotic subrank $\underset{\sim}{\mathrm{Q}}(\Phi)$ for tight $\Phi \subseteq I_1 \times I_2 \times I_3$. He proved the following. Let $\Phi \subseteq I_1 \times I_2 \times I_3$ be tight. Then

$$\underset{\sim}{\mathrm{Q}}(\Phi) = \min_{\theta\in\mathcal{P}([3])} 2^{H_\theta(\Phi)} = \max_{P\in\mathcal{P}(\Phi)} \min_{i\in[3]} 2^{H(P_i)}. \tag{1.1}$$

We study the higher-order regime $\Phi \subseteq I_1 \times \cdots \times I_k$, $k \geq 4$:

Theorem (Theorem 5.7)**.** *Let $\Phi \subseteq I_1 \times \cdots \times I_k$ be tight. Then $\underset{\sim}{\mathrm{Q}}(\Phi)$ is lower bounded by an expression that generalizes the right-hand side of* (1.1)*.*

Stating the lower bound requires a few definitions, so we do not state it here. It is not known whether our new lower bound matches the upper bound given by quantum or support functionals.

Using Theorem 5.7 we managed to exactly determine the asymptotic subranks of several new examples. These results in turn we used to obtain upper bounds on the asymptotic rank of so-called complete graph tensors, via a higher-order Strassen laser method.

1.4 Abstract asymptotic spectra

Strassen mainly studied tensors, but he developed an abstract theory of asymptotic spectra in a general setting. In the next section we apply this abstract theory to graphs. We now introduce the abstract theory. One has a semiring S (think of a semiring as a ring without additive inverses) that contains $\mathbb{N}$ and a preorder $\leqslant$ on S that (1) behaves well with respect to the semiring operations, (2) induces the natural order on $\mathbb{N}$, and (3) for any $a, b \in S, b \neq 0$ there is an $r \in \mathbb{N} \subseteq S$ with $a \leqslant r \cdot b$. We call such a preorder a Strassen preorder. The main theorem is that the asymptotic version $\lesssim$ of the Strassen preorder is characterised by the monotone semiring homomorphisms $S \to \mathbb{R}_{\geq 0}$. For $a, b \in S$, let $a \lesssim b$ if there is a sequence $x_n \in \mathbb{N}^{\mathbb{N}}$ with $x_n^{1/n} \to 1$ when $n \to \infty$ and $a^n \leqslant b^n x_n$ for all $n \in \mathbb{N}$. Let

$$\mathbf{X} \coloneqq \mathbf{X}(S, \leqslant) \coloneqq \{\phi \in \mathrm{Hom}(S, \mathbb{R}_{\geq 0}) : \forall a, b \in S \ a \leqslant b \Rightarrow \phi(a) \leq \phi(b)\}.$$

The set $\mathbf{X}$ is called the asymptotic spectrum of $(S, \leqslant)$.

Theorem (Strassen)**.** *$a \lesssim b$ iff $\forall \phi \in \mathbf{X}\ \ \phi(a) \leq \phi(b)$.*

Strassen applies this theorem to study rank and subrank of tensors. We define an abstract notion of rank $\mathrm{R}(a) \coloneqq \min\{n \in \mathbb{N} : a \leqslant n\}$ and an abstract notion of subrank $\mathrm{Q}(a) \coloneqq \max\{m \in \mathbb{N} : m \leqslant a\}$. We then naturally have an asymptotic rank $\underset{\sim}{\mathrm{R}}(a) \coloneqq \lim_{n\to\infty} \mathrm{R}(a^n)^{1/n}$ and (under certain mild conditions) an asymptotic subrank $\underset{\sim}{\mathrm{Q}}(a) \coloneqq \lim_{n\to\infty} \mathrm{Q}(a^n)^{1/n}$. In fact $\underset{\sim}{\mathrm{R}}(a) = \inf_n \mathrm{R}(a^n)^{1/n}$ and $\underset{\sim}{\mathrm{Q}}(a) = \sup_n \mathrm{Q}(a^n)^{1/n}$ by Fekete's lemma. The theorem implies the following dual characterisations.

Corollary (Section 2.8)**.** *If $a \in S$ with $a^k \geqslant 2$ for some $k \in \mathbb{N}$, then*

$$\underset{\sim}{\mathrm{Q}}(a) = \min_{\phi\in\mathbf{X}} \phi(a).$$

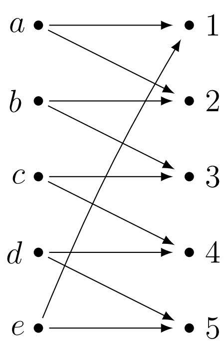
If $a \in S$ with $\phi(a) \geq 1$ for some $\phi \in \mathbf{X}$, then

$$\underset{\sim}{\mathrm{R}}(a) = \max_{\phi\in\mathbf{X}} \phi(a).$$

In Chapter 2 we will discuss the abstract theory of asymptotic spectra. We will discuss a proof of the above theorem that is obtained by integrating the proofs of Strassen in [Str88] and the proof of the Kadison–Dubois theorem of Becker and Schwarz in [BS83]. We will also discuss some basic properties of general asymptotic spectra.

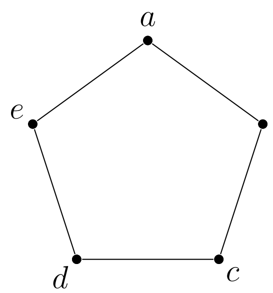
1.5 The asymptotic spectrum of graphs

In the previous section we have seen the abstract theory of asymptotic spectra. We now discuss a problem in graph theory where we can apply this abstract theory. Consider a communication channel with input alphabet $\{a, b, c, d, e\}$ and output alphabet $\{1, 2, 3, 4, 5\}$. When the sender gives an input to the channel, the receiver gets an output according to the following diagram, where an outgoing arrow is picked randomly (say uniformly randomly):



Output 2 has an incoming arrow from a and an incoming arrow from b. We say a and b are confusable, because the receiver cannot know whether a or b was given as an input to the channel. In this channel the pairs of inputs $\{a, b\}, \{b, c\}, \{c, d\}, \{d, e\}, \{e, a\}$ are confusable. If we restrict the input set to a subset of pairwise non-confusable letters, say $\{a, c\}$, then we can use the channel to communicate two messages with zero error. It is clear that for this channel any non-confusable set of inputs has size at most two. Can we make better use of the channel if we use the channel twice? Yes: now the input set is the set of two letter words $\{aa, ab, ac, ad, ae, ba, bb, \ldots\}$, and we have a set of pairwise non-confusable words $\{aa, bc, ce, db, ed\}$, which has size 5. Thus "per channel use" we can send at least $\sqrt{5}$ letters! What happens if we use the channel n times?

The situation is concisely described by drawing the confusability graph of the channel, which has the input letters as vertices and the confusable pairs of input letters as edges. For the above channel the confusability graph is the 5-cycle C_5:



A subset of inputs that are pairwise non-confusable corresponds to a subset of the vertices in the confusability graph that contains no edges, an independent set. The independence number of any graph G is the size of the largest independent set in G, and is denoted by $\alpha(G)$. If G is the confusability graph of some channel, then the confusability graph for using the channel n times is denoted by $G^{\boxtimes n}$ (the graph product $\boxtimes$ is called the strong graph product). The question of how many letters we can send asymptotically translates to computing the limit

$$\Theta(G) \coloneqq \lim_{n\to\infty} \alpha(G^{\boxtimes n})^{1/n},$$

which exists because α is supermultiplicative under $\boxtimes$. The parameter $\Theta(G)$ was introduced by Shannon [Sha56] and is called the Shannon capacity of the graph G. Computing the Shannon capacity is a nontrivial problem already for small graphs. Lovász in 1979 [Lov79] computed the value $\Theta(C_5) = \sqrt{5}$ by introducing and evaluating a new graph parameter ϑ which is now known as the Lovász theta number. Already for the 7-cycle C_7 the Shannon capacity is not known.

Duality theorem. We propose a new application of the abstract theory of asymptotic spectra to graph theory. The main theorem that results from this is a dual characterisation of the Shannon capacity of graphs. For graphs G and H we say $G \leqslant H$ if there is a graph homomorphism $\overline{G} \to \overline{H}$, i.e. from the complement of G to the complement of H. We show graphs are a semiring under the strong graph product $\boxtimes$ and the disjoint union $\sqcup$, and $\leqslant$ is a Strassen preorder on this semiring. The rank in this setting is the clique cover number $\overline{\chi}(\cdot) = \chi(\overline{\cdot})$, i.e. the chromatic number of the complement. The subrank in this setting is the independence number $\alpha(\cdot)$. Let $\mathbf{X}(\mathcal{G})$ be the set of semiring homomorphisms from graphs to $\mathbb{R}_{\geq 0}$ that are monotone under $\leqslant$. From the abstract theory of asymptotic spectra we derive the following duality theorem.

Theorem (Theorem 3.1 (ii))**.** $\Theta(G) = \min_{\phi \in \mathbf{X}(\mathcal{G})} \phi(G)$.

In Chapter 3 we will prove Theorem 3.1 and we will discuss the known elements in $\mathbf{X}(\mathcal{G})$, which are the Lovász theta number and a family of parameters obtained by "fractionalising".

1.6 Tensor degeneration

We move to the second story line that we mentioned earlier: degeneration. Degeneration is a prominent theme in algebraic complexity theory. Roughly speaking, degeneration is an algebraic notion of approximation defined via orbit closures.

For tensors, for example, degeneration is defined as follows. Let V_1, V_2, V_3 be finite-dimensional complex vector spaces and let $V = V_1 \otimes V_2 \otimes V_3$ be the tensor product space. Let $G = \mathrm{GL}(V_1) \times \mathrm{GL}(V_2) \times \mathrm{GL}(V_3)$ act naturally on V. Let $s, t \in V$. Let $G \cdot t = \{g \cdot t : g \in G\}$ be the orbit of t under G. We say t degenerates to s, and write $t \trianglerighteq s$, if s is an element in the orbit closure $\overline{G \cdot t}$. Here the closure is taken with respect to the Zariski topology, or equivalently with respect to the Euclidean topology. One should think of this degeneration $\trianglelefteq$ as a topologically closed version of the restriction preorder $\leq$ for tensors that we defined earlier. Degeneration is a "larger" preorder than restriction in the sense that $s \trianglelefteq t$ implies $s \leq t$.

In several algebraic models of computation, approximative computations correspond to certain degenerations. In some models such an approximative computation can be turned into an exact computation at a small cost, for example using the method of interpolation. The currently fastest matrix multiplication algorithms are constructed in this way, for example.

On the other hand, it turns out that if a lower bound technique for an algebraic measure of complexity is "continuous" then the lower bounds obtained with this technique are already lower bounds on the approximative version of the complexity measure. This observation turns approximative complexity and degeneration into an interesting topic itself. A research program in this direction is the geometric complexity theory program of Mulmuley and Sohoni towards separating the algebraic complexity class **VP** (and related classes) from **VNP** [MS01] (see also [Ike13]).

In this section we briefly discuss three results related to degeneration of tensors that are not discussed further in this dissertation. Then we will discuss results on combinatorial degeneration in Section 1.7 and algebraic branching program degeneration in Section 1.8.

Ratio of tensor rank and border rank. The approximative or degeneration version of tensor rank is called border rank and is denoted by $\underline{\mathrm{R}}$. It has been known since the work of Bini and Strassen that tensor rank R and border rank $\underline{\mathrm{R}}$ are different. How much can they be different? In [Zui17] we showed the following lower bound. Let $k \geq 3$. There is a sequence of k-tensors t_n in $(\mathbb{C}^{2^n})^{\otimes k}$ such that

$\mathrm{R}(t_n)/\underline{\mathrm{R}}(t_n) \geq k - o(1)$ when $n \to \infty$. This answers a question of Landsberg and Michałek [LM16b] and disproves a conjecture of Rhodes [AJRS13]. Further progress will most likely require the construction of explicit tensors with high tensor rank, which has implications in formula complexity [Raz13].

Border support rank. Support rank is a variation on tensor rank, which has its own approximative version called border support rank. A border support rank upper bound for the matrix multiplication tensor yields an upper bound on the asymptotic complexity. This was shown by Cohn and Umans in the context of the group theoretic approach towards fast matrix multiplication [CU13]. They asked: what is the border support rank of the smallest matrix multiplication tensor $\langle 2, 2, 2\rangle$? In [BCZ17a] we showed that it equals seven. Our proof uses the highest-weight vector technique (see also [HIL13]). Our original motivation to study support rank is a connection that we found between support rank and nondeterministic multiparty quantum communication complexity [BCZ17b].

Tensor rank under outer tensor product. We applied degeneration as a tool to study an outer tensor product $\overline{\otimes}$ on tensors. For $s \in \mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_k}$ and $t \in \mathbb{C}^{m_1} \otimes \cdots \otimes \mathbb{C}^{m_\ell}$ let $s \mathbin{\overline{\otimes}} t$ be the natural $(k+\ell)$-tensor in $\mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_k} \otimes \mathbb{C}^{m_1} \otimes \cdots \otimes \mathbb{C}^{m_\ell}$. The products $\otimes$ and $\overline{\otimes}$ differ by a regrouping of the tensor indices. It is well known that tensor rank is not multiplicative under $\otimes$. In [CJZ18] we showed that tensor rank is already not multiplicative under $\overline{\otimes}$, a stronger result. Nonmultiplicativity occurs when taking a power of a tensor whose border rank is strictly smaller than its tensor rank. This answers a question of Draisma [Dra15] and Saptharishi et al. [CKSV16].

1.7 Combinatorial degeneration

In the previous section we introduced the general idea of degeneration and discussed degeneration of tensors. Combinatorial degeneration is the combinatorial analogue of tensor degeneration. Consider sets $\Phi \subseteq \Psi \subseteq I_1 \times \cdots \times I_k$ of k-tuples. We say Φ is a combinatorial degeneration of Ψ and write $\Psi \trianglerighteq \Phi$ if there are maps $u_i : I_i \to \mathbb{Z}$ 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 $\sum_{i=1}^k u_i(\alpha_i) = 0$. We prove that combinatorial asymptotic subrank is nonincreasing under combinatorial degeneration:

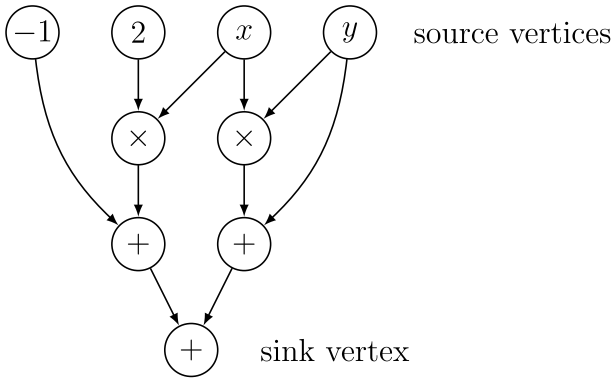
Theorem (Theorem 5.21)**.** *If* $\Psi \trianglerighteq \Phi$, *then* $\underset{\sim}{\mathrm{Q}}(\Psi) \geq \underset{\sim}{\mathrm{Q}}(\Phi)$.

The analogous statement for subrank of tensors is trivially true. The crucial point is that Theorem 5.21 is about combinatorial subrank. As an example, Theorem 5.21 combined with the CW method yields an elegant optimal construction of tri-colored sum-free sets, which are combinatorial objects related to cap sets.

1.8 Algebraic branching program degeneration

We now consider degeneration in the context of algebraic branching programs. A central theme in algebraic complexity theory is the study of the power of different algebraic models of computation and the study of the corresponding complexity classes. We have already (implicitly) used an algebraic model of computation when we discussed matrix multiplication: circuits.

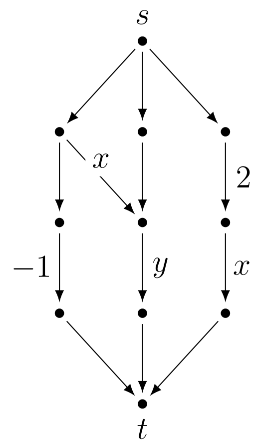
- A circuit is a directed acyclic graph G with one or more source vertices and one sink vertex. Each source vertex is labelled by a variable x_i or a constant $\alpha \in \mathbb{F}$. The other vertices are labelled by either $+$ or $\times$ and have in-degree 2 (that is, fan-in 2). Each vertex of G naturally computes a polynomial. The value of G is the element computed at the sink vertex. The size of G is the number of vertices. (One may also allow multiple sink vertices in order to compute multiple polynomials, e.g. to compute matrix multiplication.) Here is an example of a circuit computing $xy + 2x + y - 1$.



Consider the following two models.

- A formula is a circuit whose graph is a tree.

- An algebraic branching program (abp) is a directed acyclic graph G with one source vertex s, one sink vertex t and affine linear forms over the base field $\mathbb{F}$ as edge labels. Moreover each vertex is labeled with an integer (its layer) and the arrows in the abp point from vertices in layer i to vertices in layer $i + 1$. The cardinality of the largest layer we call the width of the abp. The number of vertices we call the size of the abp. The value of an abp is the sum of the values of all s–t-paths, where the value of an s–t-path is the product of its edge labels. We say that an abp computes its value. Here is an example of a width-3 abp computing $xy + 2x + y - 1$.



The above models of computation give rise to complexity classes. A complexity class consists of families of multivariate polynomials $(f_n)_n = (f(x_1, \ldots, x_{q_n})_n)_{n \in \mathbb{N}}$ over some fixed field $\mathbb{F}$. We say a family of polynomials $(f_n)_n$ is a p-family if the degree of f_n and the number of variables of f_n grow polynomially in n. Let $\mathbf{VP}$ be the class of p-families with polynomially bounded circuit size. Let $\mathbf{VP}_\mathsf{e}$ be the class of p-families with polynomially bounded formula size. For $k \in \mathbb{N}$, let $\mathbf{VP}_k$ be the class of families of polynomials computable by width-k abps of polynomially bounded size. Let $\mathbf{VP}_\mathsf{s}$ be the class of p-families computable by *skew* circuits of polynomial size. Skew circuits are a type of circuits between formulas and general circuits. The class $\mathbf{VP}_\mathsf{s}$ coincides with the class of families of polynomials computable by abps of polynomially bounded size (see e.g. [Sap16]). Ben-Or and Cleve proved that $\mathbf{VP}_3 = \mathbf{VP}_4 = \cdots = \mathbf{VP}_\mathsf{e}$ [BOC92]. Allender and Wang proved $\mathbf{VP}_2 \subsetneq \mathbf{VP}_3$ [AW16]. Thus $\mathbf{VP}_2 \subsetneq \mathbf{VP}_3 = \mathbf{VP}_4 = \cdots = \mathbf{VP}_\mathsf{e} \subseteq \mathbf{VP}_\mathsf{s}$. The following separation problem is one of the many open problems regarding algebraic complexity classes: Is the inclusion $\mathbf{VP}_\mathsf{e} \subseteq \mathbf{VP}_\mathsf{s}$ strict? Motivated by this separation problem we study the approximation closure of $\mathbf{VP}_\mathsf{e}$. We mentioned that Ben-Or and Cleve proved that formula size is polynomially equivalent to width-3 abp size [BOC92]. Regarding width-2, there are explicit polynomials that cannot be computed by any width-2 abp of any size [AW16]. The abp model has a natural notion of approximation. When we allow approximation in our abps, the situation changes completely:

Theorem (Theorem 7.8)**.** *Any polynomial can be approximated by a width-2 abp of size polynomial in the formula size.*

In terms of complexity classes this means $\overline{\mathbf{VP}_2} = \overline{\mathbf{VP}_\mathsf{e}}$, where $\overline{\cdot}$ denotes the "approximation closure" of the complexity class. The theorem suggests an approach regarding the separation of $\mathbf{VP}_\mathsf{e}$ and $\mathbf{VP}_\mathsf{s}$. Namely, superpolynomial lower bounds on formula size may be obtained from superpolynomial lower bounds on approximate width-2 abp size. We moreover study the nondeterminism closure of complexity classes and prove a new characterisation of the complexity class $\mathbf{VNP}$.

1.9 Organisation

This dissertation is divided into chapters as follows. We will begin with the abstract theory of asymptotic spectra in Chapter 2. Then we introduce the asymptotic spectra of graphs and a new characterisation of the Shannon capacity in Chapter 3. In Chapter 4 we introduce the asymptotic spectrum of tensors, discuss the support functionals of Strassen for oblique tensors and a characterisation of asymptotic slice rank of oblique tensors as the minimum over the support functionals. In Chapter 5 we discuss tight tensors, the higher-order Coppersmith–Winograd method, the combinatorial degeneration method and applications to the cap set problem, type sets and graph tensors. In Chapter 6 we introduce an infinite family of elements in the asymptotic spectrum of complex k-tensors and characterise the asymptotic slice rank as the minimum over the quantum functionals. Finally, in Chapter 7 we study algebraic branching programs, and approximation closure and nondeterminism closure of algebraic complexity classes.

Chapter 2

The theory of asymptotic spectra

2.1 Introduction

This is an expository chapter about the abstract theory of asymptotic spectra of Volker Strassen [Str88]. The theory studies semirings S that are endowed with a preorder $\leqslant$. The main result Theorem 2.13 is that under certain conditions, the asymptotic version $\lesssim$ of this preorder is characterised by the semiring homomorphisms $S \to \mathbb{R}_{\geq 0}$ that are monotone under $\leqslant$. These monotone homomorphisms make up the "asymptotic spectrum" of $(S, \leqslant)$. For the elements of S we have natural notions of rank and subrank, generalising rank and subrank of tensors. The asymptotic spectrum gives a dual characterisation of the asymptotic versions of rank and subrank. This dual description may be thought of as a "lower bound" method in the sense of computational complexity theory. In Chapter 3 and Chapter 4 we will study two specific pairs $(S, \leqslant)$.

2.2 Semirings and preorders

A *(commutative) semiring* is a set S with a binary addition operation $+$, a binary multiplication operation $\cdot$, and elements $0, 1 \in S$, such that for all $a, b, c \in S$

(1) $+$ is associative: $(a + b) + c = a + (b + c)$

(2) $+$ is commutative: $a + b = b + a$

(3) $0 + a = a$

(4) $\cdot$ is associative: $(a \cdot b) \cdot c = a \cdot (b \cdot c)$

(5) $\cdot$ is commutative: $a \cdot b = b \cdot a$

(6) $1 \cdot a = a$

(7) $\cdot$ distributes over $+$: $a \cdot (b + c) = (a \cdot b) + (a \cdot c)$

(8) $0 \cdot a = 0$.

As usual we abbreviate $a \cdot b$ as ab. We denote the sum of n times the element $1 \in S$ by $n \in S$. A *preorder* is a relation $\preccurlyeq$ on a set X such that for all $a, b, c \in X$

(1) $\preccurlyeq$ is reflexive: $a \preccurlyeq a$

(2) $\preccurlyeq$ is transitive: $a \preccurlyeq b$ and $b \preccurlyeq c$ implies $a \preccurlyeq c$.

As usual $a \preccurlyeq b$ is the same as $b \succcurlyeq a$. Let $\mathbb{N} \coloneqq \{0, 1, 2, \ldots,\}$ be the set of natural numbers and let $\mathbb{N}_{>0} \coloneqq \{1, 2, \ldots\}$ be the set of strictly-positive natural numbers. We write $\leq$ for the natural order $0 \leq 1 \leq 2 \leq 3 \leq \cdots$ on $\mathbb{N}$.

2.3 Strassen preorders

Let S be a semiring. A preorder $\preccurlyeq$ on S is a *Strassen preorder* if

(1) $\forall n, m \in \mathbb{N}\ \ n \leq m$ in $\mathbb{N}$ iff $n \preccurlyeq m$

(2) $\forall a, b, c, d \in S$ if $a \preccurlyeq b$ and $c \preccurlyeq d$, then $a + c \preccurlyeq b + d$ and $ac \preccurlyeq bd$

(3) $\forall a, b \in S, b \neq 0\ \ \exists r \in \mathbb{N}\ \ a \preccurlyeq rb$.

Recall that for any $n \in \mathbb{N}$ we denote the sum of n times the element $1 \in S$ by $n \in S$. Note that condition (1) implies that $\forall n, m \in \mathbb{N}$ we have $n \neq m$ in $\mathbb{N}$ if and only if $n \neq m$ in S. We thus view $\mathbb{N}$ as a subset of S. Note that condition (2) is equivalent to the condition: $\forall a, b, s \in S$ if $a \preccurlyeq b$, then $a + s \preccurlyeq b + s$ and $as \preccurlyeq bs$.

Let $\preccurlyeq$ be a Strassen preorder on S. Then $0 \preccurlyeq 1$ by condition (1). For $a \in S$, we have $a \preccurlyeq a$ by reflexivity and thus $0 \preccurlyeq a$, by condition (2).

Examples

We give two examples of a semiring with a Strassen preorder. Proofs and formal definitions are given later.

Graphs. Let S be the set of all (isomorphism classes of) finite simple graphs. Let $G, H \in S$. Let $G \sqcup H$ be the disjoint union of G and H. Let $G \boxtimes H$ be the strong graph product of G and H (see Chapter 3). With addition $\sqcup$ and multiplication $\boxtimes$ the set S becomes a semiring. The 0 in S is the graph with no vertices and the 1 in S is the graph with a single vertex. Let $\overline{G}$ be the complement of G. Define a preorder $\leqslant$ on S by $G \leqslant H$ if there is a graph homomorphism $\overline{G} \to \overline{H}$. Then $\leqslant$ is a Strassen preorder. We will investigate this semiring further in Chapter 3.

Tensors. Let $\mathbb{F}$ be a field. Let $k \in \mathbb{N}$. Let S be the set of all k-tensors over $\mathbb{F}$ with arbitrary format, that is, $S = \cup\{\mathbb{F}^{n_1} \otimes \cdots \otimes \mathbb{F}^{n_k} : n_1, \ldots, n_k \in \mathbb{N}\}$. For $s \in \mathbb{F}^{n_1} \otimes \cdots \otimes \mathbb{F}^{n_k}$ and $t \in \mathbb{F}^{m_1} \otimes \cdots \otimes \mathbb{F}^{m_k}$, let $s \leqslant t$ if there are linear maps $A_i : \mathbb{F}^{m_i} \to \mathbb{F}^{n_i}$ with $(A_1 \otimes \cdots \otimes A_k)t = s$. We identify any $s, t \in S$ for which $s \leqslant t$ and $t \leqslant s$. Let $\oplus$ be the direct sum of k-tensors and let $\otimes$ be the tensor product of k-tensors (see Chapter 4). With addition $\oplus$ and multiplication $\otimes$ the set S becomes a semiring. The 0 in S is the zero tensor and the 1 in S is the standard basis element $e_1 \otimes \cdots \otimes e_1 \in \mathbb{F}^1 \otimes \cdots \otimes \mathbb{F}^1$. The preorder $\leqslant$ is a Strassen preorder. We will investigate this semiring further in Chapter 4, Chapter 5, and Chapter 6.

2.4 Asymptotic preorders $\precsim$

Definition 2.1. Let $\preccurlyeq$ be a relation on S. Define the relation $\precsim$ on S by

$$a_2 \precsim a_1 \quad \text{if} \quad \exists (x_N)_N \in \mathbb{N}^{\mathbb{N}} \left(\inf_{N\in\mathbb{N}} x_N^{1/N} = 1 \text{ and } \left(\forall N \in \mathbb{N}\ a_2^N \preccurlyeq a_1^N x_N\right)\right). \tag{2.1}$$

If $\preccurlyeq$ is a Strassen preorder, then we may in (2.1) replace the infimum $\inf_N x_N^{1/N}$ by the limit $\lim_{N\to\infty} x_N^{1/N}$, since we may assume $x_{N+M} \leq x_N x_M$ (if $a_2^N \preccurlyeq a_1^N x_N$ and $a_2^M \preccurlyeq a_1^M x_M$, then $a_2^{N+M} \preccurlyeq a_1^{N+M} x_N x_M$) and then apply Fekete's lemma (Lemma 2.2):

Lemma 2.2 (Fekete's lemma, see [PS98, No. 98])**.** *Let $x_1, x_2, x_3, \ldots \in \mathbb{R}_{\geq 0}$ satisfy $x_{n+m} \leq x_n + x_m$. Then $\lim_{n\to\infty} x_n/n = \inf_n x_n/n$.*

Proof. Let $y = \inf_n x_n/n$. Let $\varepsilon > 0$. Let $m \in \mathbb{N}_{>0}$ with $x_m/m < y + \varepsilon$. Any $n \in \mathbb{N}$ can be written in the form $n = qm + r$ where r is an integer $0 \leq r \leq m - 1$. Set $x_0 = 0$. Then $x_n = x_{qm+r} \leq x_m + x_m + \cdots + x_m + x_r = qx_m + x_r$. Therefore

$$\frac{x_n}{n} = \frac{x_{qm+r}}{qm + r} \leq \frac{qx_m + x_r}{qm + r} = \frac{x_m}{m}\frac{qm}{qm + r} + \frac{x_r}{n}.$$

Thus

$$y \leq \frac{x_n}{n} < (y + \varepsilon)\frac{qm}{n} + \frac{x_r}{n}.$$

The claim follows because $x_r/n \to 0$ and $qm/n \to 1$ when $n \to \infty$. $\square$

For $a_1, a_2 \in S$, if $a_1 \preccurlyeq a_2$ then clearly $a_1 \precsim a_2$.

Lemma 2.3. *Let $\preccurlyeq$ be a Strassen preorder on S. Then $\precsim$ is a Strassen preorder on S.*

We call $\precsim$ the "asymptotic preorder" corresponding to $\preccurlyeq$.

Proof. Let $a, b, c, d \in S$. We verify that $\lesssim$ is a preorder.

First, reflexivity. We have $a \preccurlyeq a$, so $a^N \preccurlyeq a^N \cdot 1$, so $a \lesssim a$.

Second, transitivity. Let $a \lesssim b$ and $b \lesssim c$. This means $a^N \preccurlyeq b^N x_N$ and $b^N \preccurlyeq c^N y_N$ with $x_N^{1/N} \to 1$ and $y_N^{1/N} \to 1$. Then $a^N \preccurlyeq b^N x_N \preccurlyeq c^N x_N y_N$. Since $(x_N y_N)^{1/N} \to 1$, we conclude $a \lesssim c$.

We verify condition (1). Let $n, m \in \mathbb{N}$. If $n \leq m$, then $n \preccurlyeq m$, so $n \lesssim m$. If $n \lesssim m$, then $n^N \preccurlyeq n^M x_N$, so $n^N \leq m^N x_N$, which implies $n \leq m$.

We verify condition (2). Let $a \lesssim b$ and $c \lesssim d$. This means $a^N \preccurlyeq b^N x_N$ and $c^N \preccurlyeq d^N y_N$. Thus $a^N c^N \preccurlyeq b^N d^N x_N y_N$, and so $ac \lesssim bd$. Assume x_N and y_N are nondecreasing (otherwise set $x_N = \max_{n \leq N} x_n$). Then

$$
\begin{aligned}
(a+c)^N = \sum_{m=0}^{N} \binom{N}{m} a^m c^{N-m} &\preccurlyeq \sum_{m=0}^{N} \binom{N}{m} b^m d^{N-m} x_m y_{N-m} \\
&\preccurlyeq \sum_{m=0}^{N} \binom{N}{m} b^m d^{N-m} x_N y_N = (b+d)^N x_N y_N.
\end{aligned}
$$

Thus $a + c \lesssim b + d$.

We verify (3). Let $a, b \in S, b \neq 0$. Then there is an $r \in \mathbb{N}$ with $a \preccurlyeq rb$, and thus $a \lesssim rb$. $\square$

In the following lemma $\mathrel{\underset{\sim}{\lesssim}}$ denotes the asymptotic preorder associated to the asymptotic preorder associated to $\preccurlyeq$.

Lemma 2.4. *Let $\preccurlyeq$ be a Strassen preorder on S. Let $a_1, a_2, b \in S$.*

(i) *If $a_2 + b \preccurlyeq a_1 + b$, then $a_2 \lesssim a_1$.*

(ii) *If $a_2 b \preccurlyeq a_1 b$ with $b \neq 0$, then $a_2 \lesssim a_1$.*

(iii) *If $a_2 \mathrel{\underset{\sim}{\lesssim}} a_1$, then $a_2 \lesssim a_1$.*

(iv) *If $\exists s \in S$ $\forall n \in \mathbb{N}$ $na_2 \preccurlyeq na_1 + s$, then $a_2 \lesssim a_1$.*

Proof. (ii) Let $a_2 b \preccurlyeq a_1 b$. By an inductive argument similar to the argument we used to prove (2.4),

$$
\forall N \in \mathbb{N} \quad a_2^N b \preccurlyeq a_1^N b. \tag{2.2}
$$

Let $m, r \in \mathbb{N}$ with $1 \preccurlyeq mb \preccurlyeq r$. (We use $b \neq 0$.) From (2.2) follows

$$
\forall N \in \mathbb{N} \quad a_2^N \preccurlyeq a_2^N mb \preccurlyeq a_1^N mb \preccurlyeq a_1^N r.
$$

Thus we conclude $a_2 \lesssim a_1$.

(iii) Let $a_2 \mathrel{\underset{\sim}{\lesssim}} a_1$. This means $a_2^N \lesssim a_1^N x_N$ with $x_N^{1/N} \to 1$. This in turn means that $(a_2^N)^M \preccurlyeq (a_1^N x_N)^M y_{N,M}$ with $\forall N \in \mathbb{N}$ $y_{N,M}^{1/M} \to 1$, that is,

$$
a_2^{NM} \preccurlyeq a_1^{NM} x_N^M y_{N,M}.
$$

Choose a sequence $N \mapsto M_N$ such that $(y_{N,M_N})^{1/M_N} \leq 2$, e.g. given N let M_N be the smallest M for which $(y_{N,M})^{1/M} \leq 2$. Then $a_2^{NM_N} \preccurlyeq a_1^{NM_N} x_N^{M_N} y_{N,M_N}$ and

$$(x_N^{M_N} y_{N,M_N})^{1/(NM_N)} = x_N^{1/N} (y_{N,M_N})^{1/(NM_N)} \leq x_N^{1/N} 2^{1/N} \to 1.$$

We conclude $a_2 \lesssim a_1$.

(iv) Let $s \in S$ with $\forall n \in \mathbb{N}\ \ na_2 \preccurlyeq na_1 + s$. We may assume $a_1 \neq 0$. Let $k \in \mathbb{N}$ with $s \preccurlyeq ka_1$. Then

$$\forall n \in \mathbb{N} \quad kna_2 \preccurlyeq kna_1 + ka_1 = ka_1(n+1). \tag{2.3}$$

Apply (ii) to (2.3) to get

$$\forall n \in \mathbb{N} \quad a_2 n \lesssim a_1(n+1).$$

By an inductive argument,

$$\forall N \in \mathbb{N} \quad a_2^N \lesssim a_2^{N-1} a_1 2 \lesssim a_2^{N-2} a_1^2 3 \lesssim \cdots \lesssim a_1^N (N+1).$$

Since $(N+1)^{1/N} \to 1$, $a_2 \mathrel{\overset{\approx}{\lesssim}} a_1$. From (iii) follows $a_2 \lesssim a_1$.

(i) Let $a_2 + b \preccurlyeq a_1 + b$. We first prove

$$\forall q \in \mathbb{N} \quad qa_2 + b \preccurlyeq qa_1 + b. \tag{2.4}$$

By assumption the statement is true for $q = 1$; suppose the statement is true for $q - 1$, then

$$\begin{aligned} qa_2 + b &= (q-1)a_2 + (a_2 + b) \preccurlyeq (q-1)a_2 + (a_1 + b) \\ &= ((q-1)a_2 + b) + a_1 \preccurlyeq ((q-1)a_1 + b) + a_1 = qa_1 + b, \end{aligned}$$

which proves the statement by induction. Then $\forall n \in \mathbb{N}\ \ na_2 \preccurlyeq na_1 + b$. From (iv) follows $a_2 \lesssim a_1$. □

2.5 Maximal Strassen preorders

Let $\mathcal{P}$ be the set of Strassen preorders on S. For $\preccurlyeq_1, \preccurlyeq_2 \in \mathcal{P}$ we write $\preccurlyeq_2 \subseteq \preccurlyeq_1$ if for all $a, b \in S$: $a \preccurlyeq_2 b$ implies $a \preccurlyeq_1 b$. (The notation $\preccurlyeq_2 \subseteq \preccurlyeq_1$ is natural if we think of the relations $\preccurlyeq_i$ as sets of pairs (a, b) with $a \preccurlyeq_i b$.)

Lemma 2.5. *Let $\preccurlyeq \in \mathcal{P}$ with $\preccurlyeq = \lesssim$ and $a_2 \not\preccurlyeq a_1$. Then there is an element $\preccurlyeq_{a_1 a_2} \in \mathcal{P}$ with $\preccurlyeq \subseteq \preccurlyeq_{a_1 a_2}$ and $a_1 \preccurlyeq_{a_1 a_2} a_2$.*

Proof. For $x_1, x_2 \in S$, let

$$x_1 \preccurlyeq_{a_1 a_2} x_2 \quad \text{if} \quad \exists s \in S \quad x_1 + sa_2 \preccurlyeq x_2 + sa_1.$$

The relation $\preccurlyeq_{a_1a_2}$ is reflexive, since $x + 0 \cdot a_2 \preccurlyeq x + 0 \cdot a_1$. The relation $\preccurlyeq_{a_1a_2}$ is transitive: if $x_1 \preccurlyeq_{a_1a_2} x_2$ and $x_2 \preccurlyeq_{a_1a_2} x_3$, then $x_1 + sa_2 \preccurlyeq x_2 + sa_1$ and $x_2 + ta_2 \preccurlyeq x_3 + ta_1$ for some $s, t \in S$, and so $x_1 + (t+s)a_2 \preccurlyeq x_2 + ta_2 + sa_1 \preccurlyeq x_3 + ta_1 + sa_1 = x_3 + (t+s)a_1$. Thus $x_1 \preccurlyeq_{a_1a_2} x_3$. We conclude that $\preccurlyeq_{a_1a_2}$ is a preorder on S.

We prove that $\preccurlyeq_{a_1a_2}$ is a Strassen preorder. If $x_1 \preccurlyeq_{a_1a_2} x_2$ and $y_1 \preccurlyeq_{a_1a_2} y_2$, then clearly $x_1 + y_1 \preccurlyeq_{a_1a_2} x_2 + y_2$. If $x_1 \preccurlyeq_{a_1a_2} x_2$ and $y \in S$, then $x_1 y \preccurlyeq_{a_1a_2} x_2 y$. From this follows: if $x_1 \preccurlyeq_{a_1a_2} x_2$ and $y_1 \preccurlyeq_{a_1a_2} y_2$, then $x_1 y_2 \preccurlyeq_{a_1a_2} x_2 y_2$.

Let $n, m \in \mathbb{N}$. If $n \leq m$, then $n \preccurlyeq m$, so $n \preccurlyeq_{a_1a_2} m$. If $n \not\leq m$, then $n \geq m+1$. Suppose $n \preccurlyeq_{a_1a_2} m$. Let $s \in S$ with $n + sa_2 \preccurlyeq m + sa_1$. Adding $m + 1 \preccurlyeq n$ gives

$$m + 1 + n + sa_2 \preccurlyeq n + m + sa_1.$$

Since $\preccurlyeq \,=\, \gtrsim$ we may apply Lemma 2.4 (i) to obtain

$$1 + sa_2 \preccurlyeq sa_1. \tag{2.5}$$

From (2.5) follows $s \neq 0$. From (2.5) also follows

$$sa_2 \preccurlyeq sa_1. \tag{2.6}$$

Since $\preccurlyeq \,=\, \gtrsim$ we may apply Lemma 2.4 (ii) to (2.6) to obtain the contradiction

$$a_2 \preccurlyeq a_1.$$

Therefore, $n \not\preccurlyeq_{a_1a_2} m$. We conclude that $\preccurlyeq_{a_1a_2}$ is a Strassen preorder, that is, $\preccurlyeq \in \mathcal{P}$.

Finally, we have $a_1 \preccurlyeq_{a_1a_2} a_2$, since $a_1 + 1 \cdot a_2 \preccurlyeq a_2 + 1 \cdot a_1$. Also, if $x_1 \preccurlyeq x_2$, then $x_1 + 0 \cdot a_2 \preccurlyeq x_2 + 0 \cdot a_1$, that is, $\preccurlyeq \,\subseteq\, \preccurlyeq_{a_1a_2}$. □

Let $\preccurlyeq$ be a Strassen preorder. Let $\mathcal{P}_{\preccurlyeq}$ be the set of Strassen preorders containing $\preccurlyeq$ ordered by inclusion $\subseteq$. Let $\mathcal{C} \subseteq \mathcal{P}_{\preccurlyeq}$ be any chain. Then the union of all preorders in $\mathcal{C}$ is an element of $\mathcal{P}_{\preccurlyeq}$ and contains all elements of $\mathcal{C}$. Therefore, by Zorn's lemma, $\mathcal{P}_{\preccurlyeq}$ contains a maximal element (maximal with respect to inclusion $\subseteq$).

Lemma 2.6. *Let $\preccurlyeq$ be maximal in $\mathcal{P}$. Then $\preccurlyeq \,=\, \gtrsim$.*

Proof. Trivially $\preccurlyeq \,\subseteq\, \gtrsim$. From Lemma 2.3 we know $\gtrsim \in \mathcal{P}$. From maximality of $\preccurlyeq$ follows $\preccurlyeq \,=\, \gtrsim$. □

A relation $\preccurlyeq$ on S is *total* if: for all $a, b \in S$, $a \preccurlyeq b$ or $b \preccurlyeq a$.

Lemma 2.7. *Let $\preccurlyeq$ be maximal in $\mathcal{P}$. Then $\preccurlyeq$ is total.*

Proof. Suppose $\preccurlyeq$ is not total, say $a_1 \not\preccurlyeq a_2$ and $a_2 \not\preccurlyeq a_1$. By Lemma 2.5 there is an element $\preccurlyeq_{a_1a_2} \in \mathcal{P}$ with $\preccurlyeq \,\subseteq\, \preccurlyeq_{a_1a_2}$ and $a_1 \preccurlyeq_{a_1a_2} a_2$. Then $\preccurlyeq$ is strictly contained in $\preccurlyeq_{a_1a_2}$, which contradicts the maximality of $\preccurlyeq$. We conclude $\preccurlyeq$ is total. □

2.6 The asymptotic spectrum $\mathbf{X}(S, \leqslant)$

Definition 2.8. Let S be a semiring with $\mathbb{N} \subseteq S$ and let $\leqslant$ be a Strassen preorder on S. Let

$$\mathbf{X}(S, \leqslant) := \{\phi \in \mathrm{Hom}(S, \mathbb{R}_{\geq 0}) : a \leqslant b \Rightarrow \phi(a) \leq \phi(b)\}.$$

We call $\mathbf{X}(S, \leqslant)$ the *asymptotic spectrum* of $(S, \leqslant)$. We call the elements of $\mathbf{X}(S, \leqslant)$ *spectral points.*

Lemma 2.9. *Let $\preccurlyeq \in \mathcal{P}$. For $a \in S$ define*

$$\begin{aligned}
&\phi(a) := \inf\{\tfrac{r}{s} : r, s \in \mathbb{N}, sa \preccurlyeq r\}, \\
&\psi(a) := \sup\{\tfrac{u}{v} : u, v \in \mathbb{N}, u \preccurlyeq va\}.
\end{aligned}$$

Then for $a, b \in S$ holds

$$\begin{aligned}
&\phi(ab) \leq \phi(a)\phi(b), \\
&\phi(a + b) \leq \phi(a) + \phi(b), \\
&a \preccurlyeq b \Rightarrow \phi(a) \leq \phi(b), \\
&\phi(1) = 1, \\
&\phi(0) = 0
\end{aligned}$$

and

$$\begin{aligned}
&\psi(ab) \geq \psi(a)\psi(b), \\
&\psi(a + b) \geq \psi(a) + \psi(b), \\
&a \preccurlyeq b \Rightarrow \psi(a) \leq \psi(b), \\
&\psi(1) = 1, \\
&\psi(0) = 0.
\end{aligned}$$

and

$$\psi(a) \leq \phi(a).$$

If $\preccurlyeq$ is total, then $\phi = \psi$ and ϕ is the unique semiring homomorphism $S \to \mathbb{R}_{\geq 0}$ that is monotone under $\preccurlyeq$. If $\preccurlyeq$ is maximal in $\mathcal{P}$, then $a \preccurlyeq b \Leftrightarrow \phi(a) \leq \phi(b)$.

Proof. Let $a, b \in S$. We prove $\phi(a + b) \leq \phi(a) + \phi(b)$. Let $s_a, s_b, r_a, r_b \in \mathbb{N}$. Suppose $s_a a \preccurlyeq r_a$ and $s_b b \preccurlyeq r_b$. Then $s_a s_b a \preccurlyeq s_b r_a$ and $s_a s_b b \preccurlyeq s_a r_b$. By addition $s_a s_b(a+b) \preccurlyeq s_b r_a + s_a r_b$. Thus $\phi(a+b) \leq \frac{r_a}{s_a} + \frac{r_b}{s_b}$. We prove $\psi(a+b) \geq \psi(a)+\psi(b)$. Suppose $u_a \preccurlyeq v_a a$ and $u_b \preccurlyeq v_b b$. Then $v_b u_a \preccurlyeq v_a v_b a$ and $v_a u_b \preccurlyeq v_a v_b b$. By addition $v_b u_a + v_a u_b \preccurlyeq v_a v_b(a + b)$. Thus $\psi(a + b) \geq \frac{u_a}{v_a} + \frac{u_b}{v_b}$.

We prove $\phi(ab) \leq \phi(a)\phi(b)$. Suppose $s_a a \preccurlyeq r_a$ and $s_b b \preccurlyeq r_b$. Then $s_a s_b ab \preccurlyeq r_a r_b$. Thus $\phi(ab) \leq \frac{r_a}{s_a}\frac{r_b}{s_b}$. We prove $\psi(ab) \geq \psi(a)\psi(b)$. Suppose $u_a \preccurlyeq v_a a$ and $u_b \preccurlyeq v_b b$. Then $u_a u_b \preccurlyeq v_a v_b ab$. Thus $\frac{u_a}{v_a}\frac{u_b}{v_b} \leq \psi(ab)$.

We prove $a \preccurlyeq b \Rightarrow \phi(a) \leq \phi(b)$. Suppose $s_b b \preccurlyeq r_b$. From $a \preccurlyeq b$ follows $s_b a \preccurlyeq s_b b \preccurlyeq r_b$. Thus $\phi(a) \leq \frac{r_b}{s_b}$. The proof for ψ is similar.

One verifies directly that $\phi(1) = 1$, $\psi(1) = 1$, $\phi(0) = 0$ and $\psi(0) = 0$.

We prove $\psi(a) \leq \phi(a)$. Let $r, s, u, v \in \mathbb{N}$. Suppose $u \preccurlyeq va$ and $sa \preccurlyeq r$. Then follows $su \preccurlyeq vsa \preccurlyeq vr$. Thus $u/v \leq r/s$.

Assume that $\preccurlyeq$ is total.

We prove $\psi(a) \geq \phi(a)$. Suppose $\psi(a) < \phi(a)$. Let $r, s \in \mathbb{N}$ with $\psi(a) < r/s < \phi(a)$. Then $sa \not\preccurlyeq r$. From totality follows $sa \succcurlyeq r$. Thus $\psi(a) \geq r/s$, which is a contradiction. We conclude $\psi(a) = \phi(a)$.

We prove the uniqueness of ϕ. Let ϕ_1, ϕ_2 be semiring homomorphisms $S \to \mathbb{R}_{\geq 0}$ with $a \preccurlyeq b \Rightarrow \phi_i(a) \leq \phi_i(b)$. Suppose $\phi_1(a) < \phi_2(a)$. Let $u, v \in \mathbb{N}$ with $\phi_1(a) < \frac{u}{v} < \phi_2(a)$. Then $va \not\preccurlyeq u$, so by totality $va \succcurlyeq u$. Thus $\phi_1(a) \geq \frac{u}{v}$, which is a contradiction. This proves uniqueness.

Assume $\preccurlyeq$ is maximal in $\mathcal{P}$. Lemma 2.6 gives $\preccurlyeq \,=\, \precsim$. Let $a \not\preccurlyeq b$. From Lemma 2.4 (iv) follows $\exists n \;\; na \not\preccurlyeq nb + 1$. By totality $na \succcurlyeq nb + 1$. Apply ϕ to get $\phi(a) \geq \phi(b) + \frac{1}{n}$. In particular, $\phi(a) > \phi(b)$. □

Remark 2.10. Regarding Lemma 2.9, if $\preccurlyeq$ is not total, then in general $\phi \neq \psi$. For example, let $S = C(\{1, 2\}, \mathbb{R}_{\geq 0})$ be the semiring of functions from $\{1, 2\}$ to the nonnegative reals with pointwise addition and multiplication, and let $\preccurlyeq$ be the pointwise preorder on S. Then for $a \in S$ we have that $\psi(a) = \min(a)$ and $\phi(a) = \max(a)$ so $\psi \neq \phi$.

Lemma 2.11. *The map*

$$\mathbf{X}(S, \leqslant) \to \{\text{maximal elements in } \mathcal{P}_{\leqslant}\} : \phi \mapsto \preccurlyeq_\phi$$

with $\preccurlyeq_\phi$ defined by $a \preccurlyeq_\phi b$ iff $\phi(a) \leq \phi(b)$, is a bijection.

Proof. Let $\phi \in \mathbf{X}(S, \leqslant)$. One verifies that $\preccurlyeq_\phi$ is a Strassen preorder and $\leqslant \,\subseteq\, \lesssim \,\subseteq\, \preccurlyeq_\phi$. Let $\preccurlyeq$ be maximal in $\mathcal{P}_{\preccurlyeq_\phi}$. Lemma 2.7 says that $\preccurlyeq$ is total. By Lemma 2.9 there is a $\psi \in \mathbf{X}(S, \leqslant)$ with $\preccurlyeq \,\subseteq\, \preccurlyeq_\psi$. Clearly $\preccurlyeq_\phi \,\subseteq\, \preccurlyeq_\psi$. The uniqueness statement of Lemma 2.9 implies $\phi = \psi$. This means $\preccurlyeq_\phi \,=\, \preccurlyeq$, that is, $\preccurlyeq_\phi$ is maximal. We conclude that the map is well defined.

Let $\preccurlyeq$ maximal in $\mathcal{P}_{\leqslant}$. Then $\preccurlyeq$ is total. By Lemma 2.9 there is a $\phi \in \mathbf{X}(S, \leqslant)$ with $\preccurlyeq \,\subseteq\, \preccurlyeq_\phi$. We conclude the map is surjective.

Let $\phi, \psi \in \mathbf{X}(S, \leqslant)$ with $\preccurlyeq_\phi \,=\, \preccurlyeq_\psi$. From Lemma 2.9 follows $\phi = \psi$. We conclude the map is injective. □

Lemma 2.12. *Let $a, b \in S$. Then $a \lesssim b$ iff $a \preccurlyeq b$ for all maximal $\preccurlyeq \,\in \mathcal{P}_{\leqslant}$.*

Proof. Let $\preccurlyeq \in \mathcal{P}_{\leqslant}$ be maximal. Then $\lesssim \subseteq \precsim = \preccurlyeq$ by Lemma 2.6, so $a \lesssim b$ implies $a \preccurlyeq b$.

Suppose $a \not\lesssim b$. Let $n \in \mathbb{N}_{\geq 1}$ with $na \not\lesssim nb+1$ (Lemma 2.4 (iv)). By Lemma 2.5 there is an element $\preccurlyeq_{nb+1,na} \in \mathcal{P}$ with $\lesssim \subseteq \preccurlyeq_{nb+1,na}$ and we may assume $\preccurlyeq_{nb+1,na}$ is maximal. Then $nb+1 \preccurlyeq_{nb+1,na} na$ and so $a \not\preccurlyeq_{nb+1,na} b$. $\square$

2.7 The representation theorem

The following theorem is the main theorem.

Theorem 2.13 ([Str88, Th. 2.4]). *Let S be a commutative semiring with $\mathbb{N} \subseteq S$ and let $\leqslant$ be a Strassen preorder on S. Let $\mathbf{X} = \mathbf{X}(S, \leqslant)$ be the set of $\leqslant$-monotone semiring homomorphisms from S to $\mathbb{R}_{\geq 0}$,*

$$\mathbf{X} = \mathbf{X}(S, \leqslant) = \{\phi \in \mathrm{Hom}(S, \mathbb{R}_{\geq 0}) : \forall a, b \in S \;\; a \leqslant b \Rightarrow \phi(a) \leq \phi(b)\}.$$

For $a, b \in S$ let $a \lesssim b$ if there is a sequence $(x_N) \in \mathbb{N}^{\mathbb{N}}$ with $x_N^{1/N} \to 1$ when $N \to \infty$ such that $\forall N \in \mathbb{N}\;\; a^N \leqslant b^N x_N$. Then

$$\forall a, b \in S \quad a \lesssim b \;\text{ iff }\; \forall \phi \in \mathbf{X}\;\; \phi(a) \leq \phi(b).$$

Proof. Let $a, b \in S$. Suppose $a \lesssim b$. Then clearly for all $\phi \in \mathbf{X}$ we have $\phi(a) \leq \phi(b)$. Suppose $a \not\lesssim b$. By Lemma 2.12 there is a maximal $\preccurlyeq \in \mathcal{P}_{\leqslant}$ with $a \not\preccurlyeq b$. By Lemma 2.11 there is a $\phi \in \mathbf{X}$ with $\phi(a) > \phi(b)$. $\square$

2.8 Abstract rank and subrank R, Q

We generalise the notions of rank and subrank for tensors to arbitrary semirings with a Strassen preorder. Let $a \in S$. Define the *rank*

$$\mathrm{R}(a) \coloneqq \min\{r \in \mathbb{N} : a \leqslant r\}$$

and the *subrank*

$$\mathrm{Q}(a) \coloneqq \max\{r \in \mathbb{N} : r \leqslant a\}.$$

Then $\mathrm{Q}(a) \leq \mathrm{R}(a)$. Define the *asymptotic rank*

$$\underset{\sim}{\mathrm{R}}(a) \coloneqq \lim_{N \to \infty} \mathrm{R}(a^N)^{1/N}.$$

Define the *asymptotic subrank*

$$\underset{\sim}{\mathrm{Q}}(a) \coloneqq \lim_{N \to \infty} \mathrm{Q}(a^N)^{1/N}.$$

By Fekete's lemma (Lemma 2.2), asymptotic rank is an infimum and asymptotic subrank is a supremum as follows,

$$\underset{\sim}{\mathrm{R}}(a) = \inf_N \mathrm{R}(a^N)^{1/N}$$
$$\underset{\sim}{\mathrm{Q}}(a) = \sup_N \mathrm{Q}(a^N)^{1/N} \text{ when } a = 0 \text{ or } a \geq 1.$$

Theorem 2.13 implies that the asymptotic rank and asymptotic subrank have the following dual characterisation in terms of the asymptotic spectrum. (This is a straightforward generalisation of [Str88, Th. 3.8].)

Corollary 2.14 (cf. [Str88, Th. 3.8])**.** *For any $a \in S$ for which there exists an element $\phi \in \mathbf{X}$ such that $\phi(a) \geq 1$, holds*

$$\underset{\sim}{\mathrm{R}}(a) = \max_{\phi \in \mathbf{X}} \phi(a).$$

Proof. Let $\phi \in \mathbf{X}$. For $N \in \mathbb{N}$, $\mathrm{R}(a^N) \geq \phi(a)^N$. Therefore $\underset{\sim}{\mathrm{R}}(a) \geq \phi(a)$, and so $\underset{\sim}{\mathrm{R}}(a) \geq \max_{\phi \in \mathbf{X}} \phi(a)$. It remains to prove $\underset{\sim}{\mathrm{R}}(a) \leq \max_{\phi \in \mathbf{X}} \phi(a)$. We let $x \coloneqq \max_{\phi \in \mathbf{X}} \phi(a)$. By assumption $x \geq 1$. By definition of x we have

$$\forall \phi \in \mathbf{X} \ \phi(a) \leq x.$$

Take the mth power on both sides,

$$\forall \phi \in \mathbf{X}, m \in \mathbb{N} \ \phi(a^m) \leq x^m.$$

Take the ceiling on the right-hand side,

$$\forall \phi \in \mathbf{X}, m \in \mathbb{N} \ \phi(a^m) \leq \lceil x^m \rceil.$$

Apply Theorem 2.13 to get asymptotic preorders

$$\forall m \in \mathbb{N} \ a^m \lesssim \lceil x^m \rceil.$$

Then by definition of asymptotic preorder

$$\forall m, N \in \mathbb{N} \ a^{mN} \leqslant \lceil x^m \rceil^N 2^{\varepsilon_{m,N}} \quad \text{for some} \quad \varepsilon_{m,N} \in o(N).$$

Then

$$\forall m, N \in \mathbb{N} \ \ \mathrm{R}(a^{mN})^{1/mN} \leq \lceil x^m \rceil^{1/m} 2^{\varepsilon_{m,N}/mN}.$$

From $x \geq 1$ follows $\lceil x^m \rceil^{1/m} \to x$ when $m \to \infty$. Choose $m = m(N)$ with $m(N) \to \infty$ as $N \to \infty$ and $\varepsilon_{m(N),N} \in o(N)$ to get $\underset{\sim}{\mathrm{R}}(a) = \inf_N \mathrm{R}(a^N)^{1/N} \leq x$. □

Corollary 2.15 (cf. [Str88, Th. 3.8])**.** *For $a \in S$ with $\exists k \in \mathbb{N}\, a^k \geqslant 2$,*

$$\underset{\sim}{\mathrm{Q}}(a) = \min_{\phi \in \mathbf{X}} \phi(a).$$

Proof. Let $\phi \in \mathbf{X}$. For $N \in \mathbb{N}$, $\underset{\sim}{\mathrm{Q}}(a^N) \leq \phi(a)^N$. Therefore $\underset{\sim}{\mathrm{Q}}(a) \leq \phi(a)$, so $\underset{\sim}{\mathrm{Q}}(a) \leq \min_{\phi \in \mathbf{X}} \phi(a)$. It remains to prove $\underset{\sim}{\mathrm{Q}}(a) \geq \min_{\phi \in \mathbf{X}} \phi(a)$. Let $y \coloneqq \min_{\phi \in \mathbf{X}} \phi(a)$. From the assumption $a^k \geqslant 2$ follows $y \geqslant 1$. By definition of y we have

$$\forall \phi \in \mathbf{X}\ \phi(a) \geq y.$$

Take the mth power on both sides,

$$\forall \phi \in \mathbf{X}, m \in \mathbb{N}\ \phi(a^m) \geq y^m.$$

Take the floor on the right-hand side,

$$\forall \phi \in \mathbf{X}, m \in \mathbb{N}\ \phi(a^m) \geq \lfloor y^m \rfloor.$$

Apply Theorem 2.13 to get asymptotic preorders

$$\forall m \in \mathbb{N}\ a^m \gtrsim \lfloor y^m \rfloor.$$

Then by definition of asymptotic preorder

$$\forall m, N \in \mathbb{N}\ a^{mN} 2^{\varepsilon_{m,N}} \geqslant \lfloor y^m \rfloor^N \quad \text{for some} \quad \varepsilon_{m,N} \in o(N).$$

Now we use $a^k \geqslant 2$ to get

$$\forall m, N \in \mathbb{N}\ a^{mN + k\varepsilon_{m,N}} \geqslant \lfloor y^m \rfloor^N.$$

Then

$$\forall m, N \in \mathbb{N}\ \ \mathrm{Q}(a^{mN+k\varepsilon_{m,N}})^{\frac{1}{mN+k\varepsilon_{m,N}}} \geq \lfloor y^m \rfloor^{\frac{N}{mN+k\varepsilon_{m,N}}}.$$

Choose $m = m(N)$ with $m(N) \to \infty$ as $N \to \infty$ and $\varepsilon_{m(N),N} \in o(N)$ to obtain $\underset{\sim}{\mathrm{Q}}(a) = \sup_N \mathrm{Q}(a^N)^{1/N} \geq y$. □

2.9 Topological aspects

Theorem 2.13 does not tell the full story. Namely, there is also a topological component, which we will now discuss. Let S be a semiring with $\mathbb{N} \subseteq S$. Let $\leqslant$ be a Strassen preorder on S. Let $\mathbf{X} = \mathbf{X}(S, \leqslant)$ be the asymptotic spectrum of $(S, \leqslant)$. For $a \in S$, let

$$\hat{a} : \mathbf{X} \to \mathbb{R}_{\geq 0} : \phi \mapsto \phi(a). \tag{2.7}$$

The map $\hat{a}$ simply evaluates a given homomorphism ϕ at a. One may think of $\hat{a}$ as the collection $(\phi(a))_{\phi \in \mathbf{X}}$ of all evaluations of the elements of $\mathbf{X}$ at a. Let $\mathbb{R}_{\geq 0}$ have the Euclidean topology. Endow $\mathbf{X}$ with the weak topology with respect to the family of functions $\hat{a}$, $a \in S$. That is, endow $\mathbf{X}$ with the coarsest topology such that each $\hat{a}$ becomes continuous.

Let $C(\mathbf{X}, \mathbb{R}_{\geq 0})$ be the semiring of continuous functions $\mathbf{X} \to \mathbb{R}_{\geq 0}$ with addition and multiplication defined pointwise on $\mathbf{X}$, that is, $(f+g)(x) = f(x) + g(x)$ and $(f \cdot g)(x) = f(x)g(x)$ for $f, g \in C(\mathbf{X}, \mathbb{R}_{\geq 0})$ and $x \in \mathbf{X}$. Define the semiring homomorphism

$$\Phi : S \to C(\mathbf{X}, \mathbb{R}_{\geq 0}) : a \mapsto \hat{a},$$

which maps a to the evaluator $\hat{a}$ defined in (2.7).

Theorem 2.16 ([Str88, Th. 2.4])**.**

(i) $\mathbf{X}$ *is a nonempty compact Hausdorff space.*

(ii) $\forall a, b \in S \;\; a \lesssim b$ *iff* $\Phi(a) \leq \Phi(b)$ *pointwise on* $\mathbf{X}$*.*

(iii) $\Phi(S)$ *separates the points of* $\mathbf{X}$*.*

Proof. Statement (ii) follows from Theorem 2.13.

Statement (iii) is clear.

We prove statement (i). We have $2 \not\lesssim 1$, so from Theorem 2.13 follows that $\mathbf{X}$ cannot be empty.

For $a \in S$, let $n_a \in \mathbb{N}$ with $a \leq n_a$. Then for $\phi \in \mathbf{X}$, $\phi(a) \leq n_a$, and so $\phi(a) \in [0, n_a]$. Embed $\mathbf{X} \subseteq \prod_{a \in S} [0, n_a]$ as a set via $\phi \mapsto (\phi(a))_{a \in S}$. The set $\prod_{a \in S}[0, n_a]$ with the product topology is compact by the theorem of Tychonoff.

To see that $\mathbf{X}$ is closed in $\prod_{a \in S}[0, n_a]$, we write $\mathbf{X}$ as an intersection of sets,

$$\begin{aligned}
\mathbf{X} = \Big\{\phi \in \prod_{a \in S}[0, n_a] : \phi(0) = 0\Big\} \cap \Big\{\phi \in \prod_{a \in S}[0, n_a] : \phi(1) = 1\Big\} \\
\cap \bigcap_{b,c \in S} \Big\{\phi \in \prod_{a \in S}[0, n_a] : \phi(b+c) - \phi(b) - \phi(c) = 0\Big\} \\
\cap \bigcap_{b,c \in S} \Big\{\phi \in \prod_{a \in S}[0, n_a] : \phi(bc) - \phi(b)\phi(c) = 0\Big\} \\
\cap \bigcap_{\substack{b,c \in S: \\ b \leq c}} \Big\{\phi \in \prod_{a \in S}[0, n_a] : \phi(b) \leq \phi(c)\Big\},
\end{aligned}$$

and we observe that the intersected sets are closed,

$$\begin{aligned}
\mathbf{X} = \hat{0}^{-1}(\{0\}) \cap \hat{1}^{-1}(\{1\}) \\
\cap \bigcap_{b,c \in S} \big(\widehat{(b+c)} - \hat{b} - \hat{c}\big)^{-1}(\{0\}) \\
\cap \bigcap_{b,c \in S} \big(\widehat{(bc)} - \hat{b}\hat{c}\big)^{-1}(\{0\})
\end{aligned}$$

$$\cap \bigcap_{\substack{b,c \in S: \\ b \leq c}} \left(\hat{c} - \hat{b}\right)^{-1}([0,\infty)).$$

This implies $\mathbf{X}$ is also compact.

Let $\phi, \psi \in \mathbf{X}$ be distinct. Let $a \in S$ with $\phi(a) \neq \psi(a)$. Then $\hat{a}(\phi) \neq \hat{a}(\psi)$. Let $U \ni \hat{a}(\phi)$, $V \ni \hat{a}(\psi)$ be open and disjoint subsets of $\mathbb{R}_{\geq 0}$. Then $\hat{a}^{-1}(U)$ and $\hat{a}^{-1}(V)$ are open and disjoint subsets of $\mathbf{X}$. We conclude that $\mathbf{X}$ is Hausdorff. □

2.10 Uniqueness

Let S be a semiring with $\mathbb{N} \subseteq S$. Let $\leqslant$ be a Strassen preorder on S. Let $\mathbf{X} = \mathbf{X}(S, \leqslant)$ be the asymptotic spectrum of $(S, \leqslant)$. The object $\mathbf{X}$ is unique in the following sense.

Theorem 2.17 ([Str88, Cor. 2.7]). *Let $\mathbf{Y}$ be a compact Hausdorff space. Let $\Psi : S \to C(\mathbf{Y}, \mathbb{R}_{\geq 0})$ be a homomorphism of semirings such that*

$$\Psi(S) \text{ separates the points of } \mathbf{Y} \tag{2.8}$$

and

$$\forall a, b \in S \;\; a \lesssim b \Leftrightarrow \Psi(a) \leq \Psi(b) \text{ pointwise on } \mathbf{Y}. \tag{2.9}$$

Then there is a unique homeomorphism (continuous bijection with continuous inverse) $h : \mathbf{Y} \to \mathbf{X}$ such that the diagram

$$\begin{tikzcd} & S \arrow[dl, "\Phi"'] \arrow[dr, "\Psi"] & \\ C(\mathbf{X}, \mathbb{R}_{\geq 0}) \arrow[rr, "h^*"] & & C(\mathbf{Y}, \mathbb{R}_{\geq 0}) \end{tikzcd} \tag{2.10}$$

commutes, where $h^* : \phi \mapsto \phi \circ h$. Namely, let $h : y \mapsto \big(a \mapsto \Psi(a)(y)\big)$.

Proof. We prove uniqueness. Suppose there are two such homeomorphisms

$$h_1, h_2 : \mathbf{Y} \to \mathbf{X}.$$

Suppose $x \neq h_2(h_1^{-1}(x))$ for some $x \in \mathbf{X}$. Since $\Phi(S)$ separates the points of $\mathbf{X}$, there is an $a \in S$ with $\Phi(a)(x) \neq \Phi(a)(h_2(h_1^{-1}(x)))$. Let $y = h_1^{-1}(x) \in \mathbf{Y}$. Then $\Phi(a)(h_1(y)) \neq \Phi(a)(h_2(y))$. Since (2.10) commutes, $\Phi(a)(h_1(y)) = \Psi(a)(y)$ and $\Phi(a)(h_2(y)) = \Psi(a)(y)$, a contradiction.

We prove existence. Let $h : \mathbf{Y} \to \mathbf{X} : y \mapsto (a \mapsto \Psi(a)(y))$. One verifies that h is well-defined, continuous, injective and that the diagram in (2.10) commutes. It remains to show that h is surjective. We know that $\mathbb{Q} \cdot \Phi(S)$ is a $\mathbb{Q}$-subalgebra of $C(\mathbf{X}, \mathbb{R})$ which separates points and which contains the nonzero constant function $\Phi(1)$, so by the Stone–Weierstrass theorem, $\mathbb{Q} \cdot \Phi(S)$ is dense in $C(\mathbf{X}, \mathbb{R})$ under the sup-norm. Suppose h is not surjective. Then $h(\mathbf{Y}) \subsetneq \mathbf{X}$ is a proper closed subset. Let $x_0 \in \mathbf{X} \setminus h(\mathbf{Y})$ be in the complement. Since $\mathbf{X}$ is a compact Hausdorff space, there is a continuous function $f : \mathbf{X} \to [-1, 1]$ with

$$\begin{aligned} &f(h(\mathbf{Y})) = 1 \\ &f(x_0) = -1. \end{aligned}$$

We know that f can be approximated by elements from $\mathbb{Q} \cdot \Phi(S)$, i.e. let $\varepsilon > 0$, then there are $a_1, a_2 \in S$, $N \in \mathbb{N}$ such that

$$\begin{aligned} &\tfrac{1}{N}\big(\Phi(a_1)(x) - \Phi(a_2)(x)\big) > 1 - \varepsilon \quad \text{for all } x \in h(\mathbf{Y}) \\ &\tfrac{1}{N}\big(\Phi(a_1)(x_0) - \Phi(a_2)(x_0)\big) < -1 + \varepsilon. \end{aligned}$$

This means $\Psi(a_1) \geq \Psi(a_2)$ pointwise on $\mathbf{Y}$, so $a_1 \gtrsim a_2$, but also $\Phi(a_1) \not\geq \Phi(a_2)$ pointwise on $\mathbf{X}$, so $a_1 \not\gtrsim a_2$. This is a contradiction. □

2.11 Subsemirings

Let $(T, +, \cdot)$ be a semiring. A subset $S \subseteq T$ is called a *subsemiring* if $0, 1 \in S$ and moreover S is closed under $+$ and $\cdot$, i.e. for all $a, b \in S$ holds $a + b \in S$ and $a \cdot b \in S$.

Let S be a subsemiring of a semiring T and let $\leqslant$ be a Strassen preorder on T. Of course, $\leqslant$ also defines a preorder on S. Formally, we define the restriction $\leqslant|_S$ by, for all $a, b \in S$, $a \leqslant|_S b$ if and only if $a \leqslant b$. Then $\leqslant|_S$ is a Strassen preorder on S. How are the asymptotic spectra $\mathbf{X}(S, \leqslant|_S)$ and $\mathbf{X}(T, \leqslant)$ related? Obviously, for $\phi \in \mathbf{X}(T, \leqslant)$ we have $\phi|_S \in \mathbf{X}(S, \leqslant|_S)$, where $\phi|_S$ denotes the restriction of ϕ to S. In fact, the uniqueness theorem of Section 2.10 implies that all elements of $\mathbf{X}(S, \leqslant|_S)$ are restrictions of elements of $\mathbf{X}(T, \leqslant)$, as follows. Let $\mathbf{X}(T, \leqslant)|_S \coloneqq \{\phi|_S : \phi \in \mathbf{X}(T, \leqslant)\}$.

Corollary 2.18. *Let S be a subsemiring of a semiring T. Let $\leqslant$ be a Strassen preorder on T. Then*

$$\mathbf{X}(S, \leqslant|_S) = \mathbf{X}(T, \leqslant)|_S.$$

Proof. Let

$$\begin{aligned} &\mathbf{X} = \mathbf{X}(S, \leqslant|_S), \\ &\Phi : S \to C(\mathbf{X}, \mathbb{R}_{\geq 0}) : a \mapsto \hat{a} \end{aligned}$$

and let

$$\mathbf{Y} = \mathbf{X}(T, \leqslant)|_S = \{\phi|_S : \phi \in \mathbf{X}(T, \leqslant)\},$$

$$\Psi : S \to C(\mathbf{Y}, \mathbb{R}_{\geq 0}) : a \mapsto \big(\phi|_S \mapsto \phi|_S(a)\big).$$

Then $\mathbf{Y}$ is a compact Hausdorff space. Let $\phi|_S, \psi|_S \in \mathbf{Y}$ be distinct. Then there is an $a \in S$ with $\phi|_S(a) \neq \psi|_S(a)$, so (2.8) holds. For $a, b \in S$, $a \lesssim b$ iff $\Phi(a) \leq \Phi(b)$ iff $\Psi(a) \leq \Psi(b)$, so (2.9) holds. Therefore,

$$h : \mathbf{X}(T, \leqslant)|_S \to \mathbf{X}(S, \leqslant|_S) : \phi|_S \mapsto \big(a \mapsto \Psi(a)(\phi|_S)\big) = \phi|_S$$

is a homeomorphism. □

2.12 Subsemirings generated by one element

Let S be a semiring and let $\leqslant$ be a Strassen preorder on S. In this section we specialise to the simplest type of subsemiring of S. Namely, let $a \in S$ and let

$$\mathbb{N}[a] \coloneqq \Big\{\sum_{i=0}^{k} n_i\, a^i : k \in \mathbb{N}, n_i \in \mathbb{N}\Big\} \subseteq S$$

be the subsemiring of S generated by a. We call $\mathbf{X}(\mathbb{N}[a]) = \mathbf{X}(\mathbb{N}[a], \leqslant|_{\mathbb{N}[a]})$ the *asymptotic spectrum of* a.

Corollary 2.19 (cf. [Str88])**.** *If* $a^k \geqslant 2$ *for some* $k \in \mathbb{N}$*, then*

$$\underset{\sim}{\mathrm{Q}} \in \mathbf{X}(\mathbb{N}[a]).$$

If $\phi(a) \geq 1$ *for some* $\phi \in \mathbf{X}$*, then*

$$\underset{\sim}{\mathrm{R}} \in \mathbf{X}(\mathbb{N}[a]).$$

Proof. Let $\mathbf{X} = \mathbf{X}(\mathbb{N}[a])$. Let $n_1, \ldots, n_q$. By Corollary 2.15

$$\underset{\sim}{\mathrm{Q}}(a^{n_1} + \cdots + a^{n_q}) = \min_{\phi \in \mathbf{X}} \phi(a^{n_1} + \cdots + a^{n_q}).$$

Since ϕ is a homomorphism, $\phi(a^{n_1} + \cdots + a^{n_q}) = \phi(a)^{n_1} + \cdots + \phi(a)^{n_q}$. Now we observe that $x^{n_1} + \cdots + x^{n_q}$ is minimised by taking x minimal in the domain. We conclude

$$\underset{\sim}{\mathrm{Q}}(a^{n_1} + \cdots + a^{n_q}) = \sum_{i=1}^{q} \big(\min_{\phi \in \mathbf{X}} \phi(a)\big)^{n_i} = \underset{\sim}{\mathrm{Q}}(a)^{n_1} + \cdots + \underset{\sim}{\mathrm{Q}}(a)^{n_q}.$$

The claim for asymptotic rank $\underset{\sim}{\mathrm{R}}$ similarly follows from Corollary 2.14. □

Remark 2.20. In general, asymptotic subrank $\underset{\sim}{\mathrm{Q}}$ and asymptotic rank $\underset{\sim}{\mathrm{R}}$ are not elements of the asymptotic spectrum. We will see an example in Chapter 4 related to the matrix multiplication tensor.

Remark 2.21. Corollary 2.19 is closely related to Schönhage's "τ-theorem" for tensors, also called the "asymptotic sum inequality", see e.g. [BCS97, Section 15.5] and [Blä13a, Theorem 7.5]. The τ-theorem features in every recent fast matrix multiplication algorithm (more precisely, every algorithm based on the "laser method").

Remark 2.22. Every element $\phi \in \mathbf{X}(\mathbb{N}[a])$ is uniquely determined by the value of $\phi(a)$. We may thus identify $\mathbf{X}(\mathbb{N}[a])$ with the set $\{\phi(a) : \phi \in \mathbf{X}(\mathbb{N}[a])\} \subseteq \mathbb{R}_{\geq 0}$. By Theorem 2.16(i), this set is compact, so it is a union of finitely many closed intervals.

2.13 Universal spectral points

Having discussed the simplest type of subsemiring in the previous section, let us discuss the most difficult type of supersemiring. When applying the theory of asymptotic spectra to some setting, there is a natural largest semiring S in which the objects of study live. For example, we may study the semiring S of all (equivalence classes of) 3-tensors of arbitrary format over $\mathbb{F}$. Or we may study the semiring S of all (isomorphism classes of) finite simple graphs. We refer to the elements of the asymptotic spectrum $\mathbf{X}(S)$ of the "ambient" semiring S by the term *universal spectral points* (cf. [Str88, page 119]). The universal spectral points are the most useful monotone homomorphisms.

2.14 Conclusion

To a semiring S with a Strassen preorder $\leqslant$, we associated an asymptotic preorder $\lesssim$. We proved that this asymptotic preorder is characterised by the $\leqslant$-monotone semiring homomorphisms $S \to \mathbb{R}_{\geq 0}$, which make up the asymptotic spectrum $\mathbf{X}(S, \leqslant)$ of $(S, \leqslant)$. For $(S, \leqslant)$ we naturally have a rank function $\mathrm{R} : S \to \mathbb{N}$ and a subrank function $\mathrm{Q} : S \to \mathbb{N}$. Their asymptotic versions $\underaccent{\sim}{\mathrm{R}}(a) = \inf_n \mathrm{R}(a^n)^{1/n}$ and $\underaccent{\sim}{\mathrm{Q}}(a) = \sup_n \mathrm{Q}(a^n)^{1/n}$ coincide with $\max_{\phi \in \mathbf{X}(S,\leqslant)} \phi(a)$ and $\min_{\phi \in \mathbf{X}(S,\leqslant)} \phi(a)$ respectively, assuming $\exists \phi \in \mathbf{X}\, \phi(a) \geq 1$ and $\exists k \in \mathbb{N}\, a^k \geqslant 2$ respectively. Unfortunately, we have proved the existence of the asymptotic spectrum by nonconstructive means. Explicitly constructing spectral points for a given pair $(S, \leqslant)$ will be a challenging task!

Some remarks about our proof in this chapter. The proof in [Str88] uses the Kadison–Dubois theorem from the paper of Becker and Schwartz [BS83] as a black-box. Our presentation basically integrates the proof of Strassen with the proof of Becker and Schwartz. The notions of rank and subrank were in [Str88] only discussed for tensors. We considered the straightforward generalisation to arbitrary semirings with a Strassen preorder. An evident feature of our presentation is that we do not pass from the semiring to its Grothendieck ring, but instead stay in the semiring. In this way we stay close to the “real world” objects. I thank Jop Briët and Lex Schrijver for this idea. There is a large body of literature on the Kadison–Dubois theorem, for which we refer to the modern books by Prestel and Delzell [PD01, Theorem 5.2.6] and Marshall [Mar08, Theorem 5.4.4].

Chapter 3

The asymptotic spectrum of graphs; Shannon capacity

This chapter is based on the manuscript [Zui18].

3.1 Introduction

This chapter is about the Shannon capacity of graphs, which was introduced by Claude Shannon in the context of coding theory [Sha56]. More precisely, we will apply the theory of asymptotic spectra of Chapter 2 to gain a better understanding of Shannon capacity (and other asymptotic properties of graphs).

We first recall the definition of the Shannon capacity of a graph. Let G be a (finite simple) graph with vertex set $V(G)$ and edge set $E(G)$. An *independent set* or *stable set* in G is a subset of $V(G)$ that contains no edges. The *independence number* or *stability number* $\alpha(G)$ is the cardinality of the largest independent set in G. For graphs G and H, the *and-product* $G \boxtimes H$, also called strong graph product, is defined by

$$\begin{aligned} V(G \boxtimes H) &= V(G) \times V(H) \\ E(G \boxtimes H) &= \big\{\{(g,h),(g',h')\} : \big(\{g,g'\} \in E(G) \text{ or } g = g'\big) \\ &\qquad \text{and } \big(\{h,h'\} \in E(H) \text{ or } h = h'\big) \text{ and } (g,h) \neq (g',h')\big\}. \end{aligned}$$

The *Shannon capacity* $\Theta(G)$ is defined as the limit

$$\Theta(G) \coloneqq \lim_{N\to\infty} \alpha(G^{\boxtimes N})^{1/N}. \tag{3.1}$$

This limit exists and equals the supremum $\sup_N \alpha(G^{\boxtimes N})^{1/N}$ by Fekete's lemma (Lemma 2.2).

Computing the Shannon capacity is nontrivial already for small graphs. Lovász in [Lov79] computed the value $\Theta(C_5) = \sqrt{5}$, where C_k denotes the k-cycle graph, by introducing and evaluating a new graph parameter ϑ which is now known as the Lovász theta number. For example the value of $\Theta(C_7)$ is currently not known. The Shannon capacity Θ is not known to be hard to compute in the sense of computational complexity. On the other hand, deciding whether $\alpha(G) \leq k$, given a graph G and $k \in \mathbb{N}$, is NP-complete [Kar72].

New result: dual description of Shannon capacity

The new result of this chapter is a dual characterisation of the Shannon capacity of graphs. This characterisation is obtained by applying Strassen's theory of asymptotic spectra of Chapter 2. Thus this chapter also serves as an illustration of the theory of asymptotic spectra.

To state the theorem we need the standard notions *graph homomorphism*, *graph complement* and *graph disjoint union*. Let G and H be graphs. A *graph homomorphism* $f : G \to H$ is a map $f : V(G) \to V(H)$ such that for all $u, v \in V(G)$, if $\{u, v\} \in E(G)$, then $\{f(u), f(v)\} \in E(H)$. In other words, a graph homomorphism maps edges to edges. The *complement* $\overline{G}$ of G is defined by

$$\begin{aligned} V(\overline{G}) &= V(G) \\ E(\overline{G}) &= \big\{\{u, v\} : \{u, v\} \notin E(G), u \neq v\big\}. \end{aligned}$$

The *disjoint union* $G \sqcup H$ is defined by

$$\begin{aligned} V(G \sqcup H) &= V(G) \sqcup V(H) \\ E(G \sqcup H) &= E(G) \sqcup E(H). \end{aligned}$$

For $n \in \mathbb{N}$, the *complete graph* K_n is the graph with $V(K_n) = [n] \coloneqq \{1, 2, \ldots, n\}$ and $E(K_n) = \{\{i, j\} : i, j \in [n], i \neq j\}$. Thus $K_0 = \overline{K_0}$ is the empty graph and $K_1 = \overline{K_1}$ is the graph consisting of a single vertex and no edges.

We define the *cohomomorphism preorder* $\leqslant$ on graphs by $G \leqslant H$ if and only if there is a graph homomorphism $\overline{G} \to \overline{H}$ from the complement of G to the complement of H. With this definition, the independence number $\alpha(G)$ equals the maximum n such that $(K_1)^{\sqcup n} \leqslant G$.

Theorem 3.1. *Let $S \subseteq$ {graphs} be a collection of graphs which is closed under the disjoint union $\sqcup$ and the strong graph product $\boxtimes$, and which contains the graph with a single vertex, K_1. Define the asymptotic spectrum $\mathbf{X}(S)$ as the set of all maps $\phi : S \to \mathbb{R}_{\geq 0}$ such that, for all $G, H \in S$*

(1) *if $G \leqslant H$, then $\phi(G) \leq \phi(H)$*

(2) $\phi(G \sqcup H) = \phi(G) + \phi(H)$

(3) $\phi(G \boxtimes H) = \phi(G)\phi(H)$

(4) $\phi(K_1) = 1$.

Let $G \lesssim H$ *if there is a sequence* $(x_N) \in \mathbb{N}^{\mathbb{N}}$ *with* $x_N^{1/N} \to 1$ *when* $N \to \infty$ *such that for every* $N \in \mathbb{N}$

$$G^{\boxtimes N} \leqslant (H^{\boxtimes N})^{\sqcup x_N} = \underbrace{H^{\boxtimes N} \sqcup \cdots \sqcup H^{\boxtimes N}}_{x_N}.$$

Then

(i) $G \lesssim H$ *iff* $\forall \phi \in \mathbf{X}(S)\ \phi(G) \leq \phi(H)$

(ii) $\Theta(G) = \min_{\phi \in \mathbf{X}(S)} \phi(G)$.

Statement (ii) of Theorem 3.1 is nontrivial in the sense that Θ is not an element of $\mathbf{X}(\{\text{graphs}\})$. Namely, Θ is not additive under $\sqcup$ by a result of Alon [Alo98], and Θ is not multiplicative under $\boxtimes$ by a result of Haemers [Hae79], see Example 3.10. It turns out that the graph parameter $G \mapsto \max_{\phi \in \mathbf{X}(\{\text{graphs}\})} \phi(G)$ is itself an element of $\mathbf{X}(\{\text{graphs}\})$, and is equal to the fractional clique cover number $\overline{\chi}_f$ (see Section 3.3.2 and e.g. [Sch03, Eq. (67.112)]). Fritz in [Fri17] proves (independently of Strassen's line of work!) a statement that is weaker than Theorem 3.1. Namely he proves the conclusion of Theorem 3.1 without the additivity condition (2) and with the normalization condition (4) changed to $\phi(K_n) = n$ for all $n \in \mathbb{N}$.

In Section 3.2 we will prove Theorem 3.1 by applying the theory of asymptotic spectra of Chapter 2 to the appropriate semiring and preorder. In Section 3.3 we will discuss the elements in the asymptotic spectrum of graphs $\mathbf{X}(\{\text{graphs}\})$ that are currently known to me: the Lovász theta number, the fractional clique cover number, the fractional orthogonal rank of the complement, and the fractional Haemers bounds. We moreover prove a sufficient condition for the "fractionalisation" of a graph parameter to be in the asymptotic spectrum of graphs.

3.2 The asymptotic spectrum of graphs

In this section we prove Theorem 3.1 by applying the theory of asymptotic spectra to the appropriate semiring.

3.2.1 The semiring of graph isomorphism classes $\mathcal{G}$

A graph homomorphism $f : G \to H$ is a *graph isomorphism* if f is bijective as a map $V(G) \to V(H)$ and bijective as a map $E(G) \to E(H)$. We write $G \cong H$ if there is a graph isomorphism $f : G \to H$. The relation $\cong$ is an equivalence relation on $\{\text{graphs}\}$, which we call *isomorphism*. For example, the graphs G and H given by

$$V(G) = \{a, b, c, d\}, \quad E(G) = \{\{a, b\}, \{b, c\}, \{c, d\}, \{a, d\}\}$$

$$V(H) = \{1,2,3,4\}, \quad E(H) = \{\{1,3\},\{2,3\},\{2,4\},\{1,4\}\}$$

are isomorphic. Let $\mathcal{G} = \{\text{graphs}\}/\!\cong$ be the set of equivalence classes in $\{\text{graphs}\}$ under $\cong$, i.e. the isomorphism classes. The relation $\leqslant$ is a preorder on $\mathcal{G}$. Recall that K_n is the complete graph on n vertices, and thus $\overline{K_n}$ is the graph with n vertices and no edges.

Lemma 3.2. *Let $A, B, C \in \{\text{graphs}\}$.*

(i) *$\sqcup$ and $\boxtimes$ are commutative and associative operations on $\mathcal{G}$.*

(ii) *$\boxtimes$ distributes over $\sqcup$ on $\mathcal{G}$, i.e. $A \boxtimes (B \sqcup C) = (A \boxtimes B) \sqcup (A \boxtimes C)$.*

(iii) *$\overline{K_1} \boxtimes A = A$.*

(iv) *$\overline{K_0} \boxtimes A = \overline{K_0}$.*

(v) *$\overline{K_0} \sqcup A = A$.*

(vi) *$\overline{K_n} \sqcup \overline{K_m} = \overline{K_{n+m}}$.*

Proof. We leave the proof to the reader. □

In other words, Lemma 3.2 says that $(\mathcal{G}, \sqcup, \boxtimes, \overline{K_0}, \overline{K_1})$ is a (commutative) semiring in which the elements $\overline{K_0}, \overline{K_1}, \overline{K_2}, \ldots$ behave like the natural numbers $\mathbb{N}$. We will denote this semiring simply by $\mathcal{G}$.

3.2.2 Strassen preorder via graph homomorphisms

Let $\mathcal{G}$ be the semiring of graphs. Recall that $G \leqslant H$ if there is a graph homomorphism $f : \overline{G} \to \overline{H}$.

Lemma 3.3. *The preorder $\leqslant$ is a Strassen preorder on $\mathcal{G}$. That is, for graphs $A, B, C, D \in \mathcal{G}$ we have the following.*

(i) *For $n, m \in \mathbb{N}$, $\overline{K_n} \leqslant \overline{K_m}$ iff $n \leq m$.*

(ii) *If $A \leqslant B$ and $C \leqslant D$, then $A \sqcup C \leqslant B \sqcup D$ and $A \boxtimes C \leqslant B \boxtimes D$.*

(iii) *For $A, B \in \mathcal{G}$, if $B \neq K_0$, then there is an $r \in \mathbb{N}$ with $A \leqslant \overline{K_r} \boxtimes B$.*

Proof. Statement (i) is easy to verify.

We prove (ii). Let $f : \overline{A} \to \overline{B}$ and $g : \overline{C} \to \overline{D}$ be graph homomorphisms. Let the map $f \sqcup g : V(A) \sqcup V(C) \to V(B) \sqcup V(D)$ be defined by

$$(f+g)(a) = f(a) \text{ for } a \in V(A)$$
$$(f+g)(c) = g(c) \text{ for } c \in V(C).$$

One verifies directly that $f \sqcup g$ is a graph homomorphism $\overline{A \sqcup C} \to \overline{B \sqcup D}$. Let the map $f \boxtimes g : V(A) \times V(C) \to V(B) \times V(D)$ be defined by

$$(f \boxtimes g)(a, c) = (f(a), g(c)).$$

One verifies directly that $f \boxtimes g$ is a graph homomorphism $\overline{A \boxtimes C} \to \overline{B \boxtimes D}$. This proves (ii).

We prove (iii). Let $r = |V(A)|$. Then $A \leqslant \overline{K_r}$. By assumption, $B \neq K_0$, so $\overline{K_1} \leqslant B$. Therefore $A \leqslant \overline{K_r} \cong \overline{K_r} \boxtimes 1 \leqslant \overline{K_r} \boxtimes B$. This proves (iii). $\square$

3.2.3 The asymptotic spectrum of graphs $\mathbf{X}(\mathcal{G})$

We thus have a semiring $\mathcal{G}$ with a Strassen preorder $\leqslant$. We are therefore in the position to apply the theory of asymptotic spectra (Chapter 2). Let us translate the abstract terminology to this setting.

Let $G \lesssim H$ if there is a sequence $(x_N) \in \mathbb{N}^{\mathbb{N}}$ with $(x_N)^{1/N} \to 1$ such that for every $N \in \mathbb{N}$ we have $G^{\boxtimes N} \leqslant H^{\boxtimes N} \boxtimes \overline{K_{x_N}}$ i.e. $G^{\boxtimes N} \leqslant (H^{\boxtimes N})^{\sqcup x_N}$.

Let $S \subseteq \mathcal{G}$ be a subsemiring. For example, one may take $S = \mathcal{G}$, or one may choose any set $X \subseteq \mathcal{G}$ and let $S = \mathbb{N}[X]$ be the subsemiring of $\mathcal{G}$ generated by X under $\sqcup$ and $\boxtimes$.

The asymptotic spectrum of S is the set $\mathbf{X}(S)$ of $\leqslant$-monotone semiring homomorphisms $S \to \mathbb{R}_{\geq 0}$, i.e. all maps $\phi : S \to \mathbb{R}_{\geq 0}$ such that, for all $G, H \in S$

(1) if $G \leqslant H$, then $\phi(G) \leq \phi(H)$

(2) $\phi(G \sqcup H) = \phi(G) + \phi(H)$

(3) $\phi(G \boxtimes H) = \phi(G)\phi(H)$

(4) $\phi(K_1) = 1$.

We call $\mathbf{X}(\mathcal{G})$ the asymptotic spectrum of graphs.

Theorem 3.4. *Let $G, H \in S$. Then $G \lesssim H$ iff $\forall \phi \in \mathbf{X}(S)\ \phi(G) \leq \phi(H)$.*

Proof. By Lemma 3.2 we have a semigroup S and by Lemma 3.3 we have a Strassen preorder $\leqslant$, so we may apply Theorem 2.13. $\square$

We refer to Chapter 2 for a discussion of the topological properties of $\mathbf{X}(S)$.

3.2.4 Shannon capacity Θ

Let us discuss the (asymptotic) rank and (asymptotic) subrank for $(\mathcal{G}, \leqslant)$. Recall that an independent set in G is a subset of $V(G)$ that contains no edges, and the independence number $\alpha(G)$ is the cardinality of the largest independent set in G. A *colouring* of G is an assignment of colours to the elements of $V(G)$ such that connected vertices get distinct colours. The *chromatic number* $\chi(G)$ is the smallest number of colours in any colouring of G. The *clique cover number* $\overline{\chi}(G)$ is defined as the chromatic number of the complement, $\overline{\chi}(G) \coloneqq \chi(\overline{G})$.

For the semiring $\mathcal{G}$ with preorder $\leqslant$ the abstract definition of subrank of Section 2.8 becomes $\mathrm{Q}(G) = \max\{m \in \mathbb{N} : \overline{K_m} \leqslant G\}$ and the abstract definition of rank becomes $\mathrm{R}(G) = \min\{n \in \mathbb{N} : G \leqslant \overline{K_n}\}$.

Lemma 3.5.

(i) $\alpha(G) = \mathrm{Q}(G)$

(ii) $\overline{\chi}(G) = \mathrm{R}(G)$

Proof. We leave the proof to the reader. □

We see directly that the asymptotic rank is the asymptotic clique cover number,

$$\underset{\sim}{\mathrm{R}}(G) = \lim_{N\to\infty} \mathrm{R}(G^{\boxtimes N})^{1/N} = \lim_{N\to\infty} \overline{\chi}(G^{\boxtimes N})^{1/N} \eqqcolon \underset{\sim}{\overline{\chi}}(G)$$

and that the asymptotic subrank is the Shannon capacity,

$$\underset{\sim}{\mathrm{Q}}(G) = \lim_{N\to\infty} \mathrm{Q}(G^{\boxtimes N})^{1/N} = \lim_{N\to\infty} \alpha(G^{\boxtimes N})^{1/N} = \Theta(G).$$

Let $S \subseteq \mathcal{G}$ be a subsemiring. Let $G \in S$.

Corollary 3.6. $\Theta(G) = \min_{\phi\in \mathbf{X}(S)} \phi(G)$.

Proof. Let G be a graph. Either $G = K_0$ or $K_1 \leqslant G \leqslant K_1$ or $\overline{G}$ contains at least one edge. In the first two cases the claim is clearly true. In the third case $G \geqslant \overline{K_2}$ and we may thus apply Corollary 2.14. □

Corollary 3.7. $\underset{\sim}{\overline{\chi}}(G) = \max_{\phi\in \mathbf{X}(S)} \phi(G)$.

Proof. This is Corollary 2.15. □

Remark 3.8. As mentioned earlier, it turns out that $\underset{\sim}{\overline{\chi}}$ is in fact itself an element of $\mathbf{X}(\mathcal{G})$! See Section 3.3.2. (This is a striking difference with the situation for tensors, which we will discuss in Chapter 4; there, both asymptotic rank and asymptotic subrank are not in the asymptotic spectrum, see Remark 4.4.)

Shannon capacity is not in the asymptotic spectrum

Lemma 3.9. $G \boxtimes \overline{G} \geq \overline{K_{|V(G)|}}$.

Proof. Let $D = \{(u,u) : u \in V(G)\}$. Let $(u,u),(v,v) \in D$. Then either $\{u,v\} \in E(G)$ or $\{u,v\} \in E(\overline{G})$ (exclusive or), and so $\{(u,u),(v,v)\} \notin E(G\boxtimes \overline{G})$. Therefore, the subgraph in $G \boxtimes \overline{G}$ induced by D is isomorphic to $\overline{K_{|V(G)|}}$. □

Example 3.10. Let G be the Schläfli graph. This is a graph with 27 vertices. Thus $\Theta(G \boxtimes \overline{G}) \geq |V(G)| = 27$ by Lemma 3.9. On the other hand, Haemers in [Hae79] showed that $\Theta(\overline{G})\Theta(G) \leq 21$. This implies the map Θ is not in $\mathbf{X}(\mathcal{G})$, since it is not multiplicative under $\boxtimes$.

3.3 Universal spectral points

The abstract theory of asymptotic spectra of Chapter 2 does not explicitly describe the elements of $\mathbf{X}(\mathcal{G})$, i.e. the universal spectral points (cf. Section 2.13). However, several graph parameters from the literature can be shown to be universal spectral points. In fact, recently in [BC18] the first infinite family of universal spectral points was found, the fractional Haemers bounds. We give a brief (and probably incomplete) overview of currently known elements in $\mathbf{X}(\mathcal{G})$.

3.3.1 Lovász theta number ϑ

For any real symmetric matrix A let $\Lambda(A)$ be the largest eigenvalue. The Lovász theta number $\vartheta(G)$ is defined as

$$\vartheta(G) \coloneqq \min\{\Lambda(A) : A \in \mathbb{R}^{V(G)\times V(G)} \text{ symmetric}, \{u,v\} \notin E(G) \Rightarrow A_{uv} = 1\}.$$

The parameter $\vartheta(G)$ was introduced by Lovász in [Lov79]. We refer to [Knu94] and [Sch03] for a survey. It follows from well-known properties that $\vartheta \in \mathbf{X}(\mathcal{G})$.

3.3.2 Fractional graph parameters

Besides the Lovász theta number there are several elements in $\mathbf{X}(\mathcal{G})$ that are naturally obtained as fractional versions of $\boxtimes$-submultiplicative, $\sqcup$-subadditive, $\leqslant$-monotone maps $\mathcal{G} \to \mathbb{R}_{\geq 0}$. For any map $\phi : \mathcal{G} \to \mathbb{R}_{\geq 0}$ we define a fractional version ϕ_f by

$$\phi_f(G) \coloneqq \inf_d \frac{\phi\big(\overline{\overline{G} \boxtimes K_d}\big)}{d}. \tag{3.2}$$

We will discuss several fractional parameters from the literature and prove a general theorem about fractional parameters.

Fractional clique cover number

We consider the fractional version of the clique cover number $\overline{\chi}(G) = \chi(\overline{G})$. It is well-known that $\overline{\chi}_f \in \mathbf{X}(\mathcal{G})$, see e.g. [Sch03]. The fractional clique cover number $\overline{\chi}_f$ in fact equals the asymptotic clique cover number $\underset{\sim}{\overline{\chi}}(G) \coloneqq \lim_{N\to\infty} \overline{\chi}(G^{\boxtimes N})^{1/N}$ which we introduced in the previous section, see [MP71] and also [Sch03, Th. 67.17].

Fractional Haemers bound

Let $\mathrm{rank}(A)$ denote the matrix rank of any matrix A. For any set C of matrices define $\mathrm{rank}(C) \coloneqq \min\{\mathrm{rank}(A) : A \in C\}$. For a field $\mathbb{F}$ and a graph G define the set of matrices

$$M^{\mathbb{F}}(G) \coloneqq \{A \in \mathbb{F}^{V(G)\times V(G)} : \forall u,v\ A_{vv} \neq 0, \{u,v\} \notin E(G) \Rightarrow A_{uv} = 0\}.$$

Let $\mathrm{R}^{\mathbb{F}}(G) \coloneqq \mathrm{rank}(M^{\mathbb{F}}(G))$. The parameter $\mathrm{R}^{\mathbb{F}}(G)$ was introduced by Haemers in [Hae79] and is known as the Haemers bound. The fractional Haemers bound $\mathrm{R}^{\mathbb{F}}_f$ was studied by Anna Blasiak in [Bla13b] and was recently shown to be $\boxtimes$-multiplicative by Bukh and Cox in [BC18]. From this it is not hard to prove that $\mathrm{R}^{\mathbb{F}}_f \in \mathbf{X}(\mathcal{G})$. Bukh and Cox in [BC18] furthermore prove a separation result: for any field $\mathbb{F}$ of nonzero characteristic and any $\varepsilon > 0$, there is a graph G such that for any field $\mathbb{F}'$ with $\mathrm{char}(\mathbb{F}) \neq \mathrm{char}(\mathbb{F}')$ the inequality $\mathrm{R}^{\mathbb{F}}_f(G) < \varepsilon\, \mathrm{R}^{\mathbb{F}'}_f(G)$ holds. This separation result implies that there are infinitely many elements in $\mathbf{X}(\mathcal{G})$!

Fractional orthogonal rank

In [CMR+14] the orthogonal rank $\xi(G)$ and its fractional version the projective rank $\xi_f(G)$ are studied. It easily follows from results in [CMR+14] that $G \mapsto \xi_f(\overline{G})$ is in $\mathbf{X}(\mathcal{G})$.

General fractional parameters

We will prove something general about fractional parameters. Define the *lexicographic product* $G \ltimes H$ by

$$\begin{aligned} V(G \ltimes H) &= V(G) \times V(H) \\ E(G \ltimes H) &= \big\{\{(g,h),(g',h')\} : \{g,g'\} \in E(G) \\ &\qquad\qquad \text{or } (g = g' \text{ and } \{h,h'\} \in E(H))\big\}. \end{aligned}$$

The lexicographic product satisfies $G \ltimes H = \overline{\overline{G} \ltimes \overline{H}}$. Also define the *or-product* $G * H$ by

$$\begin{aligned} V(G * H) &= V(G) \times V(H) \\ E(G * H) &= \big\{\{(g,h),(g',h')\} : \{g,g'\} \in E(G) \text{ or } \{h,h'\} \in E(H)\big\}. \end{aligned}$$

The or-product and the strong graph product are related by $G * H = \overline{\overline{G} \boxtimes \overline{H}}$. The strong graph product gives a subgraph of the lexicographic product, which gives a subgraph of the or-product,

$$G \boxtimes H \subseteq G \ltimes H \subseteq G * H.$$

Therefore, $G * H \leqslant G \ltimes H \leqslant G \boxtimes H$. Finally, $G \ltimes \overline{K_d} = G * \overline{K_d}$, and of course $G \boxtimes \overline{K_d} = G^{\sqcup d}$.

We will prove: if $\phi : \mathcal{G} \to \mathbb{R}_{\geq 0}$ is $\boxtimes$-submultiplicative, $\sqcup$-subadditive and $\leqslant$-monotone, then the fractional parameter ϕ_f (defined in (3.2)) is again $\boxtimes$-submultiplicative, $\sqcup$-subadditive and $\leqslant$-monotone. Moreover, if $\phi : \mathcal{G} \to \mathbb{N}$ is $\leqslant$-monotone and satisfies

$$\forall G, H \in \mathcal{G} \quad \phi(G \ltimes H) \geq \phi(G \ltimes \overline{K_{\phi(H)}})$$

then ϕ_f is $\ltimes$-supermultiplicative and, more importantly, ϕ_f is $\boxtimes$-supermultiplicative.

Lemma 3.11. *Let $\phi : \mathcal{G} \to \mathbb{R}_{\geq 0}$.*

(i) *If ϕ is $\sqcup$-superadditive, then ϕ_f is $\sqcup$-superadditive.*

(ii) *If ϕ is $\leqslant$-monotone, then ϕ_f is $\leqslant$-monotone.*

(iii) *If ϕ is $\sqcup$-subadditive and $\leqslant$-monotone, then ϕ_f is $\sqcup$-subadditive.*

(iv) *If $\forall n \in \mathbb{N}\ \phi(\overline{K_n}) = n$, then $\forall n \in \mathbb{N}\ \phi_f(\overline{K_n}) = n$.*

(v) *If ϕ is $\boxtimes$-submultiplicative and $\leqslant$-monotone, then ϕ_f is $\boxtimes$-submultiplicative.*

Proof. Let $G, H \in \mathcal{G}$. Let $d \in \mathbb{N}$.

(i) The lexicographic product distributes over the disjoint union,

$$(G \sqcup H) \ltimes \overline{K_d} = (G \ltimes \overline{K_d}) \sqcup (H \ltimes \overline{K_d}).$$

By superadditivity,

$$\phi((G \ltimes \overline{K_d}) \sqcup (H \ltimes \overline{K_d})) \geq \phi(G \ltimes \overline{K_d}) + \phi(H \ltimes \overline{K_d}).$$

Therefore,

$$\begin{aligned}
\phi_f(G \sqcup H) &= \inf_d \frac{\phi((G \sqcup H) \ltimes \overline{K_d})}{d} \\
&= \inf_d \frac{\phi((G \ltimes \overline{K_d}) \sqcup (H \ltimes \overline{K_d}))}{d} \\
&\geq \inf_d \frac{\phi(G \ltimes \overline{K_d})}{d} + \frac{\phi(H \ltimes \overline{K_d})}{d} \\
&\geq \inf_{d_1} \frac{\phi(G \ltimes \overline{K_{d_1}})}{d_1} + \inf_{d_2} \frac{\phi(H \ltimes \overline{K_{d_2}})}{d_2} \\
&= \phi_f(G) + \phi_f(H).
\end{aligned}$$

(ii) Let $G \leqslant H$. Then $G \ltimes \overline{K_d} \leqslant H \ltimes \overline{K_d}$. Thus $\phi(G \ltimes \overline{K_d}) \leq \phi(H \ltimes \overline{K_d})$. Therefore $\phi_f(G) \leq \phi_f(H)$.

(iii) We have $G \ltimes \overline{K_d} \leqslant G \boxtimes \overline{K_d} = G^{\sqcup d}$. Thus by monotonicity and subadditivity

$$\phi(G \ltimes \overline{K_d}) \leq d\phi(G)$$

and, for $d, e \in \mathbb{N}$,

$$\phi(G \ltimes \overline{K_{de}}) = \phi((G \ltimes \overline{K_d}) \ltimes \overline{K_e}) \leq e\phi(G \ltimes \overline{K_d}).$$

We use this inequality to get, for $d_1, d_2 \in \mathbb{N}$,

$$\frac{\phi(G \ltimes \overline{K_{d_1}})}{d_1} + \frac{\phi(H \ltimes \overline{K_{d_2}})}{d_2} \geq \frac{\phi(G \ltimes \overline{K_{d_1 d_2}}) + \phi(H \ltimes \overline{K_{d_1 d_2}})}{d_1 d_2}$$

From subadditivity follows

$$\begin{aligned}\frac{\phi(G \ltimes \overline{K_{d_1 d_2}}) + \phi(H \ltimes \overline{K_{d_1 d_2}})}{d_1 d_2} &\geq \frac{\phi((G \ltimes \overline{K_{d_1 d_2}}) \sqcup (H \ltimes \overline{K_{d_1 d_2}}))}{d_1 d_2}\\ &= \frac{\phi((G \sqcup H) \ltimes \overline{K_{d_1 d_2}})}{d_1 d_2}\\ &\geq \phi_f(G \sqcup H).\end{aligned}$$

We conclude $\phi_f(G) + \phi_f(H) \geq \phi_f(G \sqcup H)$.

(iv) Let $n \in \mathbb{N}$. Then $\phi_f(\overline{K_n}) = \inf_d \phi(\overline{K_n} \ltimes \overline{K_d})/d = \inf_d \phi(\overline{K_{nd}})/d = n$.

(v) Let $d_1, d_2 \in \mathbb{N}$. We claim

$$(G \boxtimes H) \ltimes \overline{K_{d_1 d_2}} \leq (G \ltimes \overline{K_{d_1}}) \boxtimes (H \ltimes \overline{K_{d_2}}).$$

This is the same as saying there is a graph homomorphism

$$\overline{(G \boxtimes H) \ltimes \overline{K_{d_1 d_2}}} \to \overline{(G \ltimes \overline{K_{d_1}}) \boxtimes (H \ltimes \overline{K_{d_2}})},$$

which is the same as saying there is a graph homomorphism

$$(\overline{G} * \overline{H}) \ltimes K_{d_1 d_2} \to (\overline{G} \ltimes K_{d_1}) * (\overline{H} \ltimes K_{d_2}),$$

where $*$ denotes the or-product of graphs. One verifies that $(g, h, (i, j)) \mapsto ((g, i), (h, j))$ is such a graph homomorphism, proving the claim. The claim together with monotonicity and submultiplicativity gives

$$\phi((G \boxtimes H) \ltimes \overline{K_{d_1 d_2}}) \leq \phi((G \ltimes \overline{K_{d_1}}) \boxtimes (H \ltimes \overline{K_{d_2}})) \leq \phi(G \ltimes \overline{K_{d_1}})\phi(H \ltimes \overline{K_{d_2}}).$$

Therefore

$$\begin{aligned}\phi_f(G \boxtimes H) &= \inf_d \frac{\phi((G \boxtimes H) \ltimes \overline{K_d})}{d}\\ &= \inf_{d_1, d_2} \frac{\phi((G \boxtimes H) \ltimes \overline{K_{d_1 d_2}})}{d_1 d_2}\\ &\leq \inf_{d_1, d_2} \frac{\phi(G \ltimes \overline{K_{d_1}})}{d_1} \frac{\phi(H \ltimes \overline{K_{d_2}})}{d_2}\\ &= \phi_f(G)\phi_f(H).\end{aligned}$$

This concludes the proof of the lemma. $\square$

Lemma 3.12. *Let $\phi : \mathcal{G} \to \mathbb{N}$ satisfy*

$$\forall G, H \in \mathcal{G} \quad \phi(G \ltimes H) \geq \phi(G \ltimes \overline{K_{\phi(H)}}). \tag{3.3}$$

Then

$$\inf_H \frac{\phi(G \ltimes H)}{\phi(H)} = \inf_d \frac{\phi(G \ltimes \overline{K_d})}{d}.$$

Proof. From (3.3) follows

$$\frac{\phi(G \ltimes H)}{\phi(H)} \geq \frac{\phi(G \ltimes \overline{K_{\phi(H)}})}{\phi(H)},$$

and so

$$\frac{\phi(G \ltimes H)}{\phi(H)} \geq \inf_d \frac{\phi(G \ltimes \overline{K_d})}{d}.$$

We take the infimum over H to get

$$\inf_H \frac{\phi(G \ltimes H)}{\phi(H)} \geq \inf_d \frac{\phi(G \ltimes \overline{K_d})}{d}.$$

The inequality in the other direction,

$$\inf_H \frac{\phi(G \ltimes H)}{\phi(H)} \leq \inf_d \frac{\phi(G \ltimes \overline{K_d})}{d},$$

is trivially true. □

Lemma 3.13. *Let $\phi : \mathcal{G} \to \mathbb{N}$ be $\leqslant$-monotone and satisfy*

$$\forall G, H \in \mathcal{G} \quad \phi(G \ltimes H) \geq \phi(G \ltimes \overline{K_{\phi(H)}}).$$

Then ϕ_f is $\ltimes$- and $\boxtimes$-supermultiplicative.

Proof. Let $A, B \in \mathcal{G}$. We have $A \boxtimes B \geqslant A \ltimes B$, so

$$\phi_f(A \boxtimes B) \geq \phi_f(A \ltimes B).$$

It remains to show $\phi_f(A \ltimes B) \geq \phi_f(A)\phi_f(B)$. We have

$$\frac{\phi(A \ltimes B \ltimes H)}{\phi(H)} = \frac{\phi(A \ltimes (B \ltimes H))}{\phi(B \ltimes H)} \frac{\phi(B \ltimes H)}{\phi(H)},$$

which implies

$$\frac{\phi(A \ltimes B \ltimes H)}{\phi(H)} \geq \inf_{H'} \frac{\phi(A \ltimes H')}{\phi(H')} \inf_{H''} \frac{\phi(B \ltimes H'')}{\phi(H'')} = \phi_f(A)\phi_f(B).$$

Take the infimum over H to obtain $\phi_f(A \ltimes B) \geq \phi_f(A)\phi_f(B)$. □

Theorem 3.14. *Let $\phi : \mathcal{G} \to \mathbb{N}$ be $\sqcup$-additive, $\boxtimes$-submultiplicative, $\leqslant$-monotone and $\overline{K_n}$-normalised and satisfy*

$$\forall G, H \in \mathcal{G} \quad \phi(G \ltimes H) \geq \phi(G \ltimes \overline{K_{\phi(H)}}).$$

Then ϕ_f is in $\mathbf{X}(\mathcal{G})$.

Proof. This follows from Lemma 3.11, Lemma 3.12 and Lemma 3.13. □

3.4 Conclusion

In this chapter we introduced a new connection between Strassen's theory of asymptotic spectra and the Shannon capacity of graphs. In particular, we characterised the Shannon capacity (which is defined as a supremum) as a minimisation over elements in the asymptotic spectrum of graphs. Known elements in the asymptotic spectrum of graphs include the fractional clique cover number, the Lovász theta number, the projective rank and the fractional Haemers bound. We are left with a clear goal for future work: find all elements in the asymptotic spectrum of graphs.

Chapter 4

The asymptotic spectrum of tensors; exponent of matrix multiplication

This chapter is based on joint work with Matthias Christandl and Péter Vrana [CVZ18].

4.1 Introduction

This chapter is about tensors $t \in \mathbb{F}^{n_1} \otimes \cdots \otimes \mathbb{F}^{n_k}$ and their asymptotic properties. The theory of asymptotic spectra of Chapter 2 was developed by Strassen exactly for the purpose of understanding the asymptotic properties of tensors. This chapter is expository and provides the necessary background for understanding Chapter 5 and Chapter 6.

Let us first define the asymptotic properties of interest and discuss some of their applications. We need the concepts *restriction*, *tensor Kronecker product* and *diagonal* tensor. Let $s \in \mathbb{F}^{n_1} \otimes \cdots \otimes \mathbb{F}^{n_k}$ and $t \in \mathbb{F}^{m_1} \otimes \cdots \otimes \mathbb{F}^{m_k}$ be tensors. We say *s restricts to t*, and write $s \geqslant t$, if there are linear maps $A_i : \mathbb{F}^{n_i} \to \mathbb{F}^{m_i}$ such that $t = (A_1 \otimes \cdots \otimes A_k) \cdot s$. The *tensor Kronecker product* of s and t is the element $s \otimes t \in \mathbb{F}^{n_1 m_1} \otimes \cdots \otimes \mathbb{F}^{n_k m_k}$ with coordinates $(s \otimes t)_{i,j} = s_i t_j$. We naturally define the direct sum $s \oplus t \in \mathbb{F}^{n_1+m_1} \otimes \cdots \otimes \mathbb{F}^{n_k+m_k}$. We define the *diagonal tensors* $\langle n \rangle = \sum_{i=1}^{n} e_i \otimes \cdots \otimes e_i$ for $n \in \mathbb{N}$, where $e_1, \ldots, e_n$ is the standard basis of $\mathbb{F}^n$. The *tensor rank* $\mathrm{R}(t)$ is the smallest number $n \in \mathbb{N}$ such that t can be written as a sum of simple tensors; a simple tensor being a tensor of the form $v_1 \otimes \cdots \otimes v_k$. Equivalently, $\mathrm{R}(t) = \min\{n \in \mathbb{N} : t \leqslant \langle n \rangle\}$. The *asymptotic rank* is the regularisation $\underset{\sim}{\mathrm{R}}(t) = \lim_{n\to\infty} \mathrm{R}(t^{\otimes n})^{1/n}$. While tensor rank is known to be hard to compute [Hås90, Shi16], we do not know whether asymptotic rank is hard to compute.

The exponent of matrix multiplication

The motivating example for studying asymptotic rank is the problem of finding the exponent of matrix multiplication ω. Recall from the introduction that ω

is the infimum over $a \in \mathbb{R}$ such that any two $n \times n$ matrices can be multiplied using $\mathcal{O}(n^a)$ arithmetic operations (in the algebraic circuit model). It turns out (see [BCS97]) that ω is characterised by the asymptotic rank $\underset{\sim}{\mathrm{R}}(\langle 2,2,2\rangle)$ of the matrix multiplication tensor

$$\langle 2,2,2\rangle = \sum_{i,j,k\in[2]} e_{ij} \otimes e_{jk} \otimes e_{ki} \in \mathbb{F}^4 \otimes \mathbb{F}^4 \otimes \mathbb{F}^4.$$

Namely $\underset{\sim}{\mathrm{R}}(\langle 2,2,2\rangle) = 2^\omega$. We know the trivial lower bound $2 \leq \omega$, see Section 4.3. We know the (non-trivial) upper bound $\omega \leq 2.3728639$, which is by Coppersmith and Winograd [CW90] and improvements by Stothers, Williams and Le Gall [Sto10, Wil12, LG14].

Asymptotic subrank and asymptotic restriction

Besides (asymptotic) rank, we naturally define *subrank* $\mathrm{Q}(t) = \max\{m \in \mathbb{N} : \langle m\rangle \leqslant t\}$ and the *asymptotic subrank* $\underset{\sim}{\mathrm{Q}}(t) = \lim_{n\to\infty} \mathrm{Q}(t^{\otimes n})^{1/n}$. Moreover, we say *s restricts asymptotically* to t, written $s \gtrsim t$, if there is a sequence of natural numbers $a(n) \in o(n)$ such that for all $n \in \mathbb{N}$

$$s^{\otimes n} \otimes \langle 2\rangle^{\otimes a(n)} \geqslant t^{\otimes n}.$$

One can prove (see [Str91]) that

$$s^{\otimes n} \otimes \langle 2\rangle^{\otimes o(n)} \geqslant t^{\otimes n} \quad \text{iff} \quad s^{\otimes (n+o(n))} \geqslant t^{\otimes n}.$$

Our goal is to understand asymptotic restriction, asymptotic rank and asymptotic subrank.

More connections: quantum information, combinatorics, algebraic property testing

Besides matrix multiplication, other applications of asymptotic restriction of tensors, asymptotic rank of tensors and asymptotic subrank of tensors include 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], see Chapter 5, and bounding the query complexity of certain properties in algebraic property testing [KS08, BCSX10, Sha09, BX15, HX17, FK14].

This chapter is organised as follows. In Section 4.2 we briefly discuss the semiring of tensors, the asymptotic spectrum of tensors, and asymptotic rank and subrank. In Section 4.3 we discuss the gauge points, a simple construction of finitely many elements in the asymptotic spectrum of tensors. In Section 4.4 we discuss the Strassen support functionals: a family of elements in the asymptotic spectrum of "oblique" tensors. This family is parametrised by probability distributions on $[k]$. In Section 4.5 we discuss an extension of the support functionals, called the Strassen upper support functionals, which have the potential to be universal. Finally, in Section 4.6 we prove a new result: we show how asymptotic slice rank is related to the support functionals.

4.2 The asymptotic spectrum of tensors

Let us properly set up the semiring of tensors and the asymptotic spectrum. For the proofs we refer to [Str87, Str88, Str91].

4.2.1 The semiring of tensor equivalence classes $\mathcal{T}$

We begin by putting an equivalence relation on tensors. For example, we want to identify isomorphic tensors and also, for any tensor $t \in \mathbb{F}^{n_1} \otimes \cdots \otimes \mathbb{F}^{n_k}$, we want to identify t with $t \oplus 0$, where $0 \in \mathbb{F}^{m_1} \otimes \cdots \otimes \mathbb{F}^{m_k}$ is a zero tensor of any format.

We say s *is isomorphic to* t, and write $s \cong t$, if there are bijective linear maps $A_i : \mathbb{F}^{m_i} \to \mathbb{F}^{n_i}$ such that $t = (A_1, \ldots, A_k) \cdot s$.

We say s and t are *equivalent*, and write $s \sim t$, if there are zero tensors $s_0 = 0 \in \mathbb{F}^{a_1} \times \cdots \times \mathbb{F}^{a_k}$ and $t_0 = 0 \in \mathbb{F}^{b_1} \times \cdots \times \mathbb{F}^{b_k}$ such that $s \oplus s_0 \cong t \oplus t_0$. The equivalence relation $\sim$ is in fact the equivalence relation generated by the restriction preorder $\leqslant$.

Let $\mathcal{T}$ be the set of $\sim$-equivalence classes of k-tensors over $\mathbb{F}$, for some fixed k and field $\mathbb{F}$. The direct sum and the 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, but we will not make this distinction).

4.2.2 Strassen preorder via restriction

Restriction $\leqslant$ induces a partial order on $\mathcal{T}$, which behaves well with respect to the semiring operations, and naturally $n \leq m$ if and only if $\langle n \rangle \leqslant \langle m \rangle$. Therefore, restriction $\leqslant$ is a Strassen preorder on $\mathcal{T}$.

4.2.3 The asymptotic spectrum of tensors $\mathbf{X}(\mathcal{T})$

Let $S \subseteq \mathcal{T}$ be a subsemiring. Let

$$\mathbf{X}(S) = \mathbf{X}(S, \leqslant) = \{\phi \in \mathrm{Hom}(S, \mathbb{R}_{\geq 0}) : \forall a, b \in S \; a \leqslant b \Rightarrow \phi(a) \leq \phi(b)\}.$$

We call $\mathbf{X}(S)$ the asymptotic spectrum of S and we call $\mathbf{X}(\mathcal{T})$ the *asymptotic spectrum of k-tensors over $\mathbb{F}$*.

Theorem 4.1 ([Str88]). *Let $s, t \in S$. Then $s \lesssim t$ iff $\forall \phi \in \mathbf{X}(S)\ \phi(s) \leq \phi(t)$.*

Proof. This follows from Theorem 2.13. □

We refer to Chapter 2 for a discussion of the topological properties of $\mathbf{X}(S)$.

Remark 4.2. We mention that $\mathbf{X}(S)$ may equivalently be defined with degeneration $\trianglerighteq$ instead of restriction $\geq$. Over $\mathbb{C}$, we say f *degenerates to* g, written, $f \trianglerighteq g$, if $f \cong f'$ and $g \cong g'$ and g' is in the Euclidean closure (or equivalently Zariski closure) of the orbit $\mathrm{GL}_{n_1} \times \cdots \times \mathrm{GL}_{n_k} \cdot f'$. It is a nontrivial fact from algebraic geometry (see [Kra84, Lemma III.2.3.1] or [BCS97]) that there is a degeneration $f \trianglerighteq g$ if and only if there are matrices A_i with entries polynomial in ε such that $(A_1, \ldots, A_k) \cdot f = \varepsilon^d g + \varepsilon^{d+1} g_1 + \cdots + \varepsilon^{d+e} g_e$ for some elements $g_1, \ldots, g_e$. The latter definition of degeneration is valid when $\mathbb{C}$ is replaced by an arbitrary field $\mathbb{F}$ and that is how degeneration is defined for an arbitrary field. Degeneration is weaker than restriction: $f \geq g$ implies $f \trianglerighteq g$. Asymptotically, however, the notions coincide: $f \gtrsim g$ if and only if $f^{\otimes n} \otimes \langle 2 \rangle^{\otimes o(n)} \trianglerighteq g^{\otimes n}$. We mention that, analogous to restriction, degeneration gives rise to *border rank* and *border subrank*, $\underline{\mathrm{R}}(f) = \min\{r \in \mathbb{N} : f \trianglelefteq \langle r \rangle\}$ and $\underline{\mathrm{Q}}(f) = \max\{s \in \mathbb{N} : \langle s \rangle \trianglelefteq f\}$ respectively.

4.2.4 Asymptotic rank and asymptotic subrank

The abstract theory of asymptotic spectra characterises asymptotic subrank and asymptotic rank as follows.

Corollary 4.3. *Let $S \subseteq \mathcal{T}$ be a subsemiring. Let $a \in S$. Then*

$$\underset{\sim}{\mathrm{Q}}(a) = \min_{\phi \in \mathbf{X}(S)} \phi(a) \tag{4.1}$$

$$\underset{\sim}{\mathrm{R}}(a) = \max_{\phi \in \mathbf{X}(S)} \phi(a). \tag{4.2}$$

Proof. Statement (4.2) follows from Corollary 2.14, since either $a = 0$ or $a \geqslant 1$. For statement (4.1), if $t^{\otimes k} \geqslant 2$ for some $k \in \mathbb{N}$, then we apply Corollary 2.15. Otherwise, one can show that $\underset{\sim}{\mathrm{Q}}(t)$ equals 0 or 1 using the gauge points of the next section (see [Str88, Lemma 3.7]). □

Remark 4.4. One verifies that $\underset{\sim}{\mathrm{R}}$ and $\underset{\sim}{\mathrm{Q}}$ are $\leqslant$-monotones and have value n on $\langle n \rangle$. They are not universal spectral points however. Namely, the asymptotic rank of each of the three tensors

$$\langle 2,1,1 \rangle = e_1 \otimes e_1 \otimes 1 + e_2 \otimes e_2 \otimes 1 \ \in \ \mathbb{F}^2 \otimes \mathbb{F}^2 \otimes \mathbb{F}^1$$
$$\langle 1,1,2 \rangle = e_1 \otimes 1 \otimes e_1 + e_2 \otimes 1 \otimes e_2 \ \in \ \mathbb{F}^2 \otimes \mathbb{F}^1 \otimes \mathbb{F}^2$$

$$\langle 1,2,2\rangle = 1\otimes e_1\otimes e_1 + 1\otimes e_2\otimes e_2 \,\in\, \mathbb{F}^1\otimes\mathbb{F}^2\otimes\mathbb{F}^2$$

equals 2, whereas their tensor product equals the matrix multiplication tensor $\langle 2,2,2\rangle$ whose tensor rank equals 7 and whose asymptotic rank is thus at most 7, i.e. strictly smaller than 2^3. Therefore, asymptotic rank is not multiplicative. On the other hand, the asymptotic subrank of each of the above three tensors equals 1, whereas the asymptotic subrank of $\langle 2,2,2\rangle$ equals 4, see Chapter 5. Therefore, asymptotic subrank is not multiplicative.

Goal 4.5. Our goal is now to explicitly describe elements in $\mathbf{X}(\mathcal{T})$, universal spectral points, or more modestly, to describe elements in $\mathbf{X}(S)$ for interesting subsemirings $S\subseteq\mathcal{T}$.

Strassen constructed a finite family of elements in $\mathbf{X}(\mathcal{T})$, the gauge points, and an infinite family of elements in $\mathbf{X}(\{\text{oblique tensors}\})$, the support functionals. The support functionals are powerful enough to determine the asymptotic subrank of any "tight tensor". Tight tensors are discussed in Chapter 5. In Chapter 6 we construct an infinite family in $\mathbf{X}(\{k\text{-tensors over }\mathbb{C}\})$, the quantum functionals. In the rest of this chapter we discuss the gauge points and the support functionals. We will focus on the case $k=3$ for clarity of exposition.

4.3 Gauge points $\zeta^{(i)}$

Strassen in [Str88] introduced a finite family of elements in $\mathbf{X}(\mathcal{T})$, called the gauge points. We focus on 3-tensors, but the construction generalises immediately to k-tensors. Let $V_i=\mathbb{F}^{n_i}$. Let $t\in V_1\otimes V_2\otimes V_3$. Let $i\in[3]$. Let $\mathrm{flatten}_i(t)$ be the image of t under the grouping $V_1\otimes V_2\otimes V_3\to V_i\otimes(\bigotimes_{j\neq i}V_j)$. We think of $\mathrm{flatten}_i(t)$ as a matrix. Let $\zeta^{(i)}:\mathcal{T}\to\mathbb{N}: t\mapsto \mathrm{rank}(\mathrm{flatten}_i(t))$, with rank denoting matrix rank. We call $\zeta^{(1)},\zeta^{(2)},\zeta^{(3)}$ the *gauge points.* From the properties of matrix rank follows directly that $\zeta^{(i)}$ is multiplicative under $\otimes$, additive under $\oplus$, monotone under restriction $\leqslant$ (and under degeneration $\trianglelefteq$) and normalised to 1 on $\langle 1\rangle = e_1\otimes e_1\otimes e_1$.

Theorem 4.6. *$\zeta^{(1)}$, $\zeta^{(2)}$, $\zeta^{(3)}\in\mathbf{X}(\mathcal{T})$.*

Recall, $\underset{\sim}{\mathrm{Q}}(t)\leq\phi(t)\leq\underset{\sim}{\mathrm{R}}(t)$ for all $\phi\in\mathbf{X}(\mathcal{T})$. In particular, $\max_i\zeta^{(i)}(t)\leq\underset{\sim}{\mathrm{R}}(t)$. *We do not know whether* $\max_{i\in[3]}\zeta^{(i)}$ *equals* $\underset{\sim}{\mathrm{R}}$. To be precise: we do not know any t for which $\max_i\zeta^{(i)}(t)<\underset{\sim}{\mathrm{R}}(t)$ and we do not know a proof that $\max_i\zeta^{(i)}(t)=\underset{\sim}{\mathrm{R}}(t)$ for all t. There are various families of tensors t for which $\max_i\zeta^{(i)}(t)=\underset{\sim}{\mathrm{R}}(t)$ is proven. We will see such a family in Section 5.4.2. For the matrix multiplication tensor $\langle 2,2,2\rangle$ we have $4=\max_i\zeta^{(i)}(\langle 2,2,2\rangle)\leq 2^\omega$, so $\max_i\zeta^{(i)}(t)=\underset{\sim}{\mathrm{R}}(t)$ would imply that the matrix multiplication exponent ω equals 2.

On the other hand, $\underset{\sim}{\mathrm{Q}}(t)\leq\min_i\zeta^{(i)}(t)$. *There exist t for which $\underset{\sim}{\mathrm{Q}}(t)$ is strictly smaller than* $\min_{i\in[3]}\zeta^{(i)}(t)$. To show this strict inequality we need another technique of Strassen: the support functionals. These are the topic of the next section.

4.4 Support functionals ζ^θ

Strassen in [Str91] constructed an infinite family of elements in the asymptotic spectrum of "oblique" k-tensors, called the support functionals. In this section we explain the construction of the support functionals. The support functionals provide the benchmark for our new quantum functionals (Chapter 6), and are relevant in the context of combinatorial problems like the cap set problem (Section 5.4.2). For clarity of exposition we focus on 3-tensors. The ideas extend directly to k-tensors.

Oblique tensors are tensors for which in some basis the support has the following special structure. Let $t \in \mathbb{F}^{n_1} \otimes \mathbb{F}^{n_2} \otimes \mathbb{F}^{n_3}$. Let $e_1, \ldots, e_{n_i}$ be the standard basis of $\mathbb{F}^{n_i}$. Write $t = \sum_{i,j,k} t_{ijk}\, e_i \otimes e_j \otimes e_k$. Let $[n_i] \coloneqq \{1, 2, \ldots, n_i\}$. Let $\mathrm{supp}(t) \coloneqq \{(i,j,k) : t_{ijk} \neq 0\} \subseteq [n_1] \times [n_2] \times [n_3]$ be the support of t with respect to the standard basis. Let $[n_i]$ have the natural ordering $1 < 2 < \cdots < [n_i]$ and let $[n_1] \times [n_2] \times [n_3]$ have the product order, denoted by $\leq$. That is, $x \leq y$ if for all $i \in [3]$ holds $x_i \leq y_i$. We call $\mathrm{supp}(t)$ *oblique* if $\mathrm{supp}(t)$ is an antichain with respect to $\leq$, i.e. if any two elements in $\mathrm{supp}(t)$ are incomparable with respect to $\leq$. We call a tensor t *oblique* if $\mathrm{supp}(g \cdot t)$ is oblique for some group element $g \in G(t) \coloneqq \mathrm{GL}_{n_1} \times \mathrm{GL}_{n_2} \times \mathrm{GL}_{n_3}$. The family of oblique tensors is a semiring under $\oplus$ and $\otimes$.

Not all tensors are oblique. Obliqueness is not a generic property (see Proposition 6.21). However, many tensors that are of interest in algebraic complexity theory are oblique, notably the matrix multiplication tensors

$$\langle a,b,c\rangle \coloneqq \sum_{i\in[a]} \sum_{j\in[b]} \sum_{k\in[c]} e_{ij} \otimes e_{jk} \otimes e_{ki} \in \mathbb{F}^{ab} \otimes \mathbb{F}^{bc} \otimes \mathbb{F}^{ca}.$$

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. Given finite sets $X_1, \ldots, X_k$ and a probability distribution $P \in \mathcal{P}(X_1 \times \cdots \times X_k)$ on the product set $X_1 \times \cdots \times X_k$ we denote the marginal distribution of P on X_i by P_i, that is, $P_i(a) = \sum_{x : x_i = a} P(x)$ for any $a \in X_i$.

Definition 4.7. Let $\theta \in \Theta \coloneqq \mathcal{P}([3])$. For $t \in \mathbb{F}^{n_1} \otimes \mathbb{F}^{n_2} \otimes \mathbb{F}^{n_3} \setminus 0$ with $\mathrm{supp}(t)$ oblique define

$$\zeta^\theta(t) \coloneqq \max\{2^{\sum_{i=1}^3 \theta(i) H(P_i)} : P \in \mathcal{P}(\mathrm{supp}(t))\}.$$

We call the ζ^θ for $\theta \in \Theta$ the *support functionals*.

Theorem 4.8. *$\zeta^\theta \in \mathbf{X}(\{\text{oblique}\})$ for $\theta \in \Theta$.*

We work towards the proof of Theorem 4.8. For $p \in [0,1]$ let $h(p)$ be the binary entropy function, $h(p) \coloneqq -p\log_2 p - (1-p)\log_2(1-p)$, i.e. $h(p)$ is the Shannon entropy of the probability vector $(p, 1-p)$. The following properties of the Shannon entropy are well-known.

Lemma 4.9.

(i) $H(P \otimes Q) = H(P) + H(Q)$ *for* $P \in \mathcal{P}(X_1)$, $Q \in \mathcal{P}(X_2)$.

(ii) $H(P) \leq H(P_1) + H(P_2)$ *for* $P \in \mathcal{P}(X_1 \times X_2)$.

(iii) $H(pP \oplus (1-p)Q) = pH(P) + (1-p)H(Q) + h(p)$ *for* $P, Q \in \mathcal{P}(X)$, $p \in [0,1]$.

(iv) $2^a + 2^b = \max_{0\leq p\leq 1} 2^{pa+(1-p)b+h(p)}$ *for* $a, b \in \mathbb{R}$.

For $X \subseteq [n_1] \times [n_2] \times [n_3]$ let $X_{\leq} = \{y \in [n_1] \times [n_2] \times [n_3] : \exists x \in X\ y \leq x\}$ be the downward closure of X. Let $\max(X) \coloneqq \{y \in X : \forall x \in X\, y \leq x \Rightarrow y = x\}$ be the maximal points of X with respect to $\leq$. Let S_n be the symmetric group of permutations of $[n]$. Then the product group $S_{n_1} \times S_{n_2} \times S_{n_3}$ acts naturally on $[n_1] \times [n_2] \times [n_3]$.

Lemma 4.10. *Let $t \in \mathbb{F}^{n_1} \otimes \mathbb{F}^{n_2} \otimes \mathbb{F}^{n_3}$. For every $g \in G(t)$ there is a triple of permutations $w \in W(t) \coloneqq S_{n_1} \times S_{n_2} \times S_{n_3}$ with $w \cdot \max(\operatorname{supp}(g \cdot t)) \subseteq \operatorname{supp}(t)_{\leq}$.*

Proof. We prepare for the construction of w. Let $n \in \mathbb{N}$. Let $e_1, \ldots, e_n$ be the standard basis of $\mathbb{F}^n$. Let $g \in \mathrm{GL}_n$. Let $f_1, \ldots, f_n$ with $f_j = g \cdot e_j$ be the transformed basis of $\mathbb{F}^n$. Let $(E_i)_{i\in[n]}$ and $(F_j)_{j\in[n]}$ be the complete flags of $\mathbb{F}^n$ with

$$E_i = \operatorname{Span}\{e_i, e_{i+1}, \ldots, e_n\}$$
$$F_j = \operatorname{Span}\{f_j, f_{j+1}, \ldots, f_n\}.$$

Define the map

$$\pi : [n] \to [n] : j \mapsto \max\bigl\{i \in [n] : E_i \cap (f_j + F_{j+1}) \neq \emptyset\bigr\}. \tag{4.3}$$

We prove π is injective. Let $j, k \in [n]$ with $j \leq k$ and suppose $i = \pi(j) = \pi(k)$. Let $F^\times \coloneqq \mathbb{F} \setminus 0$. From (4.3) follows

$$(\mathbb{F}^\times e_i + E_{i+1}) \cap (f_j + F_{j+1}) \neq \emptyset \tag{4.4}$$
$$E_{i+1} \cap (f_j + F_{j+1}) = \emptyset \tag{4.5}$$
$$(\mathbb{F}^\times e_i + E_{i+1}) \cap (f_k + F_{k+1}) \neq \emptyset. \tag{4.6}$$

Suppose $j < k$. Then from (4.4) and (4.6) we obtain a contradiction to (4.5). We conclude that $j = k$. Thus π is injective.

For each $\mathbb{F}^{n_i}$ define as above the standard complete flag $(E^i_j)_{j\in[n_i]}$ of $\mathbb{F}^{n_i}$, the complete flag $(F^i_j)_{j\in[n_i]}$ corresponding to the basis given by g_i, and the permutation $\pi_i : [n_i] \to [n_i]$. Let $w = (\pi_1, \pi_2, \pi_3) \in W(t)$.

We will prove $w \cdot \max(\mathrm{supp}(g \cdot t)) \subseteq \mathrm{supp}(t)_{\leq}$. Let $y \in \max(\mathrm{supp}(g \cdot t))$. Let $x = w \cdot y$. By construction of π_i the intersection $E^i_{x_i} \cap (f^i_{y_i} + F^i_{y_i+1})$ is not empty. Choose

$$\overline{f^i_{y_i}} \in E^i_{x_i} \cap (f^i_{y_i} + F^i_{y_i+1}).$$

Let t^* be the multilinear map $\mathbb{F}^{n_1} \times \mathbb{F}^{n_2} \times \mathbb{F}^{n_3} \to \mathbb{F}$ with $t^*(e_i, e_j, e_k) = t_{ijk}$ for all $i \in [n_1], j \in [n_2], k \in [n_3]$. Then

$$t^*(\overline{f^1_{y_1}}, \overline{f^2_{y_2}}, \overline{f^3_{y_3}}) = t^*(f^1_{y_1}, f^2_{y_2}, f^3_{y_3}) + \sum_{\substack{z\in[n_1]\times[n_2]\times[n_3]:\\ z>y}} c_z\, t^*(f^1_{z_1}, f^2_{z_2}, f^3_{z_3}) \tag{4.7}$$

for some $c_z \in \mathbb{F}$. Since y is maximal in $\mathrm{supp}(g\cdot t)$, the sum over $z > y$ in (4.7) equals zero. We conclude $t^*(\overline{f^1_{y_1}}, \overline{f^2_{y_2}}, \overline{f^3_{y_3}}) = t^*(f^1_{y_1}, f^2_{y_2}, f^3_{y_3}) \neq 0$. Thus $t^*(E^1_{x_1} \times E^2_{x_2} \times E^3_{x_3})$ is not zero and thus $x \in \mathrm{supp}(t)_{\leq}$. □

Proof of Theorem 4.8. We prove ζ^θ on oblique tensors is $\otimes$-multiplicative, $\oplus$-additive, $\leqslant$-monotone and normalised to 1 on $\langle 1\rangle \coloneqq e_1 \otimes e_1 \otimes e_1$. The normalisation $\zeta^\theta(\langle 1\rangle) = 1$ is clear.

We prove ζ^θ is $\otimes$-supermultiplicative. Let $s \in \mathbb{F}^{n_1} \otimes \mathbb{F}^{n_2} \otimes \mathbb{F}^{n_3}$ and let $t \in \mathbb{F}^{m_1} \otimes \mathbb{F}^{m_2} \otimes \mathbb{F}^{m_3}$. Let $P \in \mathcal{P}(\mathrm{supp}(t))$ and $Q \in \mathcal{P}(\mathrm{supp}(s))$. Then the product $P \otimes Q \in \mathcal{P}(\mathrm{supp}(s \otimes t))$ has marginals $P_i \otimes Q_i$. Since $H(P_i \otimes Q_i) = H(P_i) + H(Q_i)$ (Lemma 4.9(i)), we conclude $\zeta^\theta(s)\zeta^\theta(t) \leq \zeta^\theta(s \otimes t)$.

We prove ζ^θ is $\otimes$-submultiplicative. For $P \in \mathcal{P}(\mathrm{supp}(t))$ and $\theta \in \Theta$ we use the notation $H_\theta(P) \coloneqq \sum_{i=1}^3 \theta(i) H(P_i)$. We naturally identify $\mathrm{supp}(t)$ with a subset of $[n_1] \times [n_2] \times [n_3] \times [m_1] \times [m_2] \times [m_3]$. Let $P \in \mathcal{P}(\mathrm{supp}(t))$. Let $P_{[3]}$ be the marginal distribution of P on $[n_1] \times [n_2] \times [n_3]$ and let $P_{3+[3]}$ be the marginal distribution of P on $[m_1] \times [m_2] \times [m_3]$. Then $H_\theta(P) \leq H_\theta(P_{[3]}) + H_\theta(P_{3+[3]})$ by Lemma 4.9(ii). We conclude $\zeta^\theta(s \otimes t) \leq \zeta^\theta(s)\zeta^\theta(t)$.

We prove ζ^θ is $\oplus$-additive. By definition

$$\begin{aligned}\zeta^\theta(s \oplus t) &= \max\{2^{H_\theta(P)} : P \in \mathcal{P}(\mathrm{supp}(s \oplus t))\}\\ &= \max\bigl\{\max_{0\leq p\leq 1} 2^{H_\theta(pP\oplus(1-p)Q)} : P \in \mathcal{P}(\mathrm{supp}(s)), Q \in \mathcal{P}(\mathrm{supp}(t))\bigr\}.\end{aligned}$$

From Lemma 4.9(iii) and (iv) follows

$$\begin{aligned}&\max\bigl\{\max_{0\leq p\leq 1} 2^{H_\theta(pP\oplus(1-p)Q)} : P \in \mathcal{P}(\mathrm{supp}(s)), Q \in \mathcal{P}(\mathrm{supp}(t))\bigr\}\\ &= \max\bigl\{\max_{0\leq p\leq 1} 2^{pH_\theta(P)+(1-p)H_\theta(Q)+h(p)} : P \in \mathcal{P}(\mathrm{supp}(s)), Q \in \mathcal{P}(\mathrm{supp}(t))\bigr\}\\ &= \max\bigl\{2^{H_\theta(P)} + 2^{H_\theta(Q)} : P \in \mathcal{P}(\mathrm{supp}(s)), Q \in \mathcal{P}(\mathrm{supp}(t))\bigr\}\end{aligned}$$

$$= \zeta^\theta(s) + \zeta^\theta(t).$$

We conclude $\zeta^\theta(s \oplus t) = \zeta^\theta(s) + \zeta^\theta(t)$.

We prove ζ^θ is $\leqslant$-monotone. Let $s \leqslant t$ with $\mathrm{supp}(s)$ and $\mathrm{supp}(t)$ oblique. Then there are linear maps A_i with $s = (A_1 \otimes A_2 \otimes A_3) \cdot t$. If A_1, A_2, A_3 are of the form $\mathrm{diag}(1, \ldots, 1, 0, \ldots, 0)$, then $\zeta^\theta(s) \leq \zeta^\theta(t)$. Suppose $g = (A_1, A_2, A_3) \in G(t)$. Let $P \in \mathcal{P}(\mathrm{supp}(t))$ maximise H_θ on $\mathcal{P}(\mathrm{supp}(t))$. Let $\sigma \in W$ such that $\sigma \cdot P$ has non-increasing marginals. Then $H_\theta(\sigma \cdot P) = H_\theta(P)$ and $\sigma \cdot P$ maximises H_θ on $\mathcal{P}(\mathrm{supp}(\sigma \cdot t))$. Then $\sigma \cdot P$ maximises H_θ on $\mathcal{P}(\mathrm{supp}(\sigma \cdot t)_{\leq})$ by Lemma 4.12 below. Let $Q \in \mathcal{P}(\mathrm{supp}(g \cdot t))$ maximise H_θ on $\mathcal{P}(\mathrm{supp}(g \cdot t))$. By Lemma 4.10 there is a $w \in W$ with $w \cdot \mathrm{supp}(g \cdot t) \subseteq \mathrm{supp}(\sigma \cdot t)_{\leq}$. Then $H_\theta(w \cdot Q) = H_\theta(Q) \leq H_\theta(\sigma \cdot P) = H_\theta(P)$. Thus $\max_{P \in \mathcal{P}(\mathrm{supp}(g \cdot t))} H_\theta(P) \leq \max_{P \in \mathrm{supp}(t)} H_\theta(P)$. We conclude $\zeta^\theta(g \cdot t) \leq \zeta^\theta(t)$ □

The following two lemmas finish the above proof of Theorem 4.8. Recall that in the proof we defined $H_\theta(P) \coloneqq \sum_{i=1}^3 \theta(i) H(P_i)$ for $\theta \in \Theta$.

Lemma 4.11 ([Str91, Prop. 2.1])**.** *Let $\Phi \subseteq [n_1] \times [n_2] \times [n_3]$. Let $P \in \mathcal{P}(\Phi)$. Let* $\mathrm{supp}(P)$ *be the support $\{x \in \Phi : P(x) \neq 0\}$. For $x \in \Phi$ define $h_P(x) \coloneqq -\sum_{i=1}^3 \theta(i) \log_2 P_i(x_i)$. Then P maximises H_θ on $\mathcal{P}(\Phi)$ if and only if*

$$\forall x \in \mathrm{supp}(P) \quad h_P(x) = \max_{y \in \Phi} h_P(y). \tag{4.8}$$

Proof. We write $H_\theta(P)$ in terms of h_P,

$$H_\theta(P) = \sum_{i=1}^{3} \theta(i)\, H(P_i) = \sum_{x \in \mathrm{supp}(P)} P(x)\, h_P(x). \tag{4.9}$$

For $Q \in \mathcal{P}(\Phi)$

$$\begin{aligned}
&\lim_{\varepsilon \to 0+} \frac{\mathrm{d}}{\mathrm{d}\varepsilon} H_\theta\big((1-\varepsilon)P + \varepsilon Q\big) \\
&= \lim_{\varepsilon \to 0+} \frac{\mathrm{d}}{\mathrm{d}\varepsilon} \sum_x \big((1-\varepsilon)P(x) + \varepsilon Q(x)\big)\, h_{(1-\varepsilon)P+\varepsilon Q}(x) \\
&= \sum_x P(x) \left(\sum_{i=1}^{3} \theta(i) \frac{P_i(x_i) - Q_i(x_i)}{P_i(x_i) \ln(2)} \right) + \sum_x \big(-P(x) + Q(x)\big)\, h_P(x) \\
&= \sum_x Q(x)\, h_P(x) - \sum_x P(x)\, h_P(x).
\end{aligned}$$

Therefore, since H_θ is continuous and concave, P maximises H_θ if and only if

$$\forall Q \in \mathcal{P}(\Phi) \quad \sum_x Q(x)\, h_P(x) - \sum_x P(x)\, h_P(x) \leq 0. \tag{4.10}$$

We will prove (4.10) is equivalent to (4.8). Suppose $\sum_x Q(x)h_P(x) \leq \sum_x P(x)h_P(x)$ for every $Q \in \mathcal{P}(\Phi)$. In particular $h_P(y) \leq \sum_x P(x)h_P(x)$ for every $y \in \Phi$, so $\max_{y\in\Phi} h_P(y) \leq \sum_x P(x)h_P(x)$. Then $\max_{y\in\Phi} h_P(y) = \sum_x P(x)h_P(x)$. We conclude $\max_{y\in\Phi} h_P(y) = h_P(x)$ for every $x \in \mathrm{supp}(P)$.

Suppose $\max_{y\in\Phi} h_P(y) = h_P(x)$ for every $x \in \mathrm{supp}(P)$. Then $h_P(y) \leq h_P(x)$ for every $Q \in \mathcal{P}(\Phi), y \in \mathrm{supp}(Q), x \in \mathrm{supp}(P)$. We conclude $\sum_x Q(x)h_P(x) \leq \sum_x P(x)h_P(x)$. □

Lemma 4.12 ([Str91, Cor. 2.2])**.** *Let $\Phi \subseteq [n_1] \times [n_2] \times [n_3]$. Let P maximise H_θ on $\mathcal{P}(\Phi)$. Suppose P_i is nonincreasing on $[n_i]$ for each $i \in [3]$. Then P maximises H_θ on $\mathcal{P}(\Phi_\leq)$ where $\Phi_\leq$ is the downward closure of Φ with respect to $\leq$.*

Proof. We know P satisfies (4.8). We will prove P satisfies (4.8) with Φ replaced by $\Phi_\leq$. Then we are done by Lemma 4.11. Let $x \in \Phi_\leq$. Then $x \leq y$ for some $y \in \Phi$. Then $(P_1(x_1), P_2(x_2), P_3(x_3)) \geq (P_1(y_1), P_2(y_2), P_3(y_3))$ since each P_i is nonincreasing. Then $h_P(x) \leq h_P(y)$. We conclude $\max_{\Phi_\leq} h_P \leq \max_\Phi h_P$. On the other hand, $\Phi \subseteq \Phi_\leq$. Therefore $\max_\Phi h_P \leq \max_{\Phi_\leq} h_P$. □

Using the support functionals Strassen managed to fully compute the asymptotic spectrum of several semirings generated by oblique tensors. We will see an example in Section 5.4.2.

4.5 Upper and lower support functionals $\zeta^\theta, \zeta_\theta$

In Section 4.4 we defined the support functionals $\zeta^\theta : \{\text{oblique}\} \to \mathbb{R}_{\geq 0}$ and proved that $\zeta^\theta \in \mathbf{X}(\{\text{oblique}\})$. From the general theory of asymptotic spectra (Chapter 2) we know ζ^θ is the restriction of some map $\phi : \{\text{tensors}\} \to \mathbb{R}_{\geq 0}$ in $\mathbf{X}(\mathcal{T})$. However, the proof of that fact was non-constructive. In other words, we know that ζ^θ can be extended to an element of $\mathbf{X}(\mathcal{T})$. In this short section we discuss a candidate extension proposed by Strassen, called the upper support functional. We also discuss a companion called the lower support functional.

For arbitrary $t \in \mathbb{F}^{n_1} \otimes \mathbb{F}^{n_2} \otimes \mathbb{F}^{n_3}$ the upper support functional and the lower support functional are defined as

$$\zeta^\theta(t) := \min_{g \in G(t)} \max\{2^{H_\theta(P)} : P \in \mathcal{P}(\mathrm{supp}(g \cdot t))\}$$

$$\zeta_\theta(t) := \max_{g \in G(t)} \max\{2^{H_\theta(P)} : P \in \mathcal{P}(\max(\mathrm{supp}(g \cdot t)))\}$$

with $G(t) := \mathrm{GL}_{n_1} \times \mathrm{GL}_{n_2} \times \mathrm{GL}_{n_3}$ and $H_\theta(P) := \sum_{i=1}^3 \theta(i) H(P_i)$. We summarise the known properties of the upper and lower support functional.

Theorem 4.13 ([Str91])**.** *Let $s \in \mathbb{F}^{n_1} \otimes \mathbb{F}^{n_2} \otimes \mathbb{F}^{n_3}$ and $t \in \mathbb{F}^{m_1} \otimes \mathbb{F}^{m_2} \otimes \mathbb{F}^{m_3}$. Let $\theta \in \Theta = \mathcal{P}([3])$.*

(i) $\zeta^\theta(\langle n\rangle) = n$ *for* $n \in \mathbb{N}$.

(ii) $\zeta^\theta(s \oplus t) = \zeta^\theta(s) + \zeta^\theta(t)$.

(iii) $\zeta^\theta(s \otimes t) \leq \zeta^\theta(s)\zeta^\theta(t)$.

(iv) *If* $s \geqslant t$, *then* $\zeta^\theta(s) \geq \zeta^\theta(t)$.

Theorem 4.14 ([Str91]). *Let* $s \in \mathbb{F}^{n_1} \otimes \mathbb{F}^{n_2} \otimes \mathbb{F}^{n_3}$ *and* $t \in \mathbb{F}^{m_1} \otimes \mathbb{F}^{m_2} \otimes \mathbb{F}^{m_3}$. *Let* $\theta \in \Theta$.

(i) $\zeta_\theta(\langle n\rangle) = n$ *for* $n \in \mathbb{N}$.

(ii) $\zeta_\theta(s \oplus t) \geq \zeta_\theta(s) + \zeta_\theta(t)$.

(iii) $\zeta_\theta(s \otimes t) \geq \zeta_\theta(s)\zeta_\theta(t)$.

(iv) *If* $s \geqslant t$, *then* $\zeta_\theta(s) \geq \zeta_\theta(t)$.

Theorem 4.15 ([Str91]). $\zeta^\theta(s \otimes t) \geq \zeta^\theta(s)\zeta_\theta(t)$ *and* $\zeta^\theta(t) \geq \zeta_\theta(t)$ *for* $\theta \in \Theta$.

Regarding statement (ii) in Theorem 4.14, Bürgisser [Bür90] shows that the lower support functional ζ_θ 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 $\zeta^\theta(t)$ and the lower support functional $\zeta_\theta(t)$ 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 when $\mathbb{F}$ is algebraically closed the generic value of ζ_θ on $\mathbb{F}^n \otimes \mathbb{F}^n \otimes \mathbb{F}^n$ equals $2^{(1-\min_i \theta_i)n+o(1)}$. On the other hand, Tobler [Tob91] shows that the generic value of ζ^θ on $\mathbb{F}^n \otimes \mathbb{F}^n \otimes \mathbb{F}^n$ equals n. So even generically ζ^θ and ζ_θ are different on $\mathbb{F}^n \otimes \mathbb{F}^n \otimes \mathbb{F}^n$.

For $\theta \in \Theta$ we say f is *θ-robust* if $\zeta^\theta(t) = \zeta_\theta(t)$. We say t is *robust* if t is θ-robust for all $\theta \in \Theta$. Let us try to understand what robust tensors look like. A tensor t is θ-robust if and only if

$$\zeta^\theta(t) \leq \zeta_\theta(t). \tag{4.11}$$

The set of θ-robust tensors is closed under $\oplus$ and $\otimes$, since

$$\zeta^\theta(s \oplus t) = \zeta^\theta(s) + \zeta^\theta(t) = \zeta_\theta(s) + \zeta_\theta(t) \leq \zeta_\theta(s \oplus t),$$

and

$$\zeta^\theta(s \otimes t) \leq \zeta^\theta(s)\zeta^\theta(t) = \zeta_\theta(s)\zeta_\theta(t) \leq \zeta_\theta(s \otimes t).$$

For $X \subseteq [n_1] \times [n_2] \times [n_3]$ we use the notation $H_\theta(X) := \max_{P\in\mathcal{P}(X)} H_\theta(P)$. Let $t \in \mathbb{F}^{n_1} \otimes \mathbb{F}^{n_2} \otimes \mathbb{F}^{n_3} \setminus 0$. Equation (4.11) means that there are $g, h \in G(t)$ and $P \in \mathcal{P}(\max \operatorname{supp}(h \cdot t))$ such that $H_\theta(\operatorname{supp}(g \cdot t)) \leq H_\theta(P)$. In this case we have $\zeta_\theta(t) = \zeta^\theta(t) = 2^{H_\theta(P)}$. In particular, t is θ-robust if there is a $g \in G(t)$ such that the maximisation $H_\theta(\mathrm{supp}(g \cdot t)$ is attained by a $P \in \mathcal{P}(\max(\mathrm{supp}(g \cdot t)))$. This criterion is automatically satisfied for all θ when $\mathrm{supp}(g \cdot t) = \max(\mathrm{supp}(g \cdot t))$ for some $g \in G(t)$. Suppose t is oblique. Then $\mathrm{supp}(g \cdot t)$ is an antichain for some $g \in G(t)$ and thus $\mathrm{supp}(g \cdot t) = \max \mathrm{supp}(g \cdot t)$. Then t is robust and $\zeta_\theta(t) = \zeta^\theta(t) = 2^{H_\theta(\mathrm{supp}(g \cdot t))}$.

4.6 Asymptotic slice rank

Slice rank is a variation on tensor rank that was introduced by Terence Tao in [Tao16] to study cap sets. We will look at cap sets in Section 5.4. Here we study the relationship between asymptotic slice rank and the support functionals.

Consider the following characterisation of tensor rank. Let a simple tensor be any tensor of the form $v_1 \otimes v_2 \otimes v_3 \in V_1 \otimes V_2 \otimes V_3$ with $v_i \in V_i$ for $i \in [k]$. Then the rank $\mathrm{R}(t)$ of $t \in V_1 \otimes V_2 \otimes V_3$ is the smallest number r such that t can be written as a sum of r simple tensors.

Slice rank is defined similarly, but with simple tensors replaced by slices. For $S \subseteq [k]$, let $V_S \coloneqq \bigotimes_{i \in S} V_i$. For $j \in [k]$, let $\overline{j} \coloneqq \overline{\{j\}}$. A tensor in $V_1 \otimes V_2 \otimes V_3$ is called a *slice* if it is of the form $v \otimes w$ with $v \in V_j$ and $w \in V_{\overline{j}}$ for some $j \in [k]$ (under the natural reordering of the tensor legs). Let $t \in V_1 \otimes V_2 \otimes V_3$. The *slice rank* of t, denoted by $\mathrm{SR}(t)$, is the smallest number r such that t can be written as a sum of r slices. For example, the tensor

$$W \coloneqq e_1 \otimes e_1 \otimes e_2 + e_1 \otimes e_2 \otimes e_1 + e_2 \otimes e_1 \otimes e_1 \in \mathbb{F}^2 \otimes \mathbb{F}^2 \otimes \mathbb{F}^2 \tag{4.12}$$

has slice rank 2 since we can write $W = e_1 \otimes (e_1 \otimes e_2 + e_2 \otimes e_1) + e_2 \otimes e_1 \otimes e_1$. In fact, the slice rank of any element in $V_1 \otimes V_2 \otimes V_3$ is at most $\min_i \dim V_i$. The tensor rank of W, on the other hand, is known to be 3.

Slice rank is clearly monotone under restriction. The slice rank of the diagonal tensor $\langle r \rangle$ equals r [Tao16]. It follows that subrank is at most slice rank,

$$\mathrm{Q}(t) \leq \mathrm{SR}(t).$$

The motivation for the introduction of slice rank in [Tao16] was finding upper bounds on subrank $\mathrm{Q}(t)$ and asymptotic subrank $\underset{\sim}{\mathrm{Q}}(t)$.

The main result of this section is the following theorem. Recall that a tensor t is oblique if the support $\mathrm{supp}(g \cdot t)$ is an antichain for some $g \in G(t)$.

Theorem 4.16. *Let t be oblique. Then*

$$\lim_{n \to \infty} \mathrm{SR}(t^{\otimes n})^{1/n} = \min_{\theta \in \mathcal{P}([3])} \zeta^\theta(t).$$

Our proof of Theorem 4.16 is based on a proof of Tao and Sawin in [TS16] and discussions of the author with Dion Gijswijt. The explicit connection between asymptotic slice rank and the support functionals is new.

We use Theorem 4.16, before giving its proof, to see that SR is not submultiplicative and not supermultiplicative under the tensor product $\otimes$. In particular we cannot use Fekete's lemma Lemma 2.2 to prove that the limit $\lim_{n\to\infty} \mathrm{SR}(t^{\otimes n})^{1/n}$ exists. Thus the existence of the limit is a non-trivial consequence of Theorem 4.16.

Let W as in (4.12). Then $\mathrm{SR}(W) = 2$. We have $\zeta^{(1/3,1/3,1/3)}(W) = 2^{h(1/3)} < 2$. From Theorem 4.16 follows $\mathrm{SR}(W^{\otimes n}) \leq 2^{nh(1/3)+o(1)}$. We conclude $\mathrm{SR}(W^{\otimes n}) < 2^n$ for n large enough. We conclude SR is not supermultiplicative. Now it is also clear that slice rank is not the same as (border) subrank, since (border) subrank is supermultiplicative.

Next, 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 $\langle n, n, n\rangle$ which has slice rank n^2 by Theorem 4.16 and Theorem 5.3 in the next chapter applied to the tight tensor $\langle n, n, n\rangle$. We conclude SR is not submultiplicative.

Slice rank and hitting set number

We study the hitting set number of the support of a tensor. Let $\Phi \subseteq [n_1]\times[n_2]\times[n_3]$. A *hitting set* for Φ is a 3-tuple of sets $A_1 \subseteq [n_1], A_3 \subseteq [n_2], A_3 \subseteq [n_3]$ such that for every $a \in \Phi$ there is an $i \in [3]$ with $a_i \in A_i$. We may think of Φ as a 3-partite 3-uniform hypergraph. Then the definition of hitting set says: every edge $a \in \Phi$ is hit by an element of some A_i. A hitting set is also called a vertex cover, every edge being covered by some vertex, or a transversal. The size of the hitting set (A_1, A_2, A_3) is $|A_1| + |A_2| + |A_3|$. The *hitting set number* $\tau(\Phi)$ is the size of the smallest hitting set for Φ. Let $t \in \mathbb{F}^{n_1} \otimes \mathbb{F}^{n_2} \otimes \mathbb{F}^{n_3}$.

Lemma 4.17. *Let $g \in G(t) \coloneqq \mathrm{GL}_{n_1} \times \mathrm{GL}_{n_2} \times \mathrm{GL}_{n_3}$. Then $\mathrm{SR}(t) \leq \tau(\mathrm{supp}(g\cdot t))$.*

Proof. This is clear. □

Lemma 4.18. *Let $g \in G(t)$. Then $\mathrm{SR}(t) \geq \tau(\max(\mathrm{supp}(g \cdot t)))$.*

Proof. It is sufficient to consider $g = e$. Let

$$t = \sum_{i=1}^{r_1} v_i^1 \otimes u_i^{\overline{1}} + \sum_{i=1}^{r_2} v_i^2 \otimes u_i^{\overline{2}} + \sum_{i=1}^{r_3} v_i^3 \otimes u_i^{\overline{3}}$$

be a slice decomposition. We may assume $v_1^j, \ldots, v_{r_j}^j$ are linearly independent. Let $V_j = \mathrm{Span}\{v_1^j, \ldots, v_{r_j}^j\} \subseteq \mathbb{F}^{n_j}$. Let $W_j \subseteq (\mathbb{F}^{n_j})^*$ be the elements in the dual space that vanish on V_j. Let $B_j \subseteq W_j$ be a basis with the following property: with respect to the standard basis, the matrix with the elements of B_j as columns is in reduced row echelon form, i.e. each column is of the form $(* \cdots * \, 1\, 0 \cdots 0)^T$ and the pivot elements (the 1's) are all in different rows. Let $S_j \subseteq [n_j]$ be the indices of the pivot element. Let $\overline{S_j} \coloneqq [n_j] \setminus S_j$ be the complement. Then $|\overline{S_j}| = r_j$. We claim $(\overline{S_1}, \overline{S_2}, \overline{S_3})$ is a hitting set for $\max(\mathrm{supp}(t))$.

Then $r_1 + r_2 + r_3 = |\overline{S_1}| + |\overline{S_2}| + |\overline{S_3}| \geq \tau(\max(\mathrm{supp}(t)))$. Let $x \in \max(\mathrm{supp}(t))$. Suppose $x \in S_1 \times S_2 \times S_3$. For every $j \in [3]$ let $\phi_j \in B_j$ have its pivot element at index x_j. Let $\phi = \phi_1 \otimes \phi_2 \otimes \phi_3$. Then $\phi \in W_1 \otimes W_2 \otimes W_3$, so $\phi(t) = 0$. Since x is maximal and each B_j is in reduced row echelon form,

$$\begin{aligned}\phi(t) &= \sum_{y \leq x} t_y\, \phi(e_{y_1} \otimes e_{y_2} \otimes e_{y_3})\\ &= \sum_{y < x} t_y\, \phi(e_{y_1} \otimes e_{y_2} \otimes e_{y_3}) \;+\; t_x\, e_{x_1} \otimes e_{x_2} \otimes e_{x_3}\\ &= \sum_{y < x} s_y\, e_{y_1} \otimes e_{y_2} \otimes e_{y_3} \;+\; t_x\, e_{x_1} \otimes e_{x_2} \otimes e_{x_3}\end{aligned}$$

for some $s_y \in \mathbb{F}$. From $\phi(t) = 0$ follows $t_x = 0$. This contradicts $x \in \mathrm{supp}(t)$, so $x \notin S_1 \times S_2 \times S_3$, i.e. there is a $j \in [3]$ with $x_j \in \overline{S_j}$. □

Asymptotic hitting set number

We now study the asymptotic hitting set number $\underset{\sim}{\tau}(\Phi) := \lim_{n\to\infty} \tau(\Phi^{\times n})^{1/n}$.

We will use some basic facts of types and type classes. Let X be a finite set. Let $N \in \mathbb{N}$. An *N-type on X* is a probability distribution P on X with $N \cdot P(x) \in \mathbb{N}$ for all $x \in X$. Let P be an N-type on X. The type class $T_P^N \subseteq X^N$ is the set of sequences $s = (s_1, \ldots, s_N)$ with x occuring $N \cdot P(x)$ times in s for every $x \in X$, i.e. $|\{i \in [N] : s_i = x\}| = N \cdot P(x)$.

Lemma 4.19. *The number of N-types on X equals $\binom{N+|X|-1}{|X|-1}$. Let P be an N-type. The size of the type class T_P^N equals the multinomial coefficient $\binom{N}{NP}$.*

Proof. We leave the proof to the reader. □

Lemma 4.20. *Let P be an N-type on X. Then*

$$\frac{1}{(N+1)^{|X|}} 2^{NH(P)} \leq \binom{N}{NP} \leq 2^{NH(P)}.$$

Proof. See e.g. [CT12, Theorem 11.1.3]. □

Lemma 4.21. $\log_2 \underset{\sim}{\tau}(\Phi) \leq \max_{P \in \mathcal{P}(\Phi)} \min_{i \in [3]} H(P_i)$.

Proof. Let P maximise $\max_{P \in \mathcal{P}(\Phi)} \min_i H(P_i)$. Let $n \in \mathbb{N}$. We construct a hitting set (A_1, A_2, A_3) for Φ^n as follows. Let $x \in \Phi^n$. Viewing x as an n-tuple of elements in Φ, let $Q \in \mathcal{P}_n(\Phi)$ be the type of x (i.e. the empirical distribution). Let $j \in [3]$ with $H(Q_j) = \min_{i \in [3]} H(Q_i)$. By our choice of P we have

$$H(Q_j) = \min_{i \in [3]} H(Q_i) \leq \min_{i \in [3]} H(P_i).$$

Viewing x as a 3-tuple (x_1, x_2, x_3), add x_j to A_j. We repeat this for all $x \in \Phi^n$. The final (A_1, A_2, A_3) is a hitting set for Φ^n by construction. For each $j \in [3]$,

$$|A_j| \le \sum_{Q_j} |T^n_{Q_j}| \le \sum_{Q_j} 2^{nH(Q_j)}$$

where the sum is over $Q_j \in \mathcal{P}_n(\Phi_j)$ with $H(Q_j) \le \min_{i\in[3]} H(P_i)$. Then

$$|A_j| \le |\mathcal{P}_n(\Phi_j)|\, 2^{n \min_i H(P_i)} = \mathrm{poly}(n) 2^{n \min_i H(P_i)}.$$

We conclude $|A_1| + |A_2| + |A_3| \le \mathrm{poly}(n) 2^{n \min_i H(P_i)}$. □

Lemma 4.22. $\log_2 \underset{\sim}{T}(\Phi) \ge \max_{P\in\mathcal{P}(\Phi)} \min_{i\in[3]} H(P_i)$.

Proof. Let P maximise $\max_{P\in\mathcal{P}(\Phi)} \min_i H(P_i)$. Let $n \in \mathbb{N}$. Let (A_1, A_2, A_3) be a hitting set for Φ^n. Let $Q \in \mathcal{P}_n(\Phi)$ be an n-type with $\min_i H(Q_i) = \min_i H(P_i) - o(n)$. Let $\Psi = T^n_Q \subseteq \Phi^n$ be the set of strings with type Q. Then (A_1, A_2, A_3) is a hitting set for Ψ. Let $\pi_i : \Psi \to \Phi^n_i : (x_1, x_2, x_3) \mapsto x_i$. Then

$$\Psi = \pi_1^{-1}(A_1) \cup \pi_2^{-1}(A_2) \cup \pi_3^{-1}(A_3).$$

Let $j \in [3]$ with $|\pi_j^{-1}(A_j)| \ge \frac{1}{3}|\Psi|$. The fiber $\pi_j^{-1}(a)$ has constant size over $a \in \Psi_j$. Let $c_j = |\pi_j^{-1}(a)|$ be this size. Then

$$|\Psi| = \sum_{a\in\Psi_j} |\pi_j^{-1}(a)| = \sum_{a\in\Psi_j} c_j = |\Psi_j|\, c_j.$$

And

$$|\pi_j^{-1}(A_j)| = \sum_{a\in A_j\cap\Psi_j} |\pi_j^{-1}(a)| = |A_j \cap \Psi_j|\, c_j \le |A_j|\, c_j.$$

Therefore

$$|A_j| \ge \frac{|\pi_j^{-1}(A_j)|}{c_j} \ge \frac{\frac{1}{3}|\Psi|}{c_j} = \tfrac{1}{3}|\Psi_j|\,.$$

We have $|\Psi_j| \ge 2^{nH(Q_j)-o(n)} \ge 2^{n\min_i H(Q_i)-o(n)} \ge 2^{n\min_i H(P_i)-o(n)}$. We conclude $|A_1| + |A_2| + |A_3| \ge |A_j| \ge \frac{1}{3}|\Psi_j| \ge \frac{1}{3} 2^{n\min_i H(P_i)-o(n)}$. □

Lemma 4.23. $\log_2 \underset{\sim}{T}(\Phi) = \max_{P\in\mathcal{P}(\Phi)} \min_{i\in[3]} H(P_i)$.

Proof. This follows directly from the above lemmas. □

Asymptotic slice rank

We now combine the above lemmas about slice rank and the asymptotic hitting set number to prove Theorem 4.16. First we have the following basic lemma.

Lemma 4.24. $\min_{\theta\in\Theta}\max_{P\in\mathcal{P}(\Phi)} H_\theta(P) = \max_{P\in\mathcal{P}(\Phi)}\min_{i\in[3]} H(P_i)$.

Proof. Since $H_\theta(P)$ is convex in θ and concave in P, von Neumann's minimax theorem gives $\min_\theta \max_{P\in\mathcal{P}(\Phi)} H_\theta(P) = \max_{P\in\mathcal{P}(\Phi)}\min_\theta H_\theta(P)$. Finally, we use that $\min_\theta H_\theta(P) = \min_i H(P_i)$. □

Define $f^{\sim}(t) \coloneqq \limsup_{n\to\infty} f(t^{\otimes n})^{1/n}$ and $f_{\sim}(t) \coloneqq \liminf_{n\to\infty} f(t^{\otimes n})^{1/n}$.

Lemma 4.25. *Let $t \in \mathbb{F}^{n_1}\otimes\mathbb{F}^{n_2}\otimes\mathbb{F}^{n_3}$. Then*

$$\max_{g\in G(t)}\max_{P\in\mathcal{P}(\max\operatorname{supp}(g\cdot t))}\min_i H(P_i) \leq \mathrm{SR}_{\sim}(t) \leq \mathrm{SR}^{\sim}(t) \leq \min_\theta \zeta^\theta(t).$$

Proof. By definition, $\mathrm{SR}_{\sim}(t) \leq \mathrm{SR}^{\sim}(t)$. From Lemma 4.17 follows

$$\mathrm{SR}^{\sim}(t) \leq \underset{\sim}{\tau}(\operatorname{supp}(g\cdot t))$$

for any $g \in G(t)$. Lemma 4.23 gives $\underset{\sim}{\tau}(\operatorname{supp}(g\cdot t)) = \max_{P\in\mathcal{P}(\operatorname{supp}(g\cdot t))}\min_i 2^{H(P_i)}$. Thus with the help of Lemma 4.24

$$\mathrm{SR}^{\sim}(t) \leq \min_{g\in G(t)}\max_{P\in\mathcal{P}(\operatorname{supp}(g\cdot t))}\min_i 2^{H(P_i)} = \min_\theta \zeta^\theta(t).$$

From Lemma 4.18 follows

$$\underset{\sim}{\tau}(\max(\operatorname{supp}(g\cdot t))) \leq \mathrm{SR}_{\sim}(t)$$

for any $g \in G(t)$. Lemma 4.23 gives

$$\max_{g\in G(t)}\max_{P\in\mathcal{P}(\max(\operatorname{supp}(g\cdot t)))}\min_i 2^{H(P_i)} \leq \mathrm{SR}_{\sim}(t).$$

This proves the lemma. □

Proof of Theorem 4.16. We may assume $\Phi = \operatorname{supp}(t)$ is oblique. Then, with the help of Lemma 4.24 and Lemma 4.25

$$\begin{aligned}
\min_{\theta\in\Theta}\zeta^\theta(t) &= \min_{\theta\in\Theta}\zeta_\theta(t)\\
&= \min_{\theta\in\Theta}\max_{P\in\max(\Phi)} 2^{H_\theta(P)}\\
&= \max_{P\in\max(\Phi)}\min_{i\in[3]} 2^{H(P_i)}\\
&\leq \max_{g\in G(t)}\max_{P\in\mathcal{P}(\max(\operatorname{supp}(g\cdot t)))}\min_{i\in[3]} 2^{H(P_i)}\\
&\leq \mathrm{SR}_{\sim}(t)\\
&\leq \mathrm{SR}^{\sim}(t)\\
&\leq \min_{\theta\in\Theta}\zeta^\theta(t).
\end{aligned}$$

This proves the claim. □

4.7 Conclusion

The study of asymptotic rank of tensors is motivated by the open problem of finding the exponent of matrix multiplication. Asymptotic subrank has applications in for example combinatorics and algebraic property testing. Via the theory of asymptotic spectra Strassen characterised asymptotic rank and asymptotic subrank in terms of the asymptotic spectrum of tensors. Strassen introduced the gauge points in $\mathbf{X}(\mathcal{T})$ and the support functionals in $\mathbf{X}(\{\text{oblique}\})$. More precisely, there are the lower support functionals, and the upper support functionals. 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. Finally, we showed that for oblique tensors the asymptotic slice rank exists and equals the minimum value over the support functionals. In the next chapter we will see a subfamily of the oblique 3-tensors for which the support functionals are powerful enough to compute the asymptotic subrank.

Chapter 5

Tight tensors and combinatorial subrank; cap sets

This chapter is based on joint work with Matthias Christandl and Péter Vrana [CVZ16, CVZ18].

5.1 Introduction

In the previous chapter we discussed the gauge points and the support functionals ζ^θ. The gauge points are in the asymptotic spectrum of all tensors, while the support functionals are in the asymptotic spectrum of oblique tensors.

How "powerful" are the support functionals? We know $\underset{\sim}{\mathrm{Q}}(t) \leq \zeta^\theta(t) \leq \underset{\sim}{\mathrm{R}}(t)$ for oblique t and for all θ. Thus $\max_\theta \zeta^\theta(t) \leq \underset{\sim}{\mathrm{R}}(t)$. In fact, $\max_\theta \zeta^\theta(t)$ is at most the maximum over the gauge points $\max_S \zeta_{(S)}$, and in turn $\max_S \zeta_{(S)}$ is at most $\underset{\sim}{\mathrm{R}}(t)$. As remarked earlier, it is not known whether $\max_S \zeta_{(S)}$ equals $\underset{\sim}{\mathrm{R}}(t)$ in general.

On the other hand, we have $\underset{\sim}{\mathrm{Q}}(t) \leq \min_\theta \zeta^\theta(t)$. Do we attain equality here in general; $\underset{\sim}{\mathrm{Q}}(t) = \min_\theta \zeta^\theta(t)$? The answer is "yes" for the subsemiring of tight 3-tensors. In this chapter we study tight k-tensors.

Tight tensors

Let $I_1, \ldots, I_k$ be finite sets. Let $\Phi \subseteq I_1 \times \cdots \times I_k$. We say Φ is *tight* if there are injective maps $u_i : I_i \to \mathbb{Z}$ for $i \in [k]$ such that

$$\forall \alpha \in \Phi \quad u_1(\alpha_1) + \cdots + u_k(\alpha_k) = 0.$$

We say $t \in \mathbb{F}^{n_1} \otimes \cdots \otimes \mathbb{F}^{n_k}$ is *tight* if there is a $g \in G(t) := \mathrm{GL}_{n_1} \times \cdots \times \mathrm{GL}_{n_k}$ such that the support $\mathrm{supp}(g \cdot t)$ is tight.

Recall that a tensor is oblique if the support is an antichain in some basis. 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 θ-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 5.1. Let $k \geq 3$ be fixed. For any integer $n \geq 1$ and $c \in [n]$ the set

$$\Phi_n(c) = \{\alpha \in \{0, \dots, n-1\}^k : \alpha_1 + \cdots + \alpha_k = c\}$$

is tight. For any integer $n \geq 2$ and any $c \in [n]$ the set

$$\Psi_n(c) = \{\alpha \in \{0, \dots, n-1\}^k : \alpha_1 + \cdots + \alpha_k = c \bmod n\}$$

is not tight (cf. Exercise 15.20 in [BCS97]).

Example 5.2. When $\mathbb{F}$ contains a primitive nth root of unity ζ, the tensor

$$t_n = \sum_{\alpha \in \Psi_n(0)} e_{\alpha_1} \otimes \cdots \otimes e_{\alpha_k} \in (\mathbb{F}^n)^{\otimes k},$$

which has support $\Psi_n(0)$, is tight. Namely, the elements $v_j = \sum_{i=1}^n \zeta^{ij} e_i$ for $j \in [n]$ form a basis of $\mathbb{F}^n$. Let $g \in G(t_n)$ be the corresponding basis transformation. Then we have $t_n = \sum_{j=1}^n v_j \otimes \cdots \otimes v_j$, and we see that the support $\mathrm{supp}(g^{-1} \cdot t_n) = \{\alpha \in [n]^k : \alpha_1 = \cdots = \alpha_k\}$ is tight. (See also [BCS97, Exercise 15.25].) When the characteristic of $\mathbb{F}$ equals n, the tensor t_n is also tight, as we will see in Section 5.4.2.

Combinatorial subrank and the Coppersmith–Winograd method

We care about tight tensors because of a remarkable theorem for tight 3-tensors of Strassen (Theorem 5.3 below). To understand the theorem we need the concept of combinatorial asymptotic subrank (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* Φ or simply $D \subseteq \Phi$ is a *free diagonal* if $D = \Phi \cap (D_1 \times \cdots \times D_k)$, where $D_i = \{x_i : (x_1, \dots, x_k) \in D\}$. Define the *(combinatorial) subrank* $\mathrm{Q}(\Phi)$ 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 $\Phi \times \Psi \subseteq (I_1 \times J_1) \times \cdots \times (I_k \times J_k)$ by

$$\Phi \times \Psi = \{((\alpha_1, \beta_1), \dots, (\alpha_k, \beta_k)) : \alpha \in \Phi, \beta \in \Psi\}.$$

Define the *(combinatorial) asymptotic subrank* $\underset{\sim}{\mathrm{Q}}(\Phi) = \lim_{n\to\infty} \mathrm{Q}(\Phi^{\times n})^{1/n}$. Let $t \in \mathbb{F}^{n_1} \otimes \cdots \otimes \mathbb{F}^{n_k}$ and let Φ be the support of t in the standard basis. Then $\mathrm{Q}(\Phi) \leq \mathrm{Q}(t)$ and $\underset{\sim}{\mathrm{Q}}(\Phi) \leq \underset{\sim}{\mathrm{Q}}(t)$. The number $\mathrm{Q}(\Phi)$ may be interpreted as the largest number n such that $\langle n \rangle$ can be obtained from t using a restriction that consists of matrices that have at most one nonzero entry in each row and in each column. (This is called M-restriction in [Str87, Section 6] which stands for monomial restriction.) We may also interpret Φ as a k-partite hypergraph. Then $\mathrm{Q}(\Phi)$ is the size of the largest induced k-partite matching in Φ.

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 Φ. Then the (asymptotic) subranks of Φ and t are related as follows:

$$\mathrm{Q}(\Phi) \leq \mathrm{Q}(t), \text{ and } \underset{\sim}{\mathrm{Q}}(\Phi) \leq \underset{\sim}{\mathrm{Q}}(t).$$

Strassen proved the following theorem using the method of Coppersmith and Winograd [CW90]. Recall that for $\Phi \subseteq I_1 \times I_2 \times I_3$ we let $\mathcal{P}(\Phi)$ be the set of probability distributions on Φ. For $P \in \mathcal{P}(\Phi)$, let P_1, P_2, P_3 be the marginal distributions of P on the 3 components of $I_1 \times I_2 \times I_3$.

Theorem 5.3 ([Str91, Lemma 5.1])**.** *Let $\Phi \subseteq I_1 \times I_2 \times I_3$ be tight. Then*

$$\underset{\sim}{\mathrm{Q}}(\Phi) = \max_{P \in \mathcal{P}(\Phi)} \min_{i \in [3]} 2^{H(P_i)}. \tag{5.1}$$

The consequence of Theorem 5.3 is that the support functionals are sufficiently powerful to compute the asymptotic subrank of tight 3-tensors.

Corollary 5.4 ([Str91, Proposition 5.4])**.** *Let $t \in \mathbb{F}^{n_1} \otimes \mathbb{F}^{n_2} \otimes \mathbb{F}^{n_3}$ be tight. Then*

$$\underset{\sim}{\mathrm{Q}}(t) = \min_{\theta \in \mathcal{P}([3])} \zeta^{\theta}(t).$$

Moreover, if $\Phi = \mathrm{supp}(g \cdot t)$ is tight for some $g \in G(t)$, then $\underset{\sim}{\mathrm{Q}}(t) = \underset{\sim}{\mathrm{Q}}(\Phi)$.

Remark 5.5. Strassen conjectured in [Str94, Conjecture 5.3] that for the family of tight 3-tensors the support functionals give all spectral points in the asymptotic spectrum $\mathbf{X}(\{\text{tight 3-tensors}\})$. In [Str91] numerous examples are given of subfamilies of tight 3-tensors for which this is the case.

Remark 5.6. Equation (5.1) becomes false when 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].

New results in this chapter

This chapter is an investigation of tight tensors, combinatorial asymptotic subrank and applications. More precisely this chapter contains the following new results.

Higher-order Coppersmith–Winograd method. In Section 5.2 we extend Theorem 5.3 to obtain a lower bound for $\underset{\sim}{\mathrm{Q}}(\Phi)$ for tight sets $\Phi \subseteq I_1 \times \cdots \times I_k$ with $k \geq 4$. Our lower bound is not known to be optimal in general. We compute examples for which the lower bound is optimal.

Combinatorial degeneration method. In Section 5.3 we further extend the range of application of the Coppersmith–Winograd method via a partial order $\trianglelefteq$ on supports of tensors called combinatorial degeneration. We prove that if $\Phi \trianglelefteq \Psi$, then $\underset{\sim}{\mathrm{Q}}(\Phi) \leq \underset{\sim}{\mathrm{Q}}(\Psi)$. Suppose Ψ is not tight, but Φ is tight, then we may apply the (higher-order) Coppersmith–Winograd method to obtain a lower bound on $\underset{\sim}{\mathrm{Q}}(\Phi)$ and thus on $\underset{\sim}{\mathrm{Q}}(\Psi)$.

Cap sets. In Section 5.4 we relate the theory of asymptotic spectra, the Coppersmith–Winograd method and the combinatorial degeneration method to the problem of upper bounding the maximum size of cap sets in $\mathbb{F}_p^n$.

Graph tensors. Graph tensors are generalisations of the matrix multiplication tensor $\langle 2,2,2 \rangle$ parametrised by graphs. In Section 5.5 we discuss how one can apply the higher-order Coppersmith–Winograd method to obtain upper bounds on the asymptotic rank of complete graph tensors. We also briefly discuss the surgery method, which gives good upper bounds on the asymptotic rank of graph tensors for sparse graphs like cycle graphs.

5.2 Higher-order CW method

In this section we extend Theorem 5.3 to tight $\Phi \subseteq I_1 \times \cdots \times I_k$ with $k \geq 4$. We introduce some notation. Let $\mathcal{P}(\Phi)$ be the set of probability distributions on Φ. For $P \in \mathcal{P}(\Phi)$, let $P_1, \ldots, P_k$ be the marginal distributions of P on the k components of $I_1 \times \cdots \times I_k$. Let $\mathcal{R}(\Phi)$ be the set of all subsets $R \subseteq \Phi^2$ such that: $R \not\subseteq \{(x,x) : x \in \Phi\}$ and $R \subseteq \{(x,y) \in \Phi^2 : x_i = y_i\}$ for some $i \in [k]$. For $P \in \mathcal{P}(\Phi)$ and $R \in \mathcal{R}(\Phi)$, let $\mathcal{Q}(R, (P_1, \ldots, P_k))$ be the set of probability distributions Q on R whose marginal distributions on the $2k$ components of R satisfy $Q_i = Q_{k+i} = P_i$ for $i \in [k]$.

Let $I_1, \ldots, I_k$ be finite subsets of $\mathbb{Z}$. The result of this section is a lower bound on the asymptotic subrank of any $\Phi \subseteq I_1 \times \cdots \times I_k$ satisfying $\forall a \in \Phi\ \sum_{i=1}^k a_i = 0$. For $R \subseteq \mathcal{R}(\Phi)$, let $r(R)$ be the rank over $\mathbb{Q}$ of the matrix with rows $\{x - y : (x,y) \in R\}$.

Theorem 5.7. *Let $\Phi \subseteq \mathbb{Z}^k$ be a finite set with $\forall a \in \Phi\ \sum_{i=1}^k a_i = 0$. Then*

$$\log_2 \underset{\sim}{\mathrm{Q}}(\Phi) \geq \max_P \min_{R,Q} H(P) - (k-2)\frac{H(Q) - H(P)}{r(R)}$$

with $P \in \mathcal{P}(\Phi)$, $R \in \mathcal{R}(\Phi)$ and $Q \in \mathcal{Q}(R, (P_1, \ldots, P_k))$.

5.2.1 Construction

We prepare for the proof of Theorem 5.7 by discussing some basic facts.

Average-free sets

Lemma 5.8. *Let $k \in \mathbb{N}$. Let $M \in \mathbb{N}$. We say a subset $B \subseteq \mathbb{Z}/M\mathbb{Z}$ is $(k-1)$-average-free if*

$$\forall x_1, \ldots, x_k \in B \quad x_1 + \cdots + x_{k-1} = (k-1)x_k \;\Rightarrow\; x_1 = \cdots = x_k.$$

There is a $(k-1)$-average-free set $B \subseteq \mathbb{Z}/M\mathbb{Z}$ of size $|B| = M^{1-o(1)}$.

Proof. There is a set $A \subseteq \{1, \ldots, \lfloor \frac{M-1}{k-1} \rfloor\}$ of size $|A| = M^{1-o(1)}$ with

$$\forall x_1, \ldots, x_k \in A \quad x_1 + \cdots + x_{k-1} = (k-1)x_k \;\Rightarrow\; x_1 = \cdots = x_k, \tag{5.2}$$

see [VC15, Lemma 10]. Let $B = \{a \bmod M : a \in A\} \subseteq \mathbb{Z}/M\mathbb{Z}$. Then $|B| = |A|$. Let $x_1, \ldots, x_k \in B$ with $x_1 + \cdots + x_{k-1} = (k-1)x_k$. View $x_1, \ldots, x_k$ as elements in $\{1, \ldots, \lfloor \frac{M-1}{k-1} \rfloor\}$. Then $x_1 + \cdots + x_{k-1} = (k-1)x_k$ still holds. From (5.2) follows $x_1 = \cdots = x_k$ in $\mathbb{Z}$, and hence also in $\mathbb{Z}/M\mathbb{Z}$. □

Linear combinations of uniform variables

Lemma 5.9. *Let M be a prime. Let $u_1, \ldots, u_n$ be independently uniformly distributed over $\mathbb{Z}/M\mathbb{Z}$. Let $v_1, \ldots, v_m$ be $(\mathbb{Z}/M\mathbb{Z})$-linear combinations of $u_1, \ldots, u_n$. Then the vector $v = (v_1, \ldots, v_m)$ is uniformly distributed over the range of v in $(\mathbb{Z}/M\mathbb{Z})^m$.*

Proof. Let $v_i = \sum_j c_{ij} u_j$ with $c_{ij} \in \mathbb{Z}/M\mathbb{Z}$. Then $v = Cu$ with $u = (u_1, \ldots, u_n)$ and C the matrix with entries $C_{ij} = c_{ij}$. Let y in the image of C. Then the cardinality of the preimage $C^{-1}(y)$ equals the cardinality of the kernel of C. Indeed, if $Cx = y$, then $C^{-1}(y) = x + \ker(C)$. Since u is uniform, we conclude that v is uniform on the image of C. □

Free diagonals

Lemma 5.10. *Let G be a graph with n vertices and m edges. Then G has at least $n - m$ connected components.*

Proof. A graph without edges has n connected components. For every edge that we add to the graph, we lose at most one connected component. □

Lemma 5.11. *Let $I_1, \ldots, I_k$ be finite sets. Let $\Psi \subseteq I_1 \times \cdots \times I_k$. Let*

$$C = \{\{a, b\} \subseteq \Psi : a \neq b; \exists i \in [k]\, a_i = b_i\}.$$

Then $\mathrm{Q}(\Psi) \geq |\Psi| - |C|$. Obviously, the statement remains true if we replace C by the larger set $\{(a, b) \in \Psi^2 : a \neq b; \exists i \in [k]\, a_i = b_i\}$.

Proof. Let $G = (\Psi, C)$ be the graph with vertex set Ψ and edge set C. Let $\Gamma \subseteq \Psi$ contain exactly one vertex per connected component of G. The vertices in Γ are pairwise not adjacent. So Γ is a diagonal. Of course, $\Gamma \subseteq \Psi \cap (\Gamma_1 \times \cdots \times \Gamma_k)$. Let $a \in \Psi \cap (\Gamma_1 \times \cdots \times \Gamma_k)$. Let $x_1, \ldots, x_k \in \Gamma$ with

$$(x_1)_1 = a_1,\ (x_2)_2 = a_2,\ \ldots,\ (x_k)_k = a_k.$$

Then $x_1, \ldots, x_k$ are all adjacent to a in G, i.e. they are all in the same connected component. Then $x_1 = \cdots = x_k$, since Γ contains precisely one vertex per connected component. So $a = x_1 = \cdots = x_k$. So $a \in \Gamma$. We conclude that $\Gamma \supseteq \Psi \cap (\Gamma_1 \times \cdots \times \Gamma_k)$. Finally, $|\Gamma| \geq |\Psi| - |C|$ by Lemma 5.10. □

We now give the proof of Theorem 5.7. We repeat some notation from above. Let $k \geq 3$. Let $\Phi \subseteq \mathbb{Z}^k$ be a finite set. Let $\mathcal{P}(\Phi)$ be the set of probability distributions on Φ. For $P \in \mathcal{P}(\Phi)$, let $P_1, \ldots, P_k$ be the marginal distributions of P on the k components of $\mathbb{Z}^k$. Let $\mathcal{R}(\Phi)$ be the set of all subsets $R \subseteq \Phi^2$ such that: $R \not\subseteq \{(x,x) : x \in \Phi\}$ and $R \subseteq \{(x,y) \in \Phi^2 : x_i = y_i\}$ for some $i \in [k]$. For $P \in \mathcal{P}(\Phi)$ and $R \in \mathcal{R}(\Phi)$, let $\mathcal{Q}(R, (P_1, \ldots, P_k))$ be the set of probability distributions Q on R whose marginal distributions on the $2k$ components of R satisfy $Q_i = Q_{k+i} = P_i$ for $i \in [k]$. For $R \subseteq \mathcal{R}(\Phi)$, let $r(R)$ be the rank over $\mathbb{Q}$ of the matrix with rows

$$\{x - y : (x,y) \in R\}.$$

For any prime M, let $r_M(R)$ be the rank over $\mathbb{Z}/M\mathbb{Z}$ of the same matrix.

Theorem (Theorem 5.7)**.** *Let $\Phi \subseteq \mathbb{Z}^k$ be a nonempty finite set such that for all $a \in \Phi$ holds $\sum_{i=1}^k a_i = 0$. Then*

$$\log_2 \underset{\sim}{\mathrm{Q}}(\Phi) \geq \max_P \min_{R,Q} H(P) - (k-2)\frac{H(Q) - H(P)}{r(R)}$$

with $P \in \mathcal{P}(\Phi)$, $R \in \mathcal{R}(\Phi)$ and $Q \in \mathcal{Q}(R, (P_1, \ldots, P_k))$.

Proof. Let P be a rational probability distribution on Φ, i.e. $\forall a \in \Phi\ P(a) \in \mathbb{Q}$.

Choice of parameters

This proof involves a variable N that we will let go to infinity, and a prime number M that depends on N. For the sake of rigor we first set the dependence of M on N, and make sure that N is large enough for M to have good properties.

Let $n \in \mathbb{N}$ such that P is an n-type, i.e. $\forall a \in \Phi\ nP(a) \in \mathbb{N}$. Let $N = tn$ be a multiple of n. Let

$$f(N) = \log_2\Bigl(2^{|\Phi|^2} \max_{R \in \mathcal{R}(\Phi)} \binom{N + |R| - 1}{|R| - 1}\Bigr) \in o(N). \tag{5.3}$$

Let

$$g(N) = |\Phi| \log_2(N+1) \in o(N).$$

By Lemma 4.20

$$2^{NH(P)-g(N)} \leq \binom{N}{NP}. \tag{5.4}$$

Let

$$\mu(N) = \max_{R,Q} \frac{H(Q) - H(P) + (1 + g(N) + f(N))\frac{1}{N}}{r(R)} \tag{5.5}$$

with $R \in \mathcal{R}(\Phi)$ and $Q \in \mathcal{Q}(R, (P_1, \ldots, P_k))$. Let M be a prime with

$$\lceil 2^{\mu(N)N} \rceil \leq M \leq 2\lceil 2^{\mu(N)N} \rceil. \tag{5.6}$$

Such a prime exists by Bertrand's postulate, see e.g. [AZ14]. We can make M arbitrarily large by choosing N large enough. Choose $N = tn$ large enough such that

$$M > k - 1 \tag{5.7}$$

$$\forall R \in \mathcal{R}(\Phi) \quad r_M(R) = r(R). \tag{5.8}$$

We will later let t and thus N go to infinity.

Restrict to marginal type classes

The set $\Phi^{\otimes N}$ is a finite subset of $(\mathbb{Z}^N)^k$. Let $a \in \Phi^{\otimes N}$. Then we have that $a_i = ((a_i)_1, \ldots, (a_i)_N) \in \mathbb{Z}^N$ for $i \in [k]$. We restrict to those a for which a_i is in the type class $T^N_{P_i}$ for all $i \in [k]$. Thus, let

$$\Psi = \Phi^{\otimes N} \cap (T^N_{P_1} \times \cdots \times T^N_{P_k}).$$

We prove a lower bound on the size of Ψ. Let $(s_1, \ldots, s_N) \in T^N_P$. Then $s_j \in \Phi$ for $j \in [N]$ and $((s_1)_i, \ldots, (s_N)_i) \in T^N_{P_i}$ for $i \in [k]$. So

$$\big(((s_1)_1, \ldots, (s_N)_1), \ldots, ((s_1)_k, \ldots, (s_N)_k)\big) \in \Phi^{\otimes N} \cap (T^N_{P_1} \times \cdots \times T^N_{P_k}) = \Psi.$$

Thus $|\Psi| \geq |T^N_P|$. By Lemma 4.19, $|T^N_P| = \binom{N}{NP}$. By Lemma 4.20, $\binom{N}{NP} \geq 2^{NH(P)-g(N)}$. Therefore,

$$|\Psi| \geq 2^{NH(P)-g(N)}. \tag{5.9}$$

Hashing

Let $u_1, \ldots, u_{k-1}, v_1, \ldots, v_N \in \mathbb{Z}/M\mathbb{Z}$. For $i \in [k]$ let

$$\begin{aligned} h_i : \mathbb{Z}^N &\to \mathbb{Z}/M\mathbb{Z} \\ x &\mapsto \begin{cases} u_i + \sum_{j=1}^N x_j v_j & \text{for } 1 \le i \le k-1 \\ \frac{1}{k-1}\big(u_1 + \cdots + u_{k-1} - \sum_{j=1}^N x_j v_j\big) & \text{for } i = k. \end{cases} \end{aligned}$$

Note that $k-1$ is invertible in $\mathbb{Z}/M\mathbb{Z}$ by (5.7). Let $a \in \Psi$. Then $((a_1)_j, \ldots, (a_k)_j) \in \Phi$ for $j \in [N]$. So $\sum_{i=1}^k (a_i)_j = 0$ for every $j \in [N]$. Thus

$$\sum_{i=1}^k \sum_{j=1}^N (a_i)_j v_j = \sum_{j=1}^N v_j \sum_{i=1}^k (a_i)_j = 0.$$

Therefore

$$h_1(a_1) + \cdots + h_{k-1}(a_{k-1}) = (k-1) h_k(a_k).$$

Restrict to average-free set

Let $B \subseteq \mathbb{Z}/M\mathbb{Z}$ be a $(k-1)$-average-free set of size

$$|B| \ge M^{1-\kappa(M)} \quad \text{with} \quad \kappa(M) \in o(1), \tag{5.10}$$

meaning

$$\forall x_1, \ldots, x_k \in B \;\; x_1 + \cdots + x_{k-1} = (k-1)x_k \Rightarrow x_1 = \cdots = x_k \tag{5.11}$$

(Lemma 5.8). Let $\Psi' \subseteq \Psi$ be the subset

$$\Psi' = \{a \in \Psi : \forall i \in [k] \;\; h_i(a_i) \in B\}.$$

Let $a \in \Psi'$. Then $a \in \Psi$, so

$$h_1(a_1) + \cdots + h_{k-1}(a_{k-1}) = (k-1) h_k(a_k)$$

Since $h_i(a_i) \in B$ for every $i \in [k]$, (5.11) implies

$$h_1(a_1) = \cdots = h_k(a_k).$$

Probabilistic method

Clearly $\mathrm{Q}(\Phi^{\otimes N}) \ge \mathrm{Q}(\Psi) \ge \mathrm{Q}(\Psi')$. Let

$$C' = \{(a, b) \in \Psi'^2 : a \ne b; \exists i \in [k] \;\; a_i = b_i\}.$$

Let $X = |\Psi'|$ and $Y = |C'|$. By Lemma 5.11

$$\mathrm{Q}(\Psi') \geq X - Y.$$

Let $u_1, \ldots, u_{k-1}, v_1, \ldots, v_N$ be independent uniformly random variables over the field $\mathbb{Z}/M\mathbb{Z}$. Then X and Y are random variables. Then

$$\mathrm{Q}(\Psi') \geq \mathbb{E}[X - Y] = \mathbb{E}[X] - \mathbb{E}[Y]$$

where the expectation is over $u_1, \ldots, u_{k-1}, v_1, \ldots, v_N$.

We will prove

$$\mathbb{E}[X] = |B|\,|\Psi|\,M^{-(k-1)} \tag{5.12}$$

$$\mathbb{E}[Y] \leq |B| \max_{R,Q} 2^{NH(Q)+f(N)} M^{-(k-1)-r(R)} \tag{5.13}$$

with $f(N)$ as defined in (5.3), and $R \in \mathcal{R}(\Phi)$, $Q \in \mathcal{Q}(R, (P_1, \ldots, P_k))$. Before proving (5.12) and (5.13) we derive the final bound.

Derivation of final bound

From (5.12) and (5.13) follows

$$\mathbb{E}[X] - \mathbb{E}[Y] \geq |B|\,|\Psi|\,M^{-(k-1)} - |B| \max_{R,Q} 2^{NH(Q)+f(N)} M^{-(k-1)-r(R)}.$$

We factor out $|B|$, $|\Psi|$ and $M^{-(k-1)}$,

$$\mathbb{E}[X] - \mathbb{E}[Y] \geq |B|\,|\Psi|\,M^{-(k-1)} \Big(1 - \frac{1}{|\Psi|} \max_{R,Q} 2^{NH(Q)+f(N)} M^{-r(R)}\Big).$$

From our choice of $\mu(N)$ from (5.5),

$$\mu(N) = \max_{R,Q} \frac{H(Q) - H(P) + (1 + g(N) + f(N))\frac{1}{N}}{r(R)},$$

follows

$$\max_{R,Q} 2^{N(H(Q)-H(P)-r(R)\mu(N))+g(N)+f(N)} \leq \frac{1}{2}. \tag{5.14}$$

Apply $|B| \geq M^{1-\kappa(M)}$ from (5.10) and $|\Psi| \geq 2^{NH(P)-g(N)}$ from (5.9) to get

$$\begin{aligned}\mathbb{E}[X] - \mathbb{E}[Y] &\geq M^{1-\kappa(M)} 2^{NH(P)-g(N)} M^{-(k-1)} \\ &\qquad \cdot \Big(1 - 2^{-NH(P)+g(N)} \max_{R,Q} 2^{NH(Q)+f(N)} M^{-r(R)}\Big) \\ &\geq M^{-(k-2+\kappa(M))} 2^{NH(P)-g(N)}\end{aligned}$$

$$\cdot \Big(1 - \max_{R,Q} 2^{NH(Q)-NH(P)+g(N)+f(N)} M^{-r(R)}\Big).$$

(Here we used (5.14) to see that the second factor is nonnegative.) Apply the upper bound $2^{\mu(N)N} \leq M \leq 2^{\mu(N)N+2}$ from (5.6) to get

$$\begin{aligned}\mathbb{E}[X] - \mathbb{E}[Y] &\geq (2^{\mu(N)N+2})^{-(k-2+\kappa(M))} 2^{NH(P)-g(N)} \\ &\qquad \cdot \Big(1 - \max_{R,Q} 2^{NH(Q)-NH(P)+g(N)+f(N)} (2^{\mu(N)N})^{-r(R)}\Big) \\ &= 2^{N(H(P)-(k-2+\kappa(M))\mu(N))-2(k-2+\kappa(M))-g(N)} \\ &\qquad \cdot \Big(1 - \max_{R,Q} 2^{N(H(Q)-H(P)-r(R)\mu(N))+g(N)+f(N)}\Big).\end{aligned}$$

Using (5.14) we get

$$\begin{aligned}\mathbb{E}[X] - \mathbb{E}[Y] &\geq 2^{N(H(P)-(k-2+\kappa(M))\mu(N))-2(k-2+\kappa(M))-g(N)} (1 - \frac{1}{2}) \\ &= 2^{N(H(P)-(k-2+\kappa(M))\mu(N))-2(k-2+\kappa(M))-g(N)-1}.\end{aligned}$$

Then

$$\begin{aligned}&\frac{1}{N} \log_2 \mathrm{Q}(\Phi^{\otimes N}) \\ &\quad \geq \frac{1}{N} \log_2(\mathbb{E}[X] - \mathbb{E}[Y]) \\ &\quad \geq H(P) - (k-2+\kappa(M)) \max_{R,Q} \frac{H(Q) - H(P) + (1 + g(N) + f(N))\frac{1}{N}}{r(R)} \\ &\qquad - \frac{2(k-2+\kappa(M)) + g(N) + 1}{N}.\end{aligned}$$

We let t and thus N go to infinity, and obtain

$$\log_2 \underset{\sim}{\mathrm{Q}}(\Phi) \geq H(P) - (k-2) \max_{R,Q} \frac{H(Q) - H(P)}{r(R)}.$$

This lower bound holds for any rational probability distribution P on Φ and by continuity for any real probability distribution P on Φ.

It remains to prove (5.12) and (5.13). We do this in the lemmas below. □

Lemma 5.12. $\mathbb{E}[X] = |B|\,|\Psi|\,M^{-(k-1)}$.

Proof. Let $a \in \Psi$. Then $h_1(a_1) + \cdots + h_{k-1}(a_{k-1}) = (k-1)h_k(a_k)$. The following four statements are equivalent:

$$\begin{aligned}&a \in \Psi' \\ &\forall i \in [k]\ h_i(a_i) \in B\end{aligned}$$

$$\exists b \in B\ h_1(a_1) = \cdots = h_k(a_k) = b$$
$$\exists b \in B\ h_1(a_1) = \cdots = h_{k-1}(a_{k-1}) = b.$$

Therefore,

$$\mathbb{P}[a \in \Psi'] = \sum_{b \in B} \mathbb{P}[h_1(a_1) = \cdots = h_{k-1}(a_{k-1}) = b].$$

For $b \in B$,

$$\mathbb{P}[h_1(a_1) = \cdots = h_{k-1}(a_{k-1}) = b] = (M^{-1})^{k-1}.$$

We conclude

$$\begin{aligned}
\mathbb{E}[X] &= \sum_{a \in \Psi} \mathbb{P}[a \in \Psi'] \\
&= \sum_{a \in \Psi} \sum_{b \in B} \mathbb{P}[h_1(a_1) = \cdots = h_{k-1}(a_{k-1}) = b] \\
&= \sum_{a \in \Psi} \sum_{b \in B} (M^{-1})^{k-1} \\
&= |\Psi|\, |B|\, M^{-(k-1)}.
\end{aligned}$$

This proves the lemma. □

Lemma 5.13. $\mathbb{E}[Y] \leq |B| \max_{R,Q} 2^{NH(Q)+f(N)} M^{-(k-1)-r(R)}$.

Proof. Let

$$C = \{(a, a') \in \Psi^2 : a \neq a'; \exists i \in [k]\ a_i = a'_i\}.$$

Let $(a, a') \in C$. The following statements are equivalent:

$$(a, a') \in C' \tag{5.15}$$
$$a, a' \in \Psi' \tag{5.16}$$
$$\forall i \in [k]\ h_i(a_i), h_i(a'_i) \in B \tag{5.17}$$
$$\exists b \in B\ h_1(a_1) = \cdots = h_k(a_k) = h_1(a'_1) = \cdots = h_k(a'_k) = b. \tag{5.18}$$

Therefore,

$$\begin{aligned}
\mathbb{E}[Y] &= \sum_{(a,a') \in C} \mathbb{P}[(a, a') \in C'] \\
&= \sum_{(a,a') \in C} \sum_{b \in B} \mathbb{P}[h_1(a_1) = \cdots = h_k(a_k) = h_1(a'_1) = \cdots = h_k(a'_k) = b].
\end{aligned}$$

Let $(a,a') \in C$. Then $h_i(a_i)$ and $h_i(a'_i)$ are $\mathbb{Z}/M\mathbb{Z}$-linear combinations of $u_1,\ldots,u_{k-1},v_1,\ldots,v_N$. The random variable

$$\big(h_1(a_1),\ldots,h_k(a_k),h_1(a'_1),\ldots,h_k(a'_k)\big)$$

is uniformly distributed over the image subspace $V \subseteq (\mathbb{Z}/M\mathbb{Z})^{2k}$. Let $b \in B$. Then $(b,\ldots,b) \in V$, since $u_1 = \cdots = u_k = b$, $v_1,\ldots,v_N = 0$ is a valid assignment. Therefore,

$$\mathbb{P}[h_1(a_1) = \cdots = h_k(a_k) = h_1(a'_1) = \cdots = h_k(a'_k) = b] = |V|^{-1}.$$

And $|V|$ equals M to the power the rank of the matrix

$$\begin{pmatrix} 1 & 0 & \cdots & 0 & \frac{1}{k-1} & 1 & 0 & \cdots & 0 & \frac{1}{k-1} \\ 0 & 1 & & 0 & \frac{1}{k-1} & 0 & 1 & & 0 & \frac{1}{k-1} \\ \vdots & & \ddots & & \vdots & & & \ddots & & \vdots \\ 0 & 0 & \cdots & 1 & \frac{1}{k-1} & 0 & 0 & \cdots & 1 & \frac{1}{k-1} \\ a_1 & a_2 & \cdots & a_{k-1} & -\frac{a_k}{k-1} & a'_1 & a'_2 & \cdots & a'_{k-1} & -\frac{a'_k}{k-1} \end{pmatrix} \tag{5.19}$$

over $\mathbb{Z}/M\mathbb{Z}$, with $a_1,\ldots,a_k,a'_1,\ldots,a'_k$ thought of as column vectors in $(\mathbb{Z}/M\mathbb{Z})^N$. With column operations we transform (5.19) into

$$\begin{pmatrix} 0 & 0 & \cdots & 0 & 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 & 0 & 1 & & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & & \ddots & & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 0 & 0 & & 1 & 0 \\ a_1 - a'_1 & a_2 - a'_2 & \cdots & a_{k-1} - a'_{k-1} & a_k - a'_k & a'_1 & a'_2 & \cdots & a'_{k-1} & 0 \end{pmatrix} \tag{5.20}$$

Matrix (5.20) has rank equal to $k-1$ plus $r_M(a,a') = \mathrm{rk}(A(a,a'))$, where

$$A(a,a') \coloneqq \begin{pmatrix} a_1 - a'_1 & a_2 - a'_2 & \cdots & a_k - a'_k \end{pmatrix}.$$

We obtain

$$\mathbb{E}[Y] \leq \sum_{(a,a')\in C} \sum_{b \in B} M^{-(k-1+r_M(a,a'))}.$$

Since the summands are independent of b, we get

$$\mathbb{E}[Y] \leq |B| \sum_{(a,a')\in C} M^{-(k-1+r_M(a,a'))}.$$

Let $(a,a') \in C$. Consider the rows of $A(a,a')$. The N rows are of the form $x_i - y_i$ with $(x_i,y_i) \in \Phi^2$. Let $s = ((x_1,y_1),\ldots,(x_N,y_N))$. Let $R =$

$\{(x_1, y_1), \dots, (x_N, y_N)\}$. We have $r_M(a, a') = r_M(R)$ and $r_M(R) = r(R)$ by (5.8). Let Q be the N-type with $\mathrm{supp}(Q) = R$ and $s \in T^N_Q$. From $a \neq a'$ follows $R \not\subseteq \{(x, x) : x \in \Phi\}$. From $\exists i \in [k]\ a_i = a'_i$ follows $\exists i \in [k]\ R \subseteq \{(x, y) : x_i = y_i\}$. From $a, a' \in T^N_{P_1} \times \dots \times T^N_{P_k}$ follows $Q_i = Q_{k+i} = P_i$ for all $i \in [k]$. We thus have

$$\mathbb{E}[Y] \leq |B| \sum_{R \in \mathcal{R}(\Phi)} \sum_{\substack{Q \in \mathcal{Q}(R,(P_1,\dots,P_k)):\\ \mathrm{supp}(Q)=R\\ Q \text{ is } N\text{-type}}} \sum_{s \in T^N_Q} M^{-(k-1+r(R))}.$$

The number of N-types Q with $\mathrm{supp}(Q) = R$ is at most the number of N-types on R, which is at most $\binom{N+|R|-1}{|R|-1}$ (Lemma 4.19). For any $Q \in \mathcal{Q}(R, (P_1, \dots, P_k))$, $|T^N_Q| \leq 2^{NH(Q)}$ (Lemma 4.19). Therefore,

$$\mathbb{E}[Y] \leq |B| \sum_{R \in \mathcal{R}(\Phi)} \binom{N + |R| - 1}{|R| - 1} \max_{Q \in \mathcal{Q}(R,(P_1,\dots,P_k))} 2^{NH(Q)}\ M^{-(k-1+r(R))}.$$

Also $|\mathcal{R}(\Phi)| \leq 2^{|\Phi|^2}$. Therefore,

$$\mathbb{E}[Y] \leq |B|\, 2^{|\Phi|^2} \max_{R \in \mathcal{R}(\Phi)} \binom{N + |R| - 1}{|R| - 1} \max_{Q \in \mathcal{Q}(R,(P_1,\dots,P_k))} 2^{NH(Q)}\ M^{-(k-1+r(R))}.$$

We conclude that

$$\mathbb{E}[Y] \leq |B| \max_{R,Q} 2^{NH(Q)+f(N)} M^{-(k-1)-r(R)}.$$

This proves the lemma. □

5.2.2 Computational remarks

The following two lemmas are helpful when applying Theorem 5.7. We leave the proof to the reader.

Lemma 5.14. *Let $P \in \mathcal{P}(\Phi)$. Let $R, R' \in \mathcal{R}(\Phi)$ with $R \subseteq R'$ and $r(R) = r(R')$. Then*

$$\max_{Q \in \mathcal{Q}(R,(P_1,\dots,P_k))} \frac{H(Q) - H(P)}{r(R)} \leq \max_{Q \in \mathcal{Q}(R',(P_1,\dots,P_k))} \frac{H(Q) - H(P)}{r(R')}.$$

Lemma 5.15. *Let $R \in \mathcal{R}(\Phi)$. There is an equivalence relation $R' \in \mathcal{R}(\Phi)$ with $R \subseteq R'$ and $r(R) = r(R')$.*

5.2.3 Examples: type sets

We discuss some examples. The first example we will use to get good upper bounds on the asymptotic rank of complete graph tensors in Section 5.5. We focus on one family of examples that is parametrised by partitions. Let $\lambda \vdash k$ be an integer partition of k with d parts. Let

$$\Phi_\lambda = \{a \in \{0, 1, \ldots, d-1\} : \mathrm{type}(a) = \lambda\}.$$

The set Φ_λ is tight.

Theorem 5.16. $\log_2 \underset{\sim}{\mathrm{Q}}(\Phi_{(2,2)}) = 1$.

Proof. Let $\Phi = \Phi_{(2,2)}$. Clearly, $\underset{\sim}{\mathrm{Q}}(\Phi) \leq 2$. After relabelling, $\forall a \in \Phi\ \sum_{i=1}^k a_i = 0$. We may thus apply Theorem 5.7. Let P be the uniform probability distribution on Φ. Then $H(P) = \log_2 6$.

Let $R \in \mathcal{R}(\Phi)$. We may assume that

$$\begin{aligned} R \subseteq\ & \{(1,1,0,0),(1,0,1,0),(1,0,0,1)\}^2 \\ & \cup \{(0,0,1,1),(0,1,0,1),(0,1,1,0)\}^2. \end{aligned}$$

We may assume R is an equivalence relation (Lemma 5.15). Let $(x, y) \in R$. Let $R' = R \cup \{((1,1,1,1) - x, (1,1,1,1) - y)\} \in \mathcal{R}(\Phi)\}$. Then $R \subseteq R'$ and $R' \in \mathcal{R}(\Phi)$ and $r(R) = r(R')$. We may thus assume that if $(x, y) \in R$, then also $((1,1,1,1) - x, (1,1,1,1) - y) \in R$ (Lemma 5.14).

Let $S = \{(1,1,0,0),(1,0,1,0),(1,0,0,1)\}$. By the above observation, it suffices to consider equivalence relations on S. There are three types of such equivalence relations.

Type (3): all three elements of S are equivalent. Then $|R| = 18$ and $r(R) = 2$.

Type $(2,1)$: two elements of S are equivalent and inequivalent to the third element (which is equivalent to itself). Then $|R| = 10$ and $r(R) = 1$.

Type $(1,1,1)$): all elements of S are inequivalent. Then $R \subseteq \{(x,x) : x \in \Phi\}$ which is a contradiction.

For type (3) and $(2,1)$, the uniform probability distribution Q on R has marginals $Q_i = Q_{4+i} = P_i$ for $i \in [4]$. The uniform Q is optimal. Then $H(Q) = \log_2 |R|$. Let $R_{(3)}$ and $R_{(2,1)}$ be equivalence relations of type (3) and $(2,1)$. Then

$$\begin{aligned} \log_2 \underset{\sim}{\mathrm{Q}}(\Phi) \geq \min\Big\{ & H(P) - \frac{2}{r(R_{(3)})}\big(\log_2 |R_{(3)}| - H(P)\big), \\ & H(P) - \frac{2}{r(R_{(2,1)})}\big(\log_2 |R_{(2,1)}| - H(P)\big)\Big\} \\ = \min\{ & \log_2 6 - \tfrac{2}{2}(\log_2 18 - \log_2 6), \\ & \log_2 6 - \tfrac{2}{1}(\log_2 10 - \log_2 6)\} \\ = \min\{ & 1, \log_2 \tfrac{54}{25}\} = 1 \end{aligned}$$

This proves the theorem. ∎

Theorem 5.17. $\log_2 \underset{\sim}{\mathrm{Q}}(\Phi_{(0^{k-1},1)}) = h(1/k)$.

Proof. We refer to [CVZ16]. □

With Srinivasan Arunachalam and Péter Vrana we have the following unpublished result.

Theorem 5.18. $\log_2 \underset{\sim}{\mathrm{Q}}(\Phi_{(0^{k/2},1^{k/2})}) = 1$.

5.3 Combinatorial degeneration method

In this section we extend the (higher-order) Coppersmith–Winograd method via a preorder called combinatorial degeneration. Suppose $\Psi \subseteq I_1 \times \cdots \times I_k$ 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 $\underset{\sim}{\mathrm{Q}}(\Psi)$ via Φ. This has an application in the context of tri-colored sum-free sets (Section 5.4.2) for example.

Definition 5.19 ([BCS97])**.** Let $\Phi \subseteq \Psi \subseteq I_1 \times \cdots \times I_k$. We say that Φ is a *combinatorial degeneration of* Ψ, and write $\Psi \trianglerighteq \Phi$, 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 5.20. *Let $t \in \mathbb{F}^{n_1} \otimes \cdots \otimes \mathbb{F}^{n_k}$. Let $\Psi = \mathrm{supp}(t)$. Let $\Phi \subseteq \Psi$ such that $\Psi \trianglerighteq \Phi$. Then $t \trianglerighteq t|_\Phi$.*

Proposition 5.20 brings us only slightly closer to our goal. Namely, given $t \in \mathbb{F}^{n_1} \otimes \cdots \otimes \mathbb{F}^{n_k}$ with $\Psi = \mathrm{supp}(t)$, and given $\Phi \subseteq \Psi$ such that $\Psi \trianglerighteq \Phi$, it follows directly from Proposition 5.20 that $t \trianglerighteq t|_\Phi$ and thus $\underset{\sim}{\mathrm{Q}}(t) \geq \underset{\sim}{\mathrm{Q}}(t|_\Phi)$. This, however, does not give us a lower bound on the combinatorial asymptotic subrank $\underset{\sim}{\mathrm{Q}}(\Psi)$. The following theorem does. Our theorem extends a result in [KSS16].

Theorem 5.21. *Let $\Phi \subseteq \Psi \subseteq I_1 \times \cdots \times I_k$. If $\Psi \trianglerighteq \Phi$, then*

$$\underset{\sim}{\mathrm{Q}}(\Psi) \geq \underset{\sim}{\mathrm{Q}}(\Phi).$$

Lemma 5.22. *Let $\Phi \subseteq \Psi \subseteq I_1 \times \cdots \times I_k$. If $\Psi \trianglerighteq \Phi$, then $\underset{\sim}{\mathrm{Q}}(\Psi) \geq \mathrm{Q}(\Phi)$.*

Proof. Pick maps $u_i : I_i \to \mathbb{Z}$ such that

$$\sum_{i=1}^k u_i(\alpha_i) = 0 \quad \text{for } \alpha \in \Phi$$

$$\sum_{i=1}^k u_i(\alpha_i) > 0 \quad \text{for } \alpha \in \Psi \setminus \Phi.$$

Let D be a free diagonal in Φ with $|D| = \mathrm{Q}(\Phi)$ and let

$$w_i = \sum_{x \in D_i} u_i(x).$$

Let $n \in \mathbb{N}$ and define

$$W_i = \big\{(x_1, \ldots, x_{n|D|}) \in I_i^{\times n|D|} : \sum_{j=1}^{n|D|} u_i(x_j) = n w_i\big\}.$$

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

$$\begin{matrix} u_1(x_{1,1}) & u_2(x_{2,1}) & \cdots & u_k(x_{k,1}) \\ u_1(x_{1,2}) & u_2(x_{2,2}) & \cdots & u_k(x_{k,2}) \\ \vdots & \vdots & \ddots & \vdots \\ u_1(x_{1,n|D|}) & u_2(x_{2,n|D|}) & \cdots & u_k(x_{k,n|D|}) \end{matrix}$$

The sum of the ith 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 Φ, $D^{\times n|D|}$ 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.

$$\mathrm{Q}(\Psi^{\times n|D|}) \geq |D^{\times n|D|} \cap (W_1 \times \cdots \times W_k)|.$$

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

$$\mathrm{Q}(\Psi^{\times n|D|}) \geq \binom{n|D|}{n, \ldots, n} = |D|^{n|D| - o(n)},$$

which implies $\underset{\sim}{\mathrm{Q}}(\Psi) = \lim_{n\to\infty} \mathrm{Q}(\Psi^{\times n|D|})^{\frac{1}{n|D|}} \geq |D|$. ◻

Proof of Theorem 5.21. We have $\underset{\sim}{\mathrm{Q}}(\Psi) = \lim_{n\to\infty} \underset{\sim}{\mathrm{Q}}(\Psi^{\times n})^{1/n}$. It follows from Lemma 5.22 that

$$\lim_{n\to\infty} \underset{\sim}{\mathrm{Q}}(\Psi^{\times n})^{1/n} \geq \lim_{n\to\infty} \mathrm{Q}(\Phi^{\times n})^{1/n}.$$

The right-hand side is $\underset{\sim}{\mathrm{Q}}(\Phi)$. ◻

5.4 Cap sets

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 $(u, u+v, u+2v)$. Until recently it was not known 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 question, showing that $c \leq 3(207+33\sqrt{33})^{1/3}/8 \approx 2.755$. Tao realised in [Tao16] that the cap set question 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 $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{F}_3$ 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 that the cap set tensor is in the $\mathrm{GL}_3(\mathbb{F}_3)^{\times 3}$ orbit of the "reduced polynomial multiplication tensor", which was studied in [Str91], and we show how recent results follow from this connection, using Theorem 5.21.

5.4.1 Reduced polynomial multiplication

Let t_n be 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 $\alpha_1 + \alpha_2 = \alpha_3$. We call t_n the reduced polynomial multiplication tensor, since t_n is essentially the structure tensor of the algebra $\mathbb{F}[x]/(x^n)$ of univariate polynomials modulo the ideal generated by x^n. The support of t_n equals

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

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

$$\Phi_n = \big\{(\alpha_1, \alpha_2, \alpha_3) \in \{0, \ldots, n-1\}^3 \,\big|\, \alpha_1 + \alpha_2 + \alpha_3 = n - 1\big\}. \tag{5.21}$$

The support Φ_n is tight (cf. Example 5.1). Strassen proves in [Str91, Theorem 6.7] using Corollary 5.4 that $\underset{\sim}{\mathrm{Q}}(t_n) = \underset{\sim}{\mathrm{Q}}(\Phi_n) = z(n)$, where $z(n)$ is defined as

$$z(n) := \frac{\gamma^n - 1}{\gamma - 1} \gamma^{-2(n-1)/3} \tag{5.22}$$

with γ equal to the unique positive real solution of the equation $\frac{1}{\gamma-1} - \frac{n}{\gamma^n - 1} = \frac{n-1}{3}$. The following table contains values of $z(n)$ for small n. See also [Str91, Table 1].

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

In fact, [Str91, Theorem 6.7] says that the asymptotic spectrum of t_n is completely determined by the support functionals; and that the possible values that the spectral points can take on t_n form the closed interval $[z(n), n]$ (cf. Remark 2.22),

$$\mathbf{X}(\mathbb{N}[t_n]) = \{\zeta^\theta|_{\mathbb{N}[t_n]} : \theta \in \mathcal{P}([3])\}, \quad \{\phi(t_n) : \phi \in \mathbf{X}(\mathbb{N}[t_n])\} = [z(n), n].$$

5.4.2 Cap sets

We turn to cap sets.

Definition 5.23. 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 $\underaccent{\sim}{r}_3(\mathbb{Z}/m\mathbb{Z}) = \lim_{n\to\infty} r_3((\mathbb{Z}/m\mathbb{Z})^n)^{1/n}$.

A three-term progression-free set in $(\mathbb{Z}/3\mathbb{Z})^n$ is called a *cap* or *cap set*. We next discuss 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 sets turn out to be upper bounds on the size of tri-colored sum-free sets.

Definition 5.24. Let G be an abelian group. Let $\Gamma \subseteq G \times G \times G$. For $i \in [3]$ we define the marginal sets $\Gamma_i = \{x \in G : \exists \alpha \in \Gamma : \alpha_i = x\}$. We say Γ is *tricolored sum-free* if the following holds. The set Γ 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 $\underaccent{\sim}{s}_3(G) = \lim_{n\to\infty} s_3(G^{\times n})^{1/n}$.

Equivalently, $\Gamma \subseteq G \times G \times G$ is a tricolored sum-free set if and only if Γ is a free diagonal in $\{\alpha \in G \times G \times G : \alpha_1 + \alpha_2 + \alpha_3 = 0\}$.

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 $\underset{\sim}{r}_3(\mathbb{Z}/m\mathbb{Z}) \leq \underset{\sim}{s}_3(\mathbb{Z}/m\mathbb{Z})$.

We summarise the recent history of results on cap sets. 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 $2.21739 \leq \underset{\sim}{r}_3(\mathbb{Z}/3\mathbb{Z})$. In [EG17] Ellenberg and Gijswijt proved the upper bound

$$\underset{\sim}{r}_3(\mathbb{Z}/3\mathbb{Z}) \leq 3(207 + 33\sqrt{33})^{1/3}/8 \approx 2.755,$$

Blasiak et al. [BCC+17] proved that in fact

$$\underset{\sim}{s}_3(\mathbb{Z}/3\mathbb{Z}) \leq 3(207 + 33\sqrt{33})^{1/3}/8.$$

This upper bound was shown to be an equality in [KSS16, Nor16, Peb16]:

Theorem 5.25. $\underset{\sim}{s}_3(\mathbb{Z}/3\mathbb{Z}) = 3(207 + 33\sqrt{33})^{1/3}/8.$

We reprove Theorem 5.25 by proving that $\underset{\sim}{s}_3(\mathbb{Z}/m\mathbb{Z})$ equals the asymptotic subrank $z(m)$ of t_m discussed in Section 5.4.1, when m is a prime power. 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 [BCC+17, KSS16, Nor16, Peb16].

Proof. We will prove $\underset{\sim}{s}_3(\mathbb{Z}/m\mathbb{Z}) = z(m)$ when m is a prime power. By definition, $\underset{\sim}{s}_3(\mathbb{Z}/m\mathbb{Z})$ equals the asymptotic subrank of the set

$$\{\alpha \in \{0, \ldots, m-1\}^3 : \alpha_1 + \alpha_2 + \alpha_3 = 0 \bmod m\}$$

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

$$\Psi_m = \{\alpha \in \{0, \ldots, m-1\}^3 : \alpha_1 + \alpha_2 + \alpha_3 = m - 1 \bmod m\}$$

and so $\underset{\sim}{s}_3(\mathbb{Z}/m\mathbb{Z}) = \underset{\sim}{\mathrm{Q}}(\Psi_m)$. Let

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

We know $\underset{\sim}{\mathrm{Q}}(\Phi_m) = z(m)$ (Section 5.4.1). We will show that $\underset{\sim}{\mathrm{Q}}(\Phi_m) = \underset{\sim}{\mathrm{Q}}(\Psi_m)$ when m is a prime power. This proves the theorem.

We prove $\underset{\sim}{\mathrm{Q}}(\Phi_m) \leq \underset{\sim}{\mathrm{Q}}(\Psi_m)$. There is a combinatorial degeneration $\Phi_m \trianglelefteq \Psi_m$. Indeed, let $u_i : \{0, \ldots, m-1\} \to \{0, \ldots, m-1\}$ 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 5.21 applies, and we thus obtain $\underset{\sim}{\mathrm{Q}}(\Phi_m) \leq \underset{\sim}{\mathrm{Q}}(\Psi_m)$. This proves the claim.

We show $\underset{\sim}{\mathrm{Q}}(\Psi_m) \leq \underset{\sim}{\mathrm{Q}}(\Phi_m)$ when m is a power of the prime p. Let $\mathbb{F} = \mathbb{F}_p$. Let $f_m \in \mathbb{F}^m \otimes \mathbb{F}^m \otimes \mathbb{F}^m$ have support Ψ_m with all nonzero coefficients equal to 1. Obviously, $\underset{\sim}{\mathrm{Q}}(\Psi_m) \leq \underset{\sim}{\mathrm{Q}}(f_m)$. To compute $\underset{\sim}{\mathrm{Q}}(f_m)$ we show that there is a basis in which the support of f_m equals the tight set Φ_m. Then $\underset{\sim}{\mathrm{Q}}(f_m) = \underset{\sim}{\mathrm{Q}}(\Phi_m)$ (Corollary 5.4). 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 $\mathbb{Z}/m\mathbb{Z} \to \mathbb{Z}/p\mathbb{Z}$, since for $a \in \{0, 1, \ldots, m-1\}$, if $x = y \bmod m$ then $\binom{x}{a} = \binom{y}{a} \bmod p$ by Lucas' theorem. Let $(e_x)_x$ be the standard basis of $\mathbb{F}^m$. The elements $(\sum_{x \in \mathbb{Z}/m\mathbb{Z}} \binom{x}{a} e_x)_{a \in \mathbb{Z}/m\mathbb{Z}}$ form a basis of $\mathbb{F}^m$ since the matrix $(\binom{x}{a})_{a,x}$ is upper triangular with ones on the diagonal. We will now rewrite f_m in the basis $((\sum_x \binom{x}{a} e_x)_a, (\sum_y \binom{y}{b} e_y)_b, (\sum_z \binom{z}{c} e_z)_c)$. Observe that $\binom{x}{m-1}$ equals 1 if and only if x equals $m-1$, and hence

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

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

$$\begin{aligned}
&\sum_{x,y,z \in \mathbb{Z}/m\mathbb{Z}} \binom{x+y+z}{m-1} e_x \otimes e_y \otimes e_z \\
&\quad = \sum_{x,y,z \in \mathbb{Z}/m\mathbb{Z}} \; \sum_{\substack{a,b,c \in \{0,1,\ldots,m-1\}:\\ a+b+c=m-1}} \binom{x}{a}\binom{y}{b}\binom{z}{c} e_x \otimes e_y \otimes e_z. \qquad (5.23)
\end{aligned}$$

We may simply rewrite (5.23) as

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

Therefore, with respect to the basis $((\sum_x \binom{x}{a} e_x)_a, (\sum_y \binom{y}{b} e_y)_b, (\sum_z \binom{z}{c} e_z)_c)$, the support of f_m equals the tight set Φ_m. (And even stronger, f_m is isomorphic to the tensor $\mathbb{F}[x]/(x^m)$ of Section 5.4.1.) $\square$

Remark 5.26. Why did we reprove the cap set result Theorem 5.25? Our motivation, being interested in the asymptotic spectrum of tensors, was to see if the techniques in the cap set papers are stronger than the Strassen support functionals, i.e. whether they give any new spectral points. Above we have seen that the cap set result itself can be proven with the support functionals. In fact, we show in Section 4.6 that for oblique tensors the asymptotic slice-rank, which was introduced in [Tao16] to give a concise proof of [EG17], equals the minimum value over the support functionals. In Section 6.11 we show that for all complex tensors asymptotic slice-rank equals the minimum value of the quantum functionals.

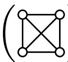
5.5 Graph tensors

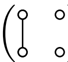
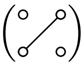
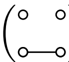
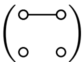
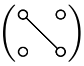
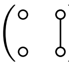
In this section we briefly discuss the application that motivated us to prove Theorem 5.7 in [CVZ16], namely upper bounding the asymptotic rank of so-called graph tensors. Graph tensors are defined as follows.

Let $G = (V, E)$ be a graph (or hypergraph) with vertex set V and edge set E. Let $n \in \mathbb{N}$. Let $(b_i)_{i\in[n]}$ be the standard basis of $\mathbb{F}^n$. We define the graph tensor $\mathrm{T}_n(G)$ as

$$\mathrm{T}_n(G) \coloneqq \sum_{i\in[n]^E} \bigotimes_{v\in V} \Big(\bigotimes_{\substack{e\in E:\\ v\in e}} b_{i_e} \Big)$$

seen as a $|V|$-tensor. Given a vertex $v \in V$ let $d(v)$ denote the degree of v, that is, $d(v)$ equals the number of edges $e \in E$ that contain v. Then $\mathrm{T}_n(G)$ is naturally in $\bigotimes_{v\in V} \mathbb{F}^{d(v)}$. We write $\mathrm{T}(G)$ for $\mathrm{T}_2(G)$. For example, for the complete graph on four vertices K_4 the graph tensor is

$\mathrm{T}(K_4) = \mathrm{T}\Big($  $\Big)$

$= \mathrm{T}\Big($  $\Big) \otimes \mathrm{T}\Big($  $\Big) \otimes \mathrm{T}\Big($  $\Big) \otimes \mathrm{T}\Big($  $\Big) \otimes \mathrm{T}\Big($  $\Big) \otimes \mathrm{T}\Big($  $\Big)$

$$= \sum_{i\in\{0,1\}^6} (b_{i_1}\otimes b_{i_2}\otimes b_{i_5}) \otimes (b_{i_2}\otimes b_{i_3}\otimes b_{i_6}) \otimes (b_{i_3}\otimes b_{i_4}\otimes b_{i_5}) \otimes (b_{i_1}\otimes b_{i_4}\otimes b_{i_6})$$

living in $(\mathbb{C}^8)^{\otimes 4}$. Let K_k be the complete graph on k vertices. The 2×2 matrix multiplication tensor $\langle 2,2,2\rangle$ equals the tensor $\mathrm{T}(K_3)$. Define the exponent $\omega(\mathrm{T}(G)) = \log_2 \underaccent{\sim}{\mathrm{R}}(\mathrm{T}(G))$. We study the exponent per edge $\tau(\mathrm{T}(G)) = \omega(\mathrm{T}(G))/\,|E(G)|$.

Our result is an upper bound on $\tau(\mathrm{T}(K_4))$ in terms of the combinatorial asymptotic subrank $\underaccent{\sim}{\mathrm{Q}}(\Phi_{(2,2)})$ which we studied in Theorem 5.16.

Theorem 5.27. *For any $q \geq 1$, $\tau(\mathrm{T}(K_4)) \leq \log_q\Big(\dfrac{q+2}{\underaccent{\sim}{\mathrm{Q}}(\Phi_{(2,2)})}\Big)$.*

Proof. We apply a generalisation of the laser method. See [CVZ16]. $\square$

Corollary 5.28. *Let $k \geq 4$. Then $\tau(\mathrm{T}(K_k)) \leq 0.772943$.*

Proof. In the bound of Theorem 5.27 we plug in the value $\underaccent{\sim}{\mathrm{Q}}(\Phi_{(2,2)}) = 2$ from Theorem 5.16. Then we optimise over q to obtain the value 0.772943. By a "covering argument" we can show that $\tau(\mathrm{T}(K_k))$ is non-increasing when k increases. $\square$

For $k \geq 4$ Corollary 5.28 improves the upper bound $\tau(\mathrm{T}(K_k)) \leq 0.790955$ that can be derived from the well-known upper bound of Le Gall [LG14] on the exponent of matrix multiplication $\omega \coloneqq \omega(\mathrm{T}(K_3))$.

A standard "flattening argument" (i.e. using the gauge points from the asymptotic spectrum) yields the lower bound $\tau(\mathrm{T}(K_k)) \geq \frac{1}{2}k/(k-1)$ if k is even and $\tau(\mathrm{T}(K_k)) \geq \frac{1}{2}(k+1)/k$ if k is odd. As a consequence, if the exponent of matrix multiplication ω equals 2, then $\tau(\mathrm{T}(K_4)) = \tau(\mathrm{T}(K_3)) = \frac{2}{3}$. We raise the following question: is there a $k \geq 5$ such that $\tau(\mathrm{T}(K_k)) < \frac{2}{3}$?

Tensor surgery; cycle graphs

For graph tensors given by sparse graphs, good upper bounds on the asymptotic rank can be obtained with an entirely different method called tensor surgery, which we introduced in [CZ18]. As an illustration let me mention the results we obtained for cycle graphs with tensor surgery. Recall $\omega = \log_2 \underset{\sim}{\mathrm{R}}(\langle 2,2,2\rangle) = \log_2 \underset{\sim}{\mathrm{R}}(\mathrm{T}(C_3))$. Let $\omega_k \coloneqq \log_2 \underset{\sim}{\mathrm{R}}(\mathrm{T}(C_k))$. First observe that $\omega_k = k$ for even k. For odd k, trivially $k-1 \leq \omega_k \leq k$. We prove the following.

Theorem 5.29. *For k, ℓ odd, $\omega_{k+\ell-1} \leq \omega_k + \omega_\ell$.*

Corollary 5.30. *Let $k \geq 5$ odd. Then, $\omega_k \leq \omega_{k-2} + \omega_3$ and thus $\omega_k \leq \frac{k-1}{2}\omega$.*

Corollary 5.31. *If $\omega = 2$, then $\omega_k = k-1$ for all odd k.*

See [CZ18] for the proofs.

5.6 Conclusion

Tight tensors are a subfamily of the oblique tensors. For tight 3-tensors the minimum over the support functionals equals the asymptotic subrank. This is proven via the Coppersmith–Winograd method. The construction is in fact of a very combinatorial nature. In this chapter we studied the combinatorial notion of subrank. We proved that combinatorial subrank is monotone under combinatorial degeneration. We studied the cap set problem via the support functionals. We extended the Coppersmith–Winograd method to higher-order tensors, and applied this method to study graph tensors.

Chapter 6

Universal points in the asymptotic spectrum of tensors; entanglement polytopes, moment polytopes

This chapter is based on joint work with Matthias Christandl and Péter Vrana [CVZ18].

6.1 Introduction

In Chapter 4, following Strassen, we introduced the asymptotic spectrum of tensors $\mathbf{X}(\mathcal{T}) = \mathbf{X}(\mathcal{T}, \leqslant)$ for $\mathcal{T}$ the semiring of k-tensors over $\mathbb{F}$ for some fixed integer k and field $\mathbb{F}$, with addition given by direct sum $\oplus$, multiplication given by tensor product $\otimes$, and preorder $\leqslant$ given by restriction (or degeneration). The asymptotic spectrum characterises the asymptotic rank $\underset{\sim}{\mathrm{R}}$ and the asymptotic subrank $\underset{\sim}{\mathrm{Q}}$. We have seen that the asymptotic rank plays an important role in algebraic complexity theory: the asymptotic rank of the matrix multiplication tensor $\langle 2,2,2\rangle = \sum_{i,j,k\in[2]} e_{ij} \otimes e_{jk} \otimes e_{ki} \in \mathbb{F}^4 \otimes \mathbb{F}^4 \otimes \mathbb{F}^4$ characterises the exponent of the arithmetic complexity of multiplying two $n \times n$ matrices over $\mathbb{F}$, that is $\underset{\sim}{\mathrm{R}}(\langle 2,2,2\rangle) = 2^\omega$. We have also seen in Chapter 5 how one may use the asymptotic subrank to upper bound the size of combinatorial objects like for example cap sets in $\mathbb{F}_3^n$.

New results in this chapter

So far, the only elements we have seen in $\mathbf{X}(\mathcal{T})$ (i.e. universal spectral points cf. Section 2.13) are the gauge points (Section 4.3). Besides that we have seen in Section 4.4 that the Strassen support functionals ζ^θ are in $\mathbf{X}(\{\text{oblique}\})$. In this chapter we introduce for the first time an explicit *infinite* family of universal spectral points (over the complex numbers), the *quantum functionals*. Our new insight is to use the moment polytope. Given a tensor $t \in \mathbb{C}^{n_1} \otimes \mathbb{C}^{n_2} \otimes \mathbb{C}^{n_3}$, the moment polytope $\mathbf{P}(t)$ is a convex polytope that carries representation-theoretic information about t. The quantum functionals are defined as maximisations over moment polytopes.

Let us immediately put a disclaimer. The quantum functionals do not give a new lower bound on the asymptotic rank of matrix multiplication $\langle 2, 2, 2\rangle$, namely they give the same lower bound as the gauge points. Also, the quantum functionals being defined for tensors over complex numbers only, we do not expect to get new upper bounds on the size of combinatorial objects that are "like cap sets".

So what have we gained? Arguably, we have found the "right" viewpoint on how to construct universal spectral points for tensors. (In fact, after writing our paper [CVZ18] we realised that Strassen had begun a study of moment polytopes in the appendix of the German survey [Str05]! Strassen did not construct new universal spectral points, however; not in that publication at least.) If there are more universal spectral points, then our viewpoint may lead the way to finding them. Moreover, whereas no efficient algorithm is known for evaluating the support functionals, the moment polytope viewpoint may open the way to having efficient algorithms for evaluating the quantum functionals.

In Sections 6.2–6.7 we work towards the construction of the quantum functionals and we give a proof that they are universal spectral points. In Sections 6.8–6.10 we compare the quantum functionals and the support functionals, and in Section 6.11 we relate asymptotic slice rank to the quantum functionals.

In this chapter we will focus on 3-tensors, but the theory naturally generalises to k-tensors.

6.2 Schur–Weyl duality

For background on representation theory we refer to [Kra84], [Ful97] and [GW09].

Let S_n be the symmetric group on n symbols. Let S_n act on the tensor space $(\mathbb{C}^d)^{\otimes n}$ by permuting the tensor legs,

$$\pi \cdot v_1 \otimes \cdots \otimes v_n = v_{\pi^{-1}(1)} \otimes \cdots \otimes v_{\pi^{-1}(n)}, \quad \pi \in S_n.$$

Let GL_d be the general linear group of $\mathbb{C}^d$. Let GL_d act on $(\mathbb{C}^d)^{\otimes n}$ via the diagonal embedding $\mathrm{GL}_d \to \mathrm{GL}_d^{\times n} : g \mapsto (g, \ldots, g)$,

$$g \cdot v_1 \otimes \cdots \otimes v_n = (gv_1) \otimes \cdots \otimes (gv_n), \quad g \in \mathrm{GL}_d.$$

The actions of S_n and GL_d commute, so we have a well-defined action of the product group $S_n \times \mathrm{GL}_d$ on $(\mathbb{C}^d)^{\otimes n}$. Schur–Weyl duality describes the decomposition of the space $(\mathbb{C}^d)^{\otimes n}$ into a direct sum of irreducible $S_n \times \mathrm{GL}_d$ representations. This decomposition is

$$(\mathbb{C}^d)^{\otimes n} \cong \bigoplus_{\lambda \vdash_d n} [\lambda] \otimes \mathbb{S}_\lambda(\mathbb{C}^d) \tag{6.1}$$

with $[\lambda]$ an irreducible S_n representation of type λ and $\mathbb{S}_\lambda(\mathbb{C}^d)$ an irreducible GL_d-representation of type λ when $\ell(\lambda) \leq d$ and 0 when $\ell(\lambda) > d$. We use the notation $\lambda \vdash_d n$ for the partitions of n with at most d parts. Let

$$P_\lambda : (\mathbb{C}^d)^{\otimes n} \to (\mathbb{C}^d)^{\otimes n}$$

be the equivariant projector onto the isotypical component of type λ, i.e. onto the subspace of $(\mathbb{C}^d)^{\otimes n}$ isomorphic to $[\lambda] \otimes \mathbb{S}_\lambda(\mathbb{C}^d)$. The projector P_λ is given by the action of the group algebra element

$$P_\lambda = \Big(\frac{\dim[\lambda]}{n!}\Big)^2 \sum_{T \in \mathrm{Tab}(\lambda)} c_T \in \mathbb{C}[S_n]$$

where $\mathrm{Tab}(\lambda)$ is the set of Young tableaux of shape λ filled with $[n]$, and with c_T the Young symmetrizer

$$c_T = \sum_{\sigma \in C(T)} \mathrm{sgn}(\sigma)\sigma \sum_{\pi \in R(T)} \pi$$

where $C(T), R(T) \subseteq S_n$ are the subgroups of permutations inside columns and permutations inside rows respectively. The element P_λ is a minimal central idempotent in $\mathbb{C}[S_n]$, and $\sum_{\lambda \vdash n} P_\lambda = e$.

Back to the decomposition of $(\mathbb{C}^d)^{\otimes n}$. We need a handle on the size of the components in the direct sum decomposition (6.1). For our application it is good to think of d as a constant and n as a large number. The number of summands in the direct sum decomposition (6.1) is upper bounded by a polynomial in n,

$$|\{\lambda \vdash_d n\}| \leq (n+1)^d,$$

i.e. there are only few summands compared to the total dimension d^n. There are the following well-known bounds on the dimensions of the irreducible representations $[\lambda]$ and $\mathbb{S}_\lambda(\mathbb{C}^d)$ that make up the summands,

$$\frac{n!}{\prod_{\ell=1}^d (\lambda_\ell + d - \ell)!} \leq \dim[\lambda] \leq \frac{n!}{\prod_{\ell=1}^d \lambda_\ell!} \tag{6.2}$$

$$\dim \mathbb{S}_\lambda(\mathbb{C}^d) \leq (n+1)^{d(d-1)/2}. \tag{6.3}$$

Let $p \in \mathbb{R}^n$ be a probability vector, i.e. $\sum_{i=1}^n p_i = 1$ and $p_i \geq 0$ for $i \in [n]$. Let $H(p)$ be the Shannon entropy of the probability vector p,

$$H(p) = \sum_{i=1}^n p_i \log_2 \frac{1}{p_i}.$$

For $\alpha \in [0,1]$, let $h(\alpha) = H((\alpha, 1-\alpha))$ be the binary entropy. For a partition $\lambda = (\lambda_1, \ldots, \lambda_\ell) \vdash n$, let $\overline{\lambda} := \lambda/n = (\lambda_1/n, \ldots, \lambda_\ell/n)$ be the probability vector obtained by normalising λ.

Let $\lambda \vdash n$. For $N \in \mathbb{N}$, let $N\lambda = (N\lambda_1, N\lambda_2, \ldots, N\lambda_\ell)$ be the stretched partition. We see that, asymptotically in the stretching factor N, the dimension of the irreducible S_{Nn}-representation $[N\lambda]$ behaves like a multinomial coefficient, and

$$2^{NnH(\overline{\lambda})-o(N)} \leq \dim[N\lambda] \leq 2^{NnH(\overline{\lambda})}. \tag{6.4}$$

6.3 Kronecker and Littlewood–Richardson coefficients $g_{\lambda\mu\nu}$, $c^{\lambda}_{\mu\nu}$

Let $\mu, \nu \vdash n$. Let $S_n \hookrightarrow S_n \times S_n : \pi \mapsto (\pi, \pi)$ be the diagonal embedding. Consider the decomposition of the tensor product $[\mu] \otimes [\nu]$ restricted along the diagonal embedding,

$$[\mu] \otimes [\nu] \downarrow^{S_n \times S_n}_{S_n} \cong \bigotimes_{\lambda \vdash n} \mathrm{Hom}_{S_n}([\lambda], [\mu] \otimes [\nu]) \otimes [\lambda].$$

Define the *Kronecker coefficient*

$$g_{\lambda\mu\nu} = \dim \mathrm{Hom}_{S_n}([\lambda], [\mu] \otimes [\nu]),$$

i.e. $g_{\lambda\mu\nu}$ is the multiplicity of $[\lambda]$ in $[\mu] \otimes [\nu]$.

Let $\lambda \vdash_{a+b}$. Let $\mathrm{GL}_a \times \mathrm{GL}_b \hookrightarrow \mathrm{GL}_{a+b} : (A, B) \mapsto A \oplus B$ be the block-diagonal embedding. Consider the decomposition of the representation $\mathbb{S}_\lambda(\mathbb{C}^{a+b})$ restricted along the block-diagonal embedding,

$$\mathbb{S}_\lambda(\mathbb{C}^{a+b}) \downarrow^{\mathrm{GL}_{a+b}}_{\mathrm{GL}_a \times \mathrm{GL}_b} \cong \bigoplus_{\substack{\mu \vdash_a \\ \nu \vdash_b}} H^{\lambda}_{\mu,\nu} \otimes \mathbb{S}_\mu(\mathbb{C}^a) \otimes \mathbb{S}_\nu(\mathbb{C}^b)$$

with

$$H^{\lambda}_{\mu,\nu} = \mathrm{Hom}_{\mathrm{GL}_a \times \mathrm{GL}_b}(\mathbb{S}_\mu(\mathbb{C}^a) \otimes \mathbb{S}_\nu(\mathbb{C}^b), \mathbb{S}_\lambda(\mathbb{C}^{a+b})).$$

Define the *Littlewood–Richardson coefficient* $c^{\lambda}_{\mu\nu} = \dim H^{\lambda}_{\mu,\nu}$.

For partitions $\lambda, \lambda' \vdash$ define $\lambda + \lambda'$ elementwise. The Kronecker and the Littlewood–Richardson coefficients have the following semigroup property (see e.g. [CHM07]).

Lemma 6.1. *Let $\lambda, \mu, \nu, \alpha, \beta, \gamma \vdash$ be partitions.*

(i) *If $g_{\lambda\mu\nu} > 0$ and $g_{\alpha\beta\gamma} > 0$, then $g_{\lambda+\alpha,\, \mu+\beta,\, \nu+\gamma} > 0$.*

(ii) *If $c^{\lambda}_{\mu\nu} > 0$ and $c^{\alpha}_{\beta\gamma} > 0$, then $c^{\lambda+\alpha}_{\mu+\beta,\, \nu+\gamma} > 0$.*

6.4 Entropy inequalities

The semigroup properties imply the following lemma. Of this lemma, the first statement can be found in a paper by Christandl and Mitchison [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.

Lemma 6.2. *Let $\lambda, \mu, \nu \vdash$ be partitions.*

(i) *If $g_{\lambda\mu\nu} > 0$, then $H(\overline{\lambda}) \leq H(\overline{\mu}) + H(\overline{\nu})$.*

(ii) *If $c^{\lambda}_{\mu\nu} > 0$, then $H(\overline{\lambda}) \leq \frac{|\mu|}{|\mu|+|\nu|} H(\overline{\mu}) + \frac{|\nu|}{|\mu|+|\nu|} H(\overline{\nu}) + h\big(\frac{|\mu|}{|\mu|+|\nu|}\big)$.*

Proof. (i) Let $g_{\lambda\mu\nu} > 0$. Suppose $\lambda, \mu, \nu \vdash n$. Let $N \in \mathbb{N}$. Then Lemma 6.1 implies $g_{N\lambda,N\mu,N\nu} > 0$. This means $\mathrm{Hom}_{S_{nN}}([N\lambda], [N\mu] \otimes [N\nu]) \neq 0$, which implies $\dim[N\lambda] \leq \dim[N\mu] \dim[N\nu]$. From (6.4) we have the dimension bounds

$$\begin{aligned} 2^{NnH(\overline{\lambda}) - o(N)} &\leq \dim[N\lambda] \\ \dim[N\mu] &\leq 2^{NnH(\overline{\mu})} \\ \dim[N\nu] &\leq 2^{NnH(\overline{\nu})}. \end{aligned}$$

Thus $NnH(\overline{\lambda}) - o(N) \leq NnH(\overline{\mu}) + NnH(\overline{\nu})$. Divide by Nn and let N go to infinity to get $H(\overline{\lambda}) \leq H(\overline{\mu}) + H(\overline{\nu})$.

(ii) We restrict the decomposition

$$(\mathbb{C}^{a+b})^{\otimes n} \cong \bigoplus_{\lambda \vdash_{a+b} n} [\lambda] \otimes \mathbb{S}_\lambda(\mathbb{C}^{a+b})$$

along the block-diagonal embedding to get

$$\begin{aligned} (\mathbb{C}^{a+b})^{\otimes n} \downarrow^{\mathrm{GL}_{a+b}}_{\mathrm{GL}_a \times \mathrm{GL}_b} &\cong \bigotimes_{\lambda \vdash_{a+b} n} [\lambda] \otimes \mathbb{S}_\lambda(\mathbb{C}^{a+b}) \downarrow^{\mathrm{GL}_{a+b}}_{\mathrm{GL}_a \times \mathrm{GL}_b} \\ &\cong \bigoplus_{\lambda \vdash_{a+b} n} [\lambda] \otimes \bigoplus_{\substack{\mu \vdash_a \\ \nu \vdash_b}} \mathbb{C}^{c^\lambda_{\mu\nu}} \otimes \mathbb{S}_\mu(\mathbb{C}^a) \otimes \mathbb{S}_\nu(\mathbb{C}^b) \\ &\cong \bigoplus_{\substack{\mu \vdash_a \\ \nu \vdash_b}} \Big(\bigoplus_{\lambda \vdash_{a+b} n} [\lambda] \otimes \mathbb{C}^{c^\lambda_{\mu\nu}} \Big) \otimes \mathbb{S}_\mu(\mathbb{C}^a) \otimes \mathbb{S}_\nu(\mathbb{C}^b). \end{aligned}$$

On the other hand,

$$\begin{aligned} (\mathbb{C}^{a+b})^{\otimes n} \downarrow &\cong (\mathbb{C}^a \oplus \mathbb{C}^b)^{\otimes n} \downarrow \cong (\mathbb{C}^a)^{\otimes n} \oplus ((\mathbb{C}^a)^{\otimes n-1} \otimes \mathbb{C}^b) \oplus \cdots \oplus (\mathbb{C}^b)^{\otimes n} \downarrow \\ &\cong \bigoplus_{k=0}^{n} \mathbb{C}^{\binom{n}{k}} \otimes \bigoplus_{\mu \vdash_a k} ([\mu] \otimes \mathbb{S}_\mu(\mathbb{C}^a)) \otimes \bigoplus_{\nu \vdash_b n-k} ([\nu] \otimes \mathbb{S}_\nu(\mathbb{C}^b)) \end{aligned}$$

$$\cong \bigoplus_{k=0}^{n} \bigoplus_{\substack{\mu \vdash_a k \\ \nu \vdash_b n-k}} \big(\mathbb{C}^{\binom{n}{k}} \otimes [\mu] \otimes [\nu]\big) \otimes \mathbb{S}_\mu(\mathbb{C}^a) \otimes \mathbb{S}_\nu(\mathbb{C}^b).$$

Suppose $c^\lambda_{\mu\nu} > 0$. Comparing the above expressions gives the inequality $\dim[\lambda] \leq \binom{n}{|\mu|} \dim[\mu] \dim[\nu]$. By the semigroup property Lemma 6.1, we have $c^{N\lambda}_{N\mu,N\nu} > 0$ for all $N \in \mathbb{N}$. Thus $\dim[N\lambda] \leq \binom{Nn}{N|\mu|} \dim[N\mu] \dim[N\nu]$ for all $N \in \mathbb{N}$. Then from (6.4) follows

$$2^{NnH(\overline{\lambda}) - o(N)} \leq 2^{Nnh(\frac{|\mu|}{n})} 2^{N|\mu|H(\overline{\mu})} 2^{N|\nu|H(\overline{\nu})}.$$

We conclude $H(\overline{\lambda}) \leq h(\frac{|\mu|}{n}) + \frac{|\mu|}{n} H(\overline{\mu}) + \frac{|\nu|}{n} H(\overline{\nu})$. □

Let $x = (x^{(1)}, x^{(2)}, x^{(3)})$ be a triple of probability vectors $x^{(i)} \in \mathbb{R}^{n_i}$. Let $\theta \in \Theta \coloneqq \mathcal{P}([3])$. Let $H_\theta(x)$ be the θ-weighted average of the Shannon entropies of the probability vectors $x^{(1)}$, $x^{(2)}$ and $x^{(3)}$,

$$H_\theta(x) \coloneqq \theta(1) H(x^{(1)}) + \theta(2) H(x^{(2)}) + \theta(3) H(x^{(3)}).$$

(Note that this notation is slightly different from the notation used in Chapter 4.) We will use the notation $\lambda \vdash^3 n$ to say that λ is a triple of partitions of n, i.e. λ equals $(\lambda^{(1)}, \lambda^{(2)}, \lambda^{(3)})$ where each $\lambda^{(i)}$ is a partition of n. We write $\overline{\lambda}$ for the component-wise normalised triple $(\overline{\lambda^{(1)}}, \overline{\lambda^{(2)}}, \overline{\lambda^{(3)}})$.

Lemma 6.3. *Let $\lambda, \mu, \nu \vdash^3$ be three triples of partitions.*

(i) *If $g_{\lambda^{(i)}\mu^{(i)}\nu^{(i)}} > 0$ for all i, then $2^{H_\theta(\overline{\lambda})} \leq 2^{H_\theta(\overline{\mu})} 2^{H_\theta(\overline{\nu})}$.*

(ii) *If $\mu \vdash^3 m, \nu \vdash^3 n - m$ and $c^{\lambda^{(i)}}_{\mu^{(i)}\nu^{(i)}} > 0$ for all i, then $2^{H_\theta(\overline{\lambda})} \leq 2^{H_\theta(\overline{\mu})} + 2^{H_\theta(\overline{\nu})}$.*

Proof. (i) Suppose $g_{\lambda^{(i)}\mu^{(i)}\nu^{(i)}} > 0$ for all i. Then $H(\overline{\lambda^{(i)}}) \leq H(\overline{\mu^{(i)}}) + H(\overline{\nu^{(i)}})$ for all i by Lemma 6.2. Thus $\sum_i \theta(i) H(\overline{\lambda^{(i)}}) \leq \sum_i \theta(i) H(\overline{\mu^{(i)}}) + \sum_i \theta(i) H(\overline{\nu^{(i)}})$. Then $H_\theta(\overline{\lambda}) \leq H_\theta(\overline{\mu}) + H_\theta(\overline{\nu})$. We conclude $2^{H_\theta(\overline{\lambda})} \leq 2^{H_\theta(\overline{\mu})} 2^{H_\theta(\overline{\nu})}$.

(ii) Suppose $c^{\lambda^{(i)}}_{\mu^{(i)}\nu^{(i)}} > 0$ for all i. Then $H(\overline{\lambda^{(i)}}) \leq \frac{m}{n} H(\overline{\mu^{(i)}}) + \frac{n-m}{n} H(\overline{\nu^{(i)}}) + h(\frac{m}{n})$ by Lemma 6.2. We take the θ-weighted average to get $H_\theta(\overline{\lambda}) \leq \frac{m}{n} H_\theta(\overline{\mu}) + \frac{n-m}{m} H_\theta(\overline{\nu}) + h(\frac{m}{n})$. We conclude $2^{H_\theta(\overline{\lambda})} \leq 2^{H_\theta(\overline{\mu})} + 2^{H_\theta(\overline{\nu})}$ by Lemma 4.9(iv). □

6.5 Hilbert spaces and density operators

Endow the vector space $\mathbb{C}^n$ with a hermitian inner product (one may take the standard hermitian inner product $\langle u, v \rangle = \sum_{i=1}^n \overline{u}_i v_i$ for $u, v \in \mathbb{C}^n$ where $\overline{\cdot}$ denotes taking the complex conjugate), so that it is a Hilbert space.

Let $(V_1, \langle\cdot,\cdot\rangle)$ and $(V_2, \langle\cdot,\cdot\rangle)$ be Hilbert spaces. On $V_1 \oplus V_2$ we define the inner product by $\langle u_1 \oplus u_2, v_1 \oplus v_2\rangle = \langle u_1, v_1\rangle + \langle u_2, v_3\rangle$. On $V_1 \otimes V_2$ we define the inner product by $\langle u_1 \otimes u_2, v_1 \otimes v_2\rangle = \langle u_1, v_1\rangle\langle u_2, v_2\rangle$ and extending linearly.

Let V be a Hilbert space. A positive semidefinite hermitian operator $\rho : V \to V$ with trace one is called a *density operator*. The sequence of eigenvalues of a density operator ρ is a probability vector. Let $\operatorname{spec}(\rho) = (p_1, \ldots, p_n)$ be the sequence of eigenvalues of ρ, ordered non-increasingly, $p_1 \geq \cdots \geq p_n$.

Let V_1 and V_2 be Hilbert spaces. Given a density operator ρ on $V_1 \otimes V_2$, the *reduced* density operator $\rho_1 = \operatorname{tr}_2 \rho$ is uniquely defined by the property that $\operatorname{tr}(\rho_1 X_1) = \operatorname{tr}(\rho(X_1 \otimes \mathrm{Id}_{V_2}))$ for all operators X_1 on V_1. The operator ρ_1 is again a density operator. The operation tr_2 is called the partial trace over V_2. Explicitly, ρ_1 is given by $\langle e_i, \rho_1(e_j)\rangle = \sum_\ell \langle e_i \otimes f_\ell, \rho(e_j \otimes f_\ell)\rangle$, where the e_i are some basis of V_1 and the f_i are some basis of V_2 (the statement is independent of basis choice).

Let V_i be a Hilbert space and consider the tensor product $V_1 \otimes V_2 \otimes V_3$. Associate with $t \in V_1 \otimes V_2 \otimes V_3$ the dual element $t^* := \langle t, \cdot\rangle \in (V_1 \otimes V_2 \otimes V_3)^*$. Then

$$\rho^t := tt^*/\langle t, t\rangle = t\langle t, \cdot\rangle/\langle t, t\rangle$$

is a density operator on $V_1 \otimes V_2 \otimes V_3$. Viewing ρ^t as a density operator on the regrouped space $V_1 \otimes (V_2 \otimes V_3)$ we may take the partial trace of ρ^t over $V_2 \otimes V_3$ as described above. We denote the resulting density operator by $\rho^t_1 = \operatorname{tr}_{23} \rho^t$. We similarly define $\rho^t_2 = \operatorname{tr}_{13} \rho^t$ and $\rho^t_3 = \operatorname{tr}_{12} \rho^t$.

6.6 Moment polytopes $\mathbf{P}(t)$

We give a brief introduction to moment polytopes. We refer to [Nes84, Bri87, Fra02, Wal14] for more information. We begin with the general setting and then specialise to orbit closures in tensor spaces.

6.6.1 General setting

Let G be a connected reductive algebraic group. (We refer to Kraft [Kra84] and Humphreys [Hum75] for an introduction to algebraic groups.) Fix a maximal torus $T \subseteq G$ and a Borel subgroup $T \subseteq B \subseteq G$. We have the character group $X(T)$, the Weyl group W, the root system $\Phi \subseteq X(T)$ and the system of positive roots $\Phi_+ \subseteq \Phi$. For $\lambda, \mu \in X(T)$, we set $\lambda \preccurlyeq \mu$ if $\mu - \lambda$ is a sum of positive roots. Let V be a rational G-representation. The restriction of the action of G to T gives a decomposition

$$V = \bigoplus_{\lambda \in X(T)} V_\lambda, \quad V_\lambda = \{v \in V : \forall t \in T\ t \cdot v = \lambda(t) v\}.$$

This decomposition is called the *weight decomposition* of V. The $\lambda \in X(T)$ with $V_\lambda \neq 0$ are called the *weights* of V with respect to T. The V_λ are the *weight spaces* of V. For $v \in V$, let v_λ be the component of v in V_λ. Let $\mathrm{supp}(v) \coloneqq \{\lambda : v_\lambda \neq 0\}$.

Let E be the real vector space $E \coloneqq X(T) \otimes \mathbb{R}$. The Weyl group W acts on $X(T)$ and thus on E. We enlarge $\preccurlyeq$ to a partial order on E as follows. For $x, y \in E$ let $x \preccurlyeq y$ if $y - x$ is a nonnegative linear combination of positive roots. Let $D \subseteq E$ be the positive Weyl chamber. For every $x \in E$ the orbit $W \cdot x$ intersects the positive Weyl chamber D in exactly one point, which we denote by $\mathrm{dom}(x)$.

Let V be a finite-dimensional rational G-module. Let $\chi \in X(T) \cap D$ be a dominant character. We denote the χ-isotypical component of V with $V_{(\chi)}$. Let $Z \subseteq V$ be a Zariski closed set. We denote the coordinate ring of Z with $\mathbb{C}[Z]$. We denote the degree d part of $\mathbb{C}[Z]$ with $\mathbb{C}[Z]_d$. If Z is G-stable, then $\mathbb{C}[Z]_d$ is a G-module.

Definition 6.4. Let V be a rational G-module and $Z \subseteq V$ a nontrivial irreducible closed G-stable cone. The *moment polytope* of Z, denoted by

$$\mathbf{P}(Z),$$

is defined as the Euclidean closure in E of the set

$$\mathbf{R}(Z) = \{\chi/d : (\mathbb{C}[Z]_d)_{(\chi^*)} \neq 0\}$$

of normalised characters χ/d for which the χ^*-isotypical component $(\mathbb{C}[Z]_d)_{(\chi^*)}$ is not zero.

Theorem 6.5 (Mumford–Ness [Nes84], Brion [Bri87], Franz [Fra02])**.** *The moment polytope is indeed a convex polytope and it is equal to the image of the so-called moment map intersected with the positive Weyl chamber,*

$$\mathbf{P}(Z) = \mu(Z \setminus 0) \cap D.$$

Let $Z = \overline{G \cdot v}$ be the orbit closure (in the Zariski topology) of a vector $v \in V \setminus 0$ and suppose $\overline{G \cdot v}$ is a cone.

Lemma 6.6 (See e.g. [Str05])**.** *Suppose $\overline{G \cdot v}$ is a cone. Then*

$$\begin{aligned}\mathbf{R}(\overline{G \cdot v}) &= \{\chi/d : (\mathbb{C}[\overline{G \cdot v}]_d)_{(\chi^*)} \neq 0\} \\ &= \{\chi/d : (\mathrm{lin}(G \cdot v^{\otimes d}))_{(\chi)} \neq 0\}.\end{aligned}$$

6.6.2 Tensor spaces

We specialise to 3-tensors. Let $V = V_1 \otimes V_2 \otimes V_3$ with $V_i = \mathbb{C}^{n_i}$. Let

$$G = \mathrm{GL}_{n_1} \times \mathrm{GL}_{n_2} \times \mathrm{GL}_{n_3}$$

$$T = T_1 \times T_2 \times T_3$$

with T_i the diagonal matrices in GL_{n_i}. The weight decomposition of V is the decomposition with respect to the standard basis elements $e_{x_1} \otimes e_{x_2} \otimes e_{x_3}$ where $x \in [n_1] \times [n_2] \times [n_3]$. The support $\mathrm{supp}(v)$ is the support of v with respect to the standard basis.

In the current setting there is a beautiful rephrasing of Theorem 6.5 in terms of ordered spectra of reduced density matrices. Recall from Section 6.5 that for $v \in V \setminus 0$ we have a density matrix ρ^v, and reduced density matrices ρ_i^v, of which we may take the non-increasingly ordered spectra, $\mathrm{spec}(\rho_i^v)$.

Theorem 6.7 (Walter–Doran–Gross–Christandl [WDGC13])**.** *Let $Z \subseteq V$ be a nontrivial irreducible closed G-stable cone. Then*

$$\mathbf{P}(Z) = \{(\mathrm{spec}\,\rho_1^z, \mathrm{spec}\,\rho_2^z, \mathrm{spec}\,\rho_3^z) : z \in Z \setminus 0\}.$$

Let $v \in V \setminus 0$. We consider the moment polytope of the orbit closure $Z = \overline{G \cdot v}$. In this setting Lemma 6.6 specialises to the following.

Lemma 6.8 (See e.g. [Str05])**.**

$$\begin{aligned}\mathbf{R}(\overline{G \cdot v}) &= \{\chi/d : (\mathbb{C}[\overline{G \cdot v}]_d)_{(\chi^*)} \neq 0,\, d \in \mathbb{N}_{\geq 1}\}\\ &= \{\chi/d : (\mathrm{lin}(G \cdot v^{\otimes d}))_{(\chi)} \neq 0,\, d \in \mathbb{N}_{\geq 1}\}\\ &= \{\chi/d : P_\chi v^{\otimes d} \neq 0,\, d \in \mathbb{N}_{\geq 1}\},\end{aligned}$$

where $P_\chi = P_{\chi^{(1)}} \otimes P_{\chi^{(2)}} \otimes P_{\chi^{(3)}}$ with $P_{\chi^{(i)}} : V_i^{\otimes d} \to V_i^{\otimes d}$ the projector onto the isotypical component of type $\chi^{(i)}$ discussed in Section 6.2.

On the other hand, Theorem 6.7 immediately gives a description of the moment polytope $\mathbf{P}(\overline{G \cdot v})$ in terms of ordered spectra of reduced density matrices.

Theorem 6.9. *Let $v \in V \setminus 0$. Then*

$$\mathbf{P}(\overline{G \cdot v}) = \{(\mathrm{spec}\,\rho_1^u, \mathrm{spec}\,\rho_2^u, \mathrm{spec}\,\rho_3^u) : u \in \overline{G \cdot v} \setminus 0\}.$$

Summarising, we have two descriptions of the moment polytope, a representation-theoretic or invariant-theoretic description (Lemma 6.8) and a quantum marginal spectra description (Theorem 6.9). These two descriptions are the key to proving the properties of the quantum functionals that we need.

6.7 Quantum functionals $F^\theta(t)$

We will now define the quantum functionals and prove that they are universal spectral points.

Let $p = (p_1, \ldots, p_n) \in \mathbb{R}^n$ be a probability vector, i.e. $\sum_{i=1}^n p_i = 1$ and $p_i \geq 0$ for all $i \in [n]$. Recall that $H(p)$ denotes the Shannon entropy of the probability vector p, $H(p) \coloneqq \sum_{i=1}^n p_i \log_2 1/p_i$. Let $x = (x^{(1)}, x^{(2)}, x^{(3)})$ be a triple of probability vectors $x^{(i)} \in \mathbb{R}^{n_i}$. Let $\theta \in \Theta \coloneqq \mathcal{P}([3])$. Recall that $H_\theta(x)$ denotes the θ-weighted average of the Shannon entropies of the three probability vectors $x^{(1)}$, $x^{(2)}$, $x^{(3)}$,

$$H_\theta(x) \coloneqq \theta(1)H(x^{(1)}) + \theta(2)H(x^{(2)}) + \theta(3)H(x^{(3)}).$$

Let $V = \mathbb{C}^{n_1} \otimes \mathbb{C}^{n_2} \otimes \mathbb{C}^{n_3}$. Let $G = \mathrm{GL}_{n_1} \times \mathrm{GL}_{n_2} \times \mathrm{GL}_{n_3}$. Let $v \in V \setminus 0$. We use the notation $\mathbf{P}(v) \coloneqq \mathbf{P}(\overline{G \cdot v})$ for the moment polytope of the orbit closure of v.

Definition 6.10. For $\theta \in \Theta$ and $v \in V \setminus 0$ let

$$F^\theta(v) \coloneqq \max\{2^{H_\theta(x)} : x \in \mathbf{P}(v)\}.$$

Let $F^\theta(0) = 0$. We call the functions F^θ the *quantum functionals*. The name quantum functional comes from the fact that the moment polytope $\mathbf{P}(t)$ consists of triples of quantum marginal entropies.

Theorem 6.11. *Let $\mathcal{T}$ be the semiring of 3-tensors over $\mathbb{C}$. Let $\leqslant$ be the restriction preorder. For $\theta \in \Theta$*

$$F^\theta \in \mathbf{X}(\mathcal{T}, \leqslant).$$

In other words, F^θ is a semiring homomorphism $\mathcal{T} \to \mathbb{R}_{\geq 0}$ which is monotone under restriction $\leqslant$. In fact, F^θ is monotone under degeneration $\trianglelefteq$.

Remark 6.12. The results in this chapter generalise to k-tensors over $\mathbb{C}$. In our paper [CVZ18] we discuss this general situation in detail and make a distinction between upper quantum functionals and lower quantum functionals.

Let $p \in \mathbb{R}^n$ and $q \in \mathbb{R}^m$ be vectors. We define the tensor product $p \otimes q \in \mathbb{R}^{nm}$ by

$$p \otimes q \coloneqq (p_i q_j : i \in [n], j \in [m]).$$

We define the direct sum $p \oplus q \in \mathbb{R}^{n+m}$ by

$$p \oplus q \coloneqq (p_1, \ldots, p_n, q_1, \ldots, q_m)$$

Let $\mathrm{dom}(p)$ be the vector obtained from p by reordering the coefficients non-increasingly.

Let $x = (x^{(1)}, x^{(2)}, x^{(3)})$ and $y = (y^{(1)}, y^{(2)}, y^{(3)})$ be triples of vectors. We define the tensor product $x \otimes y$ elementwise,

$$x \otimes y \coloneqq (x^{(1)} \otimes y^{(1)}, x^{(2)} \otimes y^{(2)}, x^{(3)} \otimes y^{(3)}).$$

We define the direct sum $x \oplus y$ elementwise,

$$x \oplus y \coloneqq (x^{(1)} \oplus y^{(1)},\, x^{(2)} \oplus y^{(2)},\, x^{(3)} \oplus y^{(3)}).$$

For a triple of vectors $x = (x^{(1)}, x^{(2)}, x^{(3)})$ let

$$\operatorname{dom}(x) \coloneqq (\operatorname{dom}(x^{(1)}),\, \operatorname{dom}(x^{(2)}),\, \operatorname{dom}(x^{(3)}))$$

be the triple of vectors obtained from x be reordering each component $x^{(i)}$ non-increasing. For any set S of triples of vectors, let

$$\operatorname{dom}(S) \coloneqq \{\operatorname{dom}(x) : x \in S\}.$$

For $v \in \mathbb{C}^{n_1} \otimes \mathbb{C}^{n_2} \otimes \mathbb{C}^{n_3}$ we will use the notation $G(v) \coloneqq \mathrm{GL}_{n_1} \times \mathrm{GL}_{n_2} \times \mathrm{GL}_{n_3}$ to denote the product of general linear groups that naturally corresponds to the space that v lives in. We will use the notation $\mathbf{P}(v) \coloneqq \mathbf{P}(\overline{G(v) \cdot v})$ for the moment polytope of the orbit closure of v.

Theorem 6.13. *Let $s \in \mathbb{C}^{n_1} \otimes \mathbb{C}^{n_2} \otimes \mathbb{C}^{n_3}$ and $t \in \mathbb{C}^{m_1} \otimes \mathbb{C}^{m_2} \otimes \mathbb{C}^{m_3}$.*

(i) $\operatorname{dom}\big(\mathbf{P}(s) \otimes \mathbf{P}(t)\big) \subseteq \mathbf{P}(s \otimes t)$

(ii) $\forall \alpha \in [0,1]\ \ \operatorname{dom}\big(\alpha\, \mathbf{P}(s) \oplus (1-\alpha)\, \mathbf{P}(t)\big) \subseteq \mathbf{P}(s \oplus t)$

(iii) *If $s, t \in \mathbb{C}^{n_1} \otimes \mathbb{C}^{n_2} \otimes \mathbb{C}^{n_3} \setminus 0$ and $s \in \overline{G(t) \cdot t}$, then $\mathbf{P}(s) \subseteq \mathbf{P}(t)$*

(iv) $\mathbf{P}(s \oplus 0) = \mathbf{P}(s) \oplus 0$

(v) $\mathbf{P}(\langle 1 \rangle) = \{((1),(1),(1))\}$ *with* $\langle 1 \rangle = e_1 \otimes e_1 \otimes e_1 \in \mathbb{C}^1 \otimes \mathbb{C}^1 \otimes \mathbb{C}^1$.

Proof. To prove statements (i) and (ii), let $x \in \mathbf{P}(s)$ and $y \in \mathbf{P}(t)$. Then there are elements $a \in \overline{G(s) \cdot s}$ and $b \in \overline{G(t) \cdot t}$ with ordered marginal spectra x and y,

$$\begin{aligned} x &= (\operatorname{spec} \rho_1^a,\, \operatorname{spec} \rho_2^a,\, \operatorname{spec} \rho_3^a) \\ y &= (\operatorname{spec} \rho_1^b,\, \operatorname{spec} \rho_2^b,\, \operatorname{spec} \rho_3^b). \end{aligned}$$

We prove statement (i). We have $a \otimes b \in \overline{G(s \otimes t) \cdot s \otimes t}$. Thus

$$\operatorname{dom}(x \otimes y) = (\operatorname{spec} \rho_1^{a\otimes b},\, \operatorname{spec} \rho_2^{a\otimes b},\, \operatorname{spec} \rho_3^{a\otimes b}) \in \mathbf{P}(s \otimes t).$$

We conclude $\operatorname{dom}(\mathbf{P}(s) \otimes \mathbf{P}(t)) \subseteq \mathbf{P}(s \otimes t)$. We prove statement (ii). Let $\alpha \in [0,1]$. Define the tensor $u(\alpha) \in \mathbb{C}^{n_1+m_1} \otimes \mathbb{C}^{n_2+m_2} \otimes \mathbb{C}^{n_3+m_3}$ by

$$u(\alpha) \coloneqq \frac{\sqrt{\alpha}}{\sqrt{\langle s, s\rangle}}\, a \oplus \frac{\sqrt{1-\alpha}}{\sqrt{\langle t, t\rangle}}\, b.$$

Then $u(\alpha) \in \overline{G(s \oplus t) \cdot s \oplus t}$. We have $\rho_i^{u(\alpha)} = \alpha\rho_i^a \oplus (1-\alpha)\rho_i^b$. From the observation

$$\mathrm{spec}(\alpha\rho_i^a \oplus (1-\alpha)\rho_i^b) = \mathrm{dom}(\alpha x \oplus (1-\alpha)y)$$

follows $\mathrm{dom}(\alpha x \oplus (1-\alpha)y) \in \mathbf{P}(\overline{G(s \oplus t) \cdot s \oplus t})$. We conclude

$$\mathrm{dom}(\alpha\mathbf{P}(s) \oplus (1-\alpha)\mathbf{P}(t)) \subseteq \mathbf{P}(s \oplus t).$$

We have thus proven statement (i) and (ii).

We prove statement (iii). Let $G = G(t) = G(s)$. Let $s \in \overline{G \cdot t}$. Then $\overline{G \cdot s} \subseteq \overline{G \cdot t}$, so we have a G-equivariant restriction map $\mathbb{C}[\overline{G \cdot s}] \twoheadleftarrow \mathbb{C}[\overline{G \cdot t}]$ on the coordinate rings. Let $\chi/d \in R(\overline{G \cdot s})$ with $(\mathbb{C}[\overline{G \cdot s}]_d)_{(\chi^*)} \neq 0$. Then also $(\mathbb{C}[\overline{G \cdot t}]_d)_{(\chi^*)} \neq 0$ by Schur's lemma. Thus $\chi/d \in \mathbf{R}(\overline{G \cdot t}) \subseteq \mathbf{P}(\overline{G \cdot t})$. We conclude $\mathbf{P}(s) \subseteq \mathbf{P}(t)$.

We prove statement (iv). Let $\chi/d \in \mathbf{R}(\overline{G(s \oplus 0) \cdot (s \oplus 0)})$ with $P_\chi(s \oplus 0)^{\otimes d} \neq 0$. Recall from Section 6.2 that P_χ is given by the action of an element in the group algebra $\mathbb{C}[S_d]$ which we also denoted by P_χ. From this viewpoint we see that also $P_\chi s^{\otimes d} \neq 0$. So $\chi/d \in \mathbf{R}(\overline{G(s) \cdot s})$.

Statement (v) is a direct observation. □

Corollary 6.14.

(i) $F^\theta(s)F^\theta(t) \leq F^\theta(s \otimes t)$

(ii) $F^\theta(s) + F^\theta(t) \leq F^\theta(s \oplus t)$

(iii) *If* $s \trianglelefteq t$, *then* $F^\theta(s) \leq F^\theta(t)$

(iv) $F^\theta(\langle 1 \rangle) = 1$

Proof. (i) Let $x \in \mathbf{P}(s)$ and $y \in \mathbf{P}(t)$. Then $x \otimes y \in \mathbf{P}(s \otimes t)$ by Theorem 6.13. It is a basic fact that $H_\theta(x) + H_\theta(y) = H_\theta(x \otimes y)$ (Lemma 4.9), so $2^{H_\theta(x)}2^{H_\theta(y)} = 2^{H_\theta(x \otimes y)}$. We conclude $F^\theta(s)F^\theta(t) \leq F^\theta(s \otimes t)$.

(ii) Let $x \in \mathbf{P}(s)$ and $y \in \mathbf{P}(t)$. Then by Theorem 6.13 for all $\alpha \in [0,1]$

$$\mathrm{dom}(\alpha x \oplus (1-\alpha)y) \in \mathbf{P}(s \oplus t).$$

It is a basic fact that $\alpha H_\theta(x) + (1-\alpha)H_\theta(y) + h(\alpha) = H_\theta(\alpha x \oplus (1-\alpha)y)$ (Lemma 4.9). Thus for any $\alpha \in [0,1]$ we have $2^{\alpha H_\theta(x) + (1-\alpha)H_\theta(y) + h(\alpha)} \leq F^\theta(s \oplus t)$. Using Lemma 4.9(iv) we conclude $F^\theta(s) + F^\theta(t) \leq F^\theta(s \oplus t)$.

(iii) This follows from statement (iii) and (iv) of Theorem 6.13, since, by definition, degeneration $s \trianglelefteq t$ means $s \oplus 0 \in \overline{G(t \oplus 0) \cdot (t \oplus 0)}$.

(iv) This follows from statement (v) of Theorem 6.13. □

Theorem 6.15.

(i) $\mathbf{R}(s \otimes t) \subseteq \{\lambda/N : \exists \mu/N \in \mathbf{R}(s), \nu/N \in \mathbf{R}(t);\ g_{\lambda^{(i)}\mu^{(i)}\nu^{(i)}} > 0 \text{ for all } i\}$

(ii) $\mathbf{R}(s \oplus t) \subseteq \{\lambda/N : \exists \mu/m \in \mathbf{R}(s), \nu/(N-m) \in \mathbf{R}(t);\ c^{\lambda^{(i)}}_{\mu^{(i)}\nu^{(i)}} > 0 \text{ for all } i\}$

Proof. (i) Let $s \in V_1 \otimes V_2 \otimes V_3$ and let $t \in W_1 \otimes W_2 \otimes W_3$. Let $\lambda/N \in \mathbf{R}(s \otimes t)$ with $P_\lambda (s \otimes t)^{\otimes N} \neq 0$. Let π be the natural regrouping map

$$\pi : ((V_1 \otimes W_1) \otimes (V_2 \otimes W_2) \otimes (V_3 \otimes W_3))^{\otimes N} \\ \to (V_1 \otimes V_2 \otimes V_3)^{\otimes N} \otimes (W_1 \otimes W_2 \otimes W_3)^{\otimes N}.$$

Then

$$(s \otimes t)^{\otimes N} = \sum_{\mu,\nu} \pi^{-1}(P_\mu \otimes P_\nu)\pi(s \otimes t)^{\otimes N}.$$

Let $\mu, \nu \vdash^3 N$ with $P_\lambda \pi^{-1}(P_\mu \otimes P_\nu)\pi(s \otimes t)^{\otimes N} \neq 0$. Then $P_\mu s^{\otimes N} \neq 0$ and $P_\nu t^{\otimes N} \neq 0$, i.e. $\mu/N \in \mathbf{R}(s)$ and $\nu/N \in \mathbf{R}(t)$. Moreover $P_\lambda \pi^{-1}(P_\mu \otimes P_\nu)\pi \neq 0$, which means the Kronecker coefficients $g_{\lambda^{(i)}\mu^{(i)}\nu^{(i)}}$ are nonzero.

(ii) Let $\lambda/N \in \mathbf{R}(s \oplus t)$ with $P_\lambda (s \oplus t)^{\otimes N} \neq 0$. Let us expand $(s \oplus t)^{\otimes N}$ as

$$(s \oplus t)^{\otimes N} = s^{\otimes N} \oplus (s^{\otimes N-1} \otimes t) \oplus \cdots \oplus t^{\otimes N}.$$

Then P_λ does not vanish on some summand, which we may assume to be of the form $s^{\otimes m} \otimes t^{\otimes N-m}$. Let π be the natural projection

$$\pi : ((V_1 \oplus W_1) \otimes (V_2 \oplus W_2) \otimes (V_3 \oplus W_3))^{\otimes N} \\ \to (V_1 \otimes V_2 \otimes V_3)^{\otimes m} \otimes (W_1 \otimes W_2 \otimes W_3)^{\otimes N-m}.$$

Let μ, ν with $P_\lambda \pi^{-1}(P_\mu \otimes P_\nu)\pi(s \oplus t)^{\otimes N} \neq 0$. Then $P_\mu s^{\otimes m} \neq 0$ and $P_\nu t^{\otimes N-m} \neq 0$. Moreover $P_\lambda \pi^{-1}(P_\mu \otimes P_\nu)\pi \neq 0$. Therefore, the Littlewood–Richardson coefficients $c^{\lambda^{(i)}}_{\mu^{(i)}\nu^{(i)}}$ are nonzero. □

Corollary 6.16.

(i) $F^\theta(s \otimes t) \leq F^\theta(s) F^\theta(t)$

(ii) $F^\theta(s \oplus t) \leq F^\theta(s) + F^\theta(t)$

Proof. (i) Let $\lambda/N \in \mathbf{R}(s \otimes t)$. By Theorem 6.15 there is a $\mu/N \in \mathbf{R}(s)$ and a $\nu/N \in \mathbf{R}(t)$ such that the Kronecker coefficient $g_{\lambda^{(i)}\mu^{(i)}\nu^{(i)}}$ is nonzero for every i. Then $2^{H_\theta(\overline{\mu})} \leq F^\theta(s)$ and $2^{H_\theta(\overline{\nu})} \leq F^\theta(t)$ by definition of F^θ. The Kronecker coefficients being nonzero implies

$$2^{H_\theta(\overline{\lambda})} \leq 2^{H_\theta(\overline{\mu})} 2^{H_\theta(\overline{\nu})}$$

by Lemma 6.3. We conclude $F^\theta(s \otimes t) \leq F^\theta(s) F^\theta(t)$.

(ii) Let $\lambda/N \in \mathbf{R}(s \oplus t)$. Then by Theorem 6.15 there are $\mu/m \in \mathbf{R}(s)$ and $\nu/(N-m) \in \mathbf{R}(t)$ such that the Littlewood–Richardson coefficient $c^{\lambda^{(i)}}_{\mu^{(i)}\nu^{(i)}}$ is nonzero for every i. This means

$$2^{H_\theta(\overline{\lambda})} \leq 2^{H_\theta(\overline{\mu})} + 2^{H_\theta(\overline{\nu})}$$

by Lemma 6.3. We conclude $F^\theta(s \oplus t) \leq F^\theta(s) + F^\theta(t)$. □

Proof of Theorem 6.11. Corollary 6.14 and Corollary 6.16 together prove Theorem 6.11. □

6.8 Outer approximation

In this section we discuss an outer approximation of $\mathbf{P}(t)$. We will use this outer approximation to show that the quantum functionals are at most the support functionals.

Let $\preccurlyeq$ be the dominance order i.e. majorization order on triples of probability vectors. For any set $S \subseteq \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \times \mathbb{R}^{n_3}$ of triples of probability vectors, let $S^{\preccurlyeq}$ denote the upward closure with respect to $\preccurlyeq$,

$$S^{\preccurlyeq} = \{y \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \times \mathbb{R}^{n_3} : \exists x \in S,\, x \preccurlyeq y\}.$$

Let $\mathrm{conv}(S)$ denote the convex hull of S in $\mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \times \mathbb{R}^{n_3}$. Recall that for $x \in S$ we defined $\mathrm{dom}(x)$ as the triple of probability vectors obtained from $x = (x^{(1)}, x^{(2)}, x^{(3)})$ by reordering the components $x^{(i)}$ such that they become non-increasing, and $\mathrm{dom}(S) = \{\mathrm{dom}(x) : x \in S\}$.

Theorem 6.17 (Strassen [Str05])**.** *Let $v \in V \setminus 0$. Then*

$$\mathbf{P}(v) \subseteq (\mathrm{dom}\,\mathrm{conv}\,\mathrm{supp}\, v)^{\preccurlyeq}. \tag{6.5}$$

Proof. We give the proof for the convenience of the reader. Let $\chi/d \in \mathbf{R}(\overline{G \cdot v})$. Then $(\mathrm{lin}(G \cdot v^{\otimes d}))_{(\chi)} \neq 0$. Let $M_\chi \subseteq \mathrm{lin}(G \cdot v^{\otimes d})$ be a simple G-submodule with highest weight χ. Let $N \subseteq V^{\otimes d}$ be the G-module complement, $N \oplus M_\chi = V^{\otimes d}$. Then $v^{\otimes d}$ is not in N. Let $v = \bigoplus_{\gamma \in \mathrm{supp}\, v} v_\gamma$ be the weight decomposition. Then $v^{\otimes d}$ is a sum of tensor products of the v_γ. At least one summand is not in N, say of weight $\eta \coloneqq \sum_\gamma d_\gamma \gamma$ with $\sum_\gamma d_\gamma = d$. The projection $V^{\otimes d} \to M_\chi$ along N maps this summand onto a nonzero weight vector of weight η. So η is a weight of M_χ. Then also $\mathrm{dom}(\eta)$ is a weight of M_χ. Since χ is the highest weight of M_χ, $\mathrm{dom}(\eta) \preccurlyeq \chi$. Then $\mathrm{dom}(\eta/d) \preccurlyeq \chi/d$. We have $\eta/d = \sum_\gamma \frac{d_\gamma}{d} \gamma \in \mathrm{conv}\,\mathrm{supp}\, v$. We conclude $\mathbf{R}(\overline{G \cdot v}) \subseteq (\mathrm{dom}\,\mathrm{conv}\,\mathrm{supp}\, v)^{\preccurlyeq}$ and thus $\mathbf{P}(\overline{G \cdot v}) \subseteq (\mathrm{dom}\,\mathrm{conv}\,\mathrm{supp}\, v)^{\preccurlyeq}$. □

6.9 Inner approximation for free tensors

In this section we discuss an inner approximation for the moment polytope of a free tensor. We will use this inner approximation in the next section to prove that the quantum functionals coincide with the support functionals when restricted to free tensors. We will prove that not all tensors are free.

We say a set $\Phi \subseteq [n_1] \times [n_2] \times [n_3]$ is *free* if every two different elements of Φ differ in at least two coordinates, in other words if the elements of Φ have Hamming distance at least two. We say $v \in V = \mathbb{C}^{n_1} \otimes \mathbb{C}^{n_2} \otimes \mathbb{C}^{n_3}$ is *free* if for some $g \in G(v) = \mathrm{GL}_{n_1} \times \mathrm{GL}_{n_2} \times \mathrm{GL}_{n_3}$ the support $\mathrm{supp}(g \cdot v) \subseteq [n_1] \times [n_2] \times [n_3]$ is free. (Free is called "schlicht" in [Str05].)

Theorem 6.18 (Strassen [Str05])**.** *Let $v \in V \setminus 0$ with* $\mathrm{supp}(v)$ *free. Then*

$$\mathrm{dom}\,\mathrm{conv}\,\mathrm{supp}\, v \subseteq \mathbf{P}(v).$$

Proof. We refer to [Str05]. □

Corollary 6.19. *Let $v \in V \setminus 0$ with* $\mathrm{supp}(v)$ *free. Then*

$$\mathbf{P}(v)^{\preccurlyeq} = \big(\mathrm{dom}\,\mathrm{conv}\,\mathrm{supp}\, v\big)^{\preccurlyeq}.$$

Proof. By Theorem 6.18 $\mathrm{dom}\,\mathrm{conv}\,\mathrm{supp}\, v \subseteq \mathbf{P}(v)$. We take the upward closure on both sides to get $(\mathrm{dom}\,\mathrm{conv}\,\mathrm{supp}\, v)^{\preccurlyeq} \subseteq \mathbf{P}(v)^{\preccurlyeq}$. On the other hand, from Theorem 6.17 follows $\mathbf{P}(v)^{\preccurlyeq} \subseteq (\mathrm{dom}\,\mathrm{conv}\,\mathrm{supp}\, v)^{\preccurlyeq}$. □

Remark 6.20. Recall that $v \in V$ is oblique if the support $\mathrm{supp}(g \cdot v)$ is an antichain for some $g \in G(v)$ (Section 4.4). Such antichains are free, so oblique tensors are free. Thus $\{\text{tight}\} \subseteq \{\text{oblique}\} \subseteq \{\text{free}\}$. Like the tight tensors and oblique tensors, free tensors from a semigroup under $\otimes$ and $\oplus$.

Proposition 6.21. *For $n \geq 5$ there exists a tensor that is not free in $\mathbb{C}^n \otimes \mathbb{C}^n \otimes \mathbb{C}^n$.*

Proof. We upper bound the maximal size of a free support. Let $\Phi \subseteq [n] \times [n] \times [n]$ be free. Any two distinct elements in Φ are still distinct if we forget the third coefficient of each. Therefore, $|\Phi| = |\{(\alpha_1, \alpha_2) : \alpha \in \Phi\}| \leq n^2$. (This is a special case of the Singleton bound [Sin64] from coding theory. This upper bound is tight, since $\Phi = \{(a, b, c) : a, b, c \in [n], c = a + b \bmod n\}$ is free and has size n^2.) Second we apply the following observation of Bürgisser [Bür90, page 3]. Let

$$Z_n = \{t \in \mathbb{C}^n \otimes \mathbb{C}^n \otimes \mathbb{C}^n : \exists g \in G(t)\ |\mathrm{supp}(g \cdot t)| < n^3 - 3n^2\}.$$

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

6.10 Quantum functionals versus support functionals

We discussed the support functionals $\zeta^\theta \in \mathbf{X}(\{\text{oblique 3-tensors over } \mathbb{F}\})$ in Chapter 4. We recall its definition over $\mathbb{C}$. Let $V = \mathbb{C}^{n_1} \otimes \mathbb{C}^{n_2} \otimes \mathbb{C}^{n_3}$. For $\theta \in \Theta := \mathcal{P}([3])$ and $t \in V \setminus 0$ with $\mathrm{supp}(t)$ oblique,

$$\zeta^\theta(t) = \max\{2^{H_\theta(P)} : P \in \mathcal{P}(\mathrm{supp}(t))\}.$$

We also discussed an extension of ζ^θ to all 3-tensors over $\mathbb{C}$: the upper support functional:

$$\zeta^\theta(t) = \min_{g \in G(t)} \max\{2^{H_\theta(P)} : P \in \mathcal{P}(\mathrm{supp}(g \cdot t))\}.$$

We know $\zeta^\theta(s \otimes t) \leq \zeta^\theta(s)\zeta^\theta(t)$, $\zeta^\theta(s \oplus t) = \zeta^\theta(s) + \zeta^\theta(t)$, $\zeta^\theta(\langle 1 \rangle) = 1$ and $s \leqslant t \Rightarrow \zeta^\theta(s) \leq \zeta^\theta(t)$ for any $s, t \in V$.

The set $\mathrm{conv}\,\mathrm{supp}(g \cdot t)$ is the set of marginals of probability distributions on $\mathrm{supp}(g \cdot t)$. Thus $\mathrm{dom}\,\mathrm{conv}\,\mathrm{supp}(g \cdot t)$ is the set of ordered marginals of probability distributions on $\mathrm{supp}(g \cdot t)$. Therefore

$$\zeta^\theta(t) = \min_{g \in G(t)} \max_{x \in S(g \cdot t)} 2^{H_\theta(x)}$$

with $S(w) = \mathrm{dom}\,\mathrm{conv}\,\mathrm{supp}\,w$. Let $X \subseteq \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \times \mathbb{R}^{n_3}$ be a set of triples of probability vectors. From Schur-convexity of the Shannon entropy function follows $\max_{x \in X} 2^{H_\theta(x)} = \max_{x \in X^{\preccurlyeq}} 2^{H_\theta(x)}$. Also $H_\theta(x) = H_\theta(\mathrm{dom}\, x)$.

Theorem 6.22. $\zeta^\theta(t) \geq F^\theta(t)$.

Proof. Let $g \in G(t)$ such that

$$\max_{x \in S} 2^{H_\theta(x)} = \zeta^\theta(t)$$

with $S = \mathrm{dom}\,\mathrm{conv}\,\mathrm{supp}(g \cdot t)$. We have

$$\max_{x \in S} 2^{H_\theta(x)} = \max_{x \in S^{\preccurlyeq}} 2^{H_\theta(x)}.$$

By Theorem 6.17, $\mathbf{P}(t) \subseteq S^{\preccurlyeq}$. We conclude $F^\theta(t) \leq \zeta^\theta(t)$. □

Theorem 6.23. *Let $t \in V$ be free. Then $\zeta^\theta(t) = F^\theta(t)$.*

Proof. We know from Theorem 6.22 that $\zeta^\theta(t) \geq F^\theta(t)$. We prove $\zeta^\theta(t) \leq F^\theta(t)$. Let $g \in G(t)$ such that $\mathrm{supp}(g \cdot t)$ is free. Let $S = \mathrm{dom}\,\mathrm{conv}\,\mathrm{supp}(g \cdot t)$. Then $\zeta^\theta(t) \leq \max_{x \in S} 2^{H_\theta(x)} = \max_{x \in S^{\preccurlyeq}} 2^{H_\theta(x)}$. By Theorem 6.18 we have $S^{\preccurlyeq} = \mathbf{P}(t)^{\preccurlyeq}$. We conclude $\zeta^\theta(t) \leq F^\theta(t)$. □

We can show that the regularised upper support functional equals the quantum support functional. As a consequence, the quantum functional is at least the lower support functional which was discussed in Chapter 4.

Theorem 6.24. $\lim_{n\to\infty} \frac{1}{n}\zeta^\theta(t^{\otimes n})^{1/n} = F^\theta(t)$.

Proof. We refer the reader to [CVZ18]. □

Corollary 6.25. $F^\theta(v) \geq \zeta_\theta(v)$.

Proof. By Theorem 6.24, $F^\theta(v) = \lim_{n\to\infty} \zeta^\theta(v^{\otimes n})^{1/n}$. We know $\zeta^\theta(v) \geq \zeta_\theta(v)$ by Theorem 4.15 and thus $\lim_{n\to\infty} \zeta^\theta(v^{\otimes n})^{1/n} \geq \lim_{n\to\infty} \zeta_\theta(v^{\otimes n})^{1/n}$. The lower support functional ζ_θ is supermultiplicative under $\otimes$ (Theorem 4.14), so

$$\lim_{n\to\infty} \zeta_\theta(v^{\otimes n})^{1/n} \geq \zeta_\theta(v).$$

Combining these three inequalities proves the theorem. □

6.11 Asymptotic slice rank

We proved in Section 4.6 that for oblique $t \in \mathbb{F}^{n_1} \otimes \mathbb{F}^{n_2} \otimes \mathbb{F}^{n_3}$ the asymptotic slice rank $\lim_{n\to\infty} \mathrm{SR}(t^{\otimes n})^{1/n}$ exists and equals $\min_{\theta\in\Theta} \zeta^\theta(t)$ with $\Theta \coloneqq \mathcal{P}([3])$. In this section we prove the analogous statement for the quantum functionals.

Theorem 6.26. *Let $t \in \mathbb{C}^{n_1} \otimes \mathbb{C}^{n_2} \otimes \mathbb{C}^{n_3}$. Then*

$$\lim_{n\to\infty} \mathrm{SR}(t^{\otimes n})^{1/n} = \min_{\theta\in\Theta} F^\theta(t).$$

We work towards the proof of Theorem 6.26. Let $t \in \mathbb{C}^{n_1} \otimes \mathbb{C}^{n_2} \otimes \mathbb{C}^{n_3} \setminus 0$. Let $E^\theta(t) \coloneqq \log_2 F^\theta(t)$.

Lemma 6.27. *For any $\varepsilon > 0$ there is an $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$ there is a $\lambda/n \in \mathbf{R}(t)$ with $\min_{i\in[3]} H(\overline{\lambda^{(i)}}) \geq \min_{\theta\in\Theta} E^\theta(t) - \varepsilon$.*

Proof. By definition

$$\min_{\theta\in\Theta} E^\theta(t) = \min_{\theta\in\Theta} \max_{x\in\mathbf{P}(t)} \sum_{j\in[3]} \theta(j) H(x^{(j)}).$$

By Von Neumann's minimax theorem, the right-hand side equals

$$\max_{x\in\mathbf{P}(t)} \min_{\theta\in\Theta} \sum_{j\in[3]} \theta(j) H(x^{(j)})$$

which equals

$$\max_{x\in\mathbf{P}(t)} \min_{j\in[3]} H(x^{(j)}).$$

Let $\varepsilon > 0$. Let $\mu/m \in \mathbf{R}(t)$ with $\min_{j\in[3]} H(\overline{\mu^{(j)}}) \geq \min_{\theta\in\Theta} E^\theta(t) - \varepsilon/2$. We will use two facts. We have $(P_{(1)} \otimes P_{(1)} \otimes P_{(1)})t = t \neq 0$. The triples of partitions λ with $P_\lambda t^{\otimes n} \neq 0$ for some n form a semigroup. Let $n \in \mathbb{N}$. We can write $n = qm+r$ with $q, r \in \mathbb{N}$, $0 \leq r < m$. Let $\lambda^{(j)} = q\mu^{(j)} + (r)$. Then by the semigroup property $P_\lambda t^{\otimes n} \neq 0$, i.e. $\lambda/n \in \mathbf{R}(t)$. We have $\frac{1}{n}(q\mu^{(j)} + (r)) = \frac{qm}{n}\overline{\mu^{(j)}} + \frac{r}{n}\overline{(r)}$. By concavity of Shannon entropy

$$\begin{aligned} H(\tfrac{1}{n}(q\mu^{(j)} + (r))) &= H(\tfrac{qm}{n}\overline{\mu^{(j)}} + \tfrac{r}{n}\overline{(r)}) \\ &\geq \tfrac{qm}{n} H(\overline{\mu^{(j)}}) \\ &\geq (1 - \tfrac{m}{n}) H(\overline{\mu^{(j)}}). \end{aligned}$$

When n is large enough $(1 - \frac{m}{n})H(\overline{\mu^{(j)}})$ is at least $H(\overline{\mu^{(j)}}) - \varepsilon/2$. Let $n_0 \in \mathbb{N}$ such that this is the case for all $j \in [3]$. □

Lemma 6.28. *Let* $\lambda/n \in \mathbf{R}(t)$. *Then* $\mathrm{SR}(t^{\otimes n}) \geq \min_{i\in[3]} \dim[\lambda^{(i)}]$.

Proof. We have the restriction $t^{\otimes n} \geq P_\lambda t^{\otimes n} \neq 0$. Choose rank-one projections A_j in the vector spaces $\mathbb{S}_{\lambda^{(j)}}(\mathbb{C}^{n_j})$ with

$$s \coloneqq (\mathrm{id}_{[\lambda^{(1)}]} \otimes A_1) \otimes (\mathrm{id}_{[\lambda^{(2)}]} \otimes A_2) \otimes (\mathrm{id}_{[\lambda^{(3)}]} \otimes A_3) P_\lambda t^{\otimes n} \neq 0.$$

The tensor s is invariant under S_n acting diagonally on $(\mathbb{C}^{n_1})^{\otimes n} \otimes (\mathbb{C}^{n_2})^{\otimes n} \otimes (\mathbb{C}^{n_3})^{\otimes n}$. Thus the marginal spectra $\operatorname{spec} \rho_i^s$ are uniform. This implies s is semistable. From [BCC+17, Theorem 4.6] follows that $\mathrm{SR}(s)$ equals $\min_{i\in[3]} \dim[\lambda^{(i)}]$. □

Lemma 6.29. $\liminf_{n\to\infty} \mathrm{SR}(t^{\otimes n})^{1/n} \geq \min_{\theta\in\Theta} F^\theta(t)$.

Proof. Let $\varepsilon > 0$. For n large enough choose $\lambda/n \in \mathbf{R}(t)$ as in Lemma 6.27. By Lemma 6.28, $\mathrm{SR}(t^{\otimes n}) \geq \min_{i\in[3]} \dim[\lambda^{(i)}]$. The right-hand side we lower bound by

$$\min_{i\in[3]} \dim[\lambda^{(i)}] \geq \min_{i\in[3]} 2^{nH(\overline{\lambda^{(i)}})} 2^{-o(n)} \geq 2^{n(\min_{\theta\in\Theta} E^\theta(t) - \varepsilon)} 2^{-o(n)}.$$

Then $\liminf_{n\to\infty} \mathrm{SR}(t^{\otimes n})^{1/n} \geq 2^{\min_{\theta\in\Theta} E^\theta(t) - \varepsilon}$. □

Lemma 6.30. $\limsup_{n\to\infty} \mathrm{SR}(t^{\otimes n})^{1/n} \leq F^\theta(t)$.

Proof. Let $n \in \mathbb{N}$. Define $s_1, s_2, s_3 \in \mathbb{C}^{n_1} \otimes \mathbb{C}^{n_2} \otimes \mathbb{C}^{n_3}$ by

$$s_1 = \Big(\sum_{\substack{\lambda^{(1)} \vdash n: \\ H(\overline{\lambda^{(1)}}) \leq E^\theta(t)}} P_{\lambda^{(1)}} \otimes \mathrm{Id} \otimes \mathrm{Id}\Big) t^{\otimes n}$$

$$s_2 = \Big(\sum_{\substack{\lambda^{(2)} \vdash n: \\ H(\overline{\lambda^{(2)}}) \leq E^\theta(t)}} \mathrm{Id} \otimes P_{\lambda^{(2)}} \otimes \mathrm{Id}\Big)(t^{\otimes n} - s_1)$$

$$s_3 = \Big(\sum_{\substack{\lambda^{(3)} \vdash n: \\ H(\overline{\lambda^{(3)}}) \leq E^\theta(t)}} \mathrm{Id} \otimes \mathrm{Id} \otimes P_{\lambda^{(3)}}\Big)(t^{\otimes n} - s_1 - s_2).$$

Then $t^{\otimes n} = s_1 + s_2 + s_3$. The slice rank of an element in the image of $P_{\lambda^{(1)}} \otimes \mathrm{Id} \otimes \mathrm{Id}$ is at most $\dim[\lambda^{(1)}] \otimes \mathbb{S}_{\lambda^{(1)}}(\mathbb{C}^{n_1})$, which is at most $2^{nH(\overline{\lambda^{(1)}})+o(n)}$ (Section 6.2). Similarly for $\mathrm{Id} \otimes P_{\lambda^{(2)}} \otimes \mathrm{Id}$ and $\mathrm{Id} \otimes \mathrm{Id} \otimes P_{\lambda^{(3)}}$. The tensor s_1 is in the image of the sum $\sum_{\lambda^{(1)}} P_{\lambda^{(1)}} \otimes \mathrm{Id} \otimes \mathrm{Id}$ over $\lambda^{(1)} \vdash n$ with at most n_1 parts. There are at most $(n+1)^{n_1}$ such partitions. Thus $\mathrm{SR}(s_1) \leq (n+1)^{n_1} 2^{nE^\theta(t)+o(n)}$. Similarly for s_2 and s_3. Therefore

$$\limsup_{n\to\infty} \mathrm{SR}(t^{\otimes n})^{1/n} \leq \limsup_{n\to\infty} \big(3(n+1)^{\max_{i\in[3]} n_i}\, 2^{nE^\theta(t)+o(n)}\big)^{1/n}. \tag{6.6}$$

The right-hand side of (6.6) equals $F^\theta(t)$. □

Proof of Theorem 6.26. Lemma 6.29 and Lemma 6.30 together prove Theorem 6.26. □

6.12 Conclusion

In this chapter we constructed the first infinite family of spectral points for 3-tensors over $\mathbb{C}$, the quantum functionals. For 30 years the only explicit spectral points known were the gauge points. The constructions in this chapter naturally generalise to higher-order tensors, for which we refer to our paper [CVZ18]. We do not know whether the quantum functionals are all of the spectral points for 3-tensors over $\mathbb{C}$. Finally, we showed that for complex tensors the asymptotic slice rank exists and equals the minimum value over all the quantum functionals.

Chapter 7

Algebraic branching programs; approximation and nondeterminism

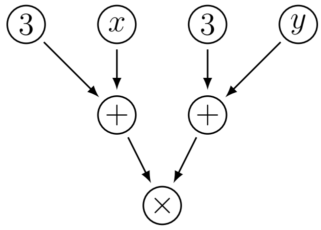
This chapter is based on joint work with Karl Bringmann and Christian Ikenmeyer [BIZ17].

7.1 Introduction

The study of asymptotic tensor rank in previous chapters was originally motivated by the study of the complexity of matrix multiplication in the algebraic circuit model, an algebraic model of computation. In this chapter we will study several other algebraic models of computation and algebraic complexity classes.

Formulas, the class $\mathbf{VP}_{\mathsf{e}}$ and the determinant

An (arithmetic) formula is a rooted binary tree whose leaves are each labeled with a variable or a field constant, and whose root and intermediate vertices are labeled with either $+$ (addition) or $\times$ (multiplication). In the natural way, via recursion over the tree structure, a formula computes a multivariate polynomial f. The *formula size* of a multivariate polynomial f is the smallest number of vertices required for any formula to compute f. Here is an example of a formula of size 7 computing the polynomial $(3+x)(3+y)$.



A sequence of multivariate polynomials $(f_n)_{n\in\mathbb{N}}$ is called a *family*. Valiant in his seminal paper [Val79] introduced the complexity class $\mathbf{VP}_{\mathsf{e}}$ that is defined as the set of all families whose formula size is polynomially bounded. (We say a sequence $(a_n)_n \in \mathbb{N}^{\mathbb{N}}$ of natural numbers is *polynomially bounded* if there exists a univariate polynomial q such that $a_n \leq q(n)$ for all n.) For example, the family $((x_1)^n + (x_2)^n + \cdots + (x_n)^n)_n$ is in $\mathbf{VP}_{\mathsf{e}}$, because the formula size of this family grows quadratically.

The smallest known formulas for the determinant family $\det_n$ have size $n^{\mathcal{O}(\log n)}$. This follows from Berkowitz' algorithm [Ber84], which gives an algebraic circuit of depth $\mathcal{O}(\log^2 n)$, and thus by expanding we get an algebraic formula of depth $\mathcal{O}(\log^2 n)$ whose size is then trivially bounded by $2^{\mathcal{O}(\log^2 n)} = n^{\mathcal{O}(\log n)}$. It is a major open question in algebraic complexity theory whether formulas of polynomially bounded size exist for $\det_n$. This question can be phrased in terms of complexity classes as asking whether or not the inclusion $\mathbf{VP}_{\mathsf{e}} \subseteq \mathbf{VP}_{\mathsf{s}}$ is strict. (We will define $\mathbf{VP}_{\mathsf{s}}$ shortly.)

Motivated by this question we study the closure class $\overline{\mathbf{VP}_{\mathsf{e}}}$ of families of polynomials that can be approximated arbitrarily closely by families in $\mathbf{VP}_{\mathsf{e}}$ (see Section 7.2.4 for the formal definition). Over the field $\mathbb{R}$ or $\mathbb{C}$ one can think of $\overline{\mathbf{VP}_{\mathsf{e}}}$ as the set of families whose *border formula size* is polynomially bounded. The border formula size of a polynomial f is the smallest number c such that there exists a sequence g_i of polynomials with formula size at most c and $\lim_{i\to\infty} g_i = f$.

Continuous lower bounds

In algebraic complexity theory, problem instances correspond to vectors $v \in \mathbb{F}^n$. A complexity lower bound often takes the form of a function $f : \mathbb{F}^n \to \mathbb{F}$ that is zero on the vectors of "low complexity" and nonzero on v. We refer to Grochow [Gro13] for a discussion of settings where complexity lower bounds are obtained in this way (e.g. [NW97, Raz09, LO15, GKKS13, LMR13, BI13]). Over the complex numbers we can in fact assume that these functions f are continuous [Gro13] (and even so-called highest-weight vector polynomials). If $\mathbf{C}$ and $\mathbf{D}$ are algebraic complexity classes with $\mathbf{C} \subseteq \mathbf{D}$ (for example, $\mathbf{C} = \mathbf{VP}_{\mathsf{e}}$ and $\mathbf{D} = \mathbf{VP}_{\mathsf{s}}$), then a proof of separation $\mathbf{D} \not\subseteq \mathbf{C}$ in this continuous manner implies the stronger separation $\mathbf{D} \not\subseteq \overline{\mathbf{C}}$. In our case, it is thus natural to aim for the separation $\mathbf{VP}_{\mathsf{s}} \not\subseteq \overline{\mathbf{VP}_{\mathsf{e}}}$ instead of the slightly weaker $\mathbf{VP}_{\mathsf{s}} \not\subseteq \mathbf{VP}_{\mathsf{e}}$, which provides further motivation for studying $\overline{\mathbf{VP}_{\mathsf{e}}}$. This is exactly analogous to the geometric complexity theory approach of Mulmuley and Sohoni (see e.g. [MS01, MS08] and the exposition [BLMW11, Sec. 9]) which aims to prove the separation $\mathbf{VNP} \not\subseteq \overline{\mathbf{VP}_{\mathsf{s}}}$ to attack Valiant's famous conjecture $\mathbf{VP}_{\mathsf{s}} \neq \mathbf{VNP}$ [Val79]. (Here $\mathbf{VNP}$ is the class of p-definable families, see Section 7.2.4.)

New results in this chapter

We prove two new results in this chapter

Algebraic branching programs of width 2. An *algebraic branching program* (abp) is a directed acyclic graph with a source vertex s and a sink vertex t that has affine linear forms (in one or more variables) over the base field $\mathbb{F}$ as edge labels. Moreover, we require that each vertex is labeled with an integer (its layer) and that edges in the abp only point from vertices in layer i to vertices in layer $i+1$. The *width* of an abp is the cardinality of its largest layer. The *size* of an abp is the number of its vertices. The *value* of an abp is the sum of the values of all s–t-paths, where the value of an s–t-path is the product of its edge labels. We say that an abp *computes* its value. The class $\mathbf{VP}_\mathsf{s}$ coincides with the class of families of polynomials that can be computed by abps of polynomially bounded size, see e.g. [Sap16].

For $k \in \mathbb{N}$ we introduce the class $\mathbf{VP}_k$ as the class of families of polynomials computable by width-k abps of polynomially bounded size. It is well-known (see Lemma 7.2) that $\mathbf{VP}_k \subseteq \mathbf{VP}_\mathsf{e}$ for all $k \geq 1$. In 1992, Ben-Or and Cleve [BOC92] showed that $\mathbf{VP}_k = \mathbf{VP}_\mathsf{e}$ for all $k \geq 3$. In 2011, Allender and Wang [AW16] showed that width-2 abps cannot compute every polynomial, so in particular we have a strict inclusion $\mathbf{VP}_2 \subsetneq \mathbf{VP}_3$.

We prove that the closure of $\mathbf{VP}_2$ and the closure of $\mathbf{VP}_\mathsf{e}$ are equal,

$$\overline{\mathbf{VP}_2} = \overline{\mathbf{VP}_\mathsf{e}}, \tag{7.1}$$

when $\mathrm{char}(\mathbb{F}) \neq 2$. From (7.1) and the result of Allender and Wang follows directly that the inclusion $\mathbf{VP}_2 \subsetneq \overline{\mathbf{VP}_2}$ is strict. We have thus separated a complexity class from its approximation closure.

VNP via affine linear forms. Every algebraic complexity class has a nondeterministic closure (see Section 7.2.5 for the definition). The nondeterministic closure of $\mathbf{VP}$ is called $\mathbf{VNP}$, and the nondeterministic closure of $\mathbf{VP}_\mathsf{e}$ is called $\mathbf{VNP}_\mathsf{e}$. In 1980, Valiant [Val80] proved $\mathbf{VNP}_\mathsf{e} = \mathbf{VNP}$. The nondeterministic closure of $\mathbf{VP}_1$ and $\mathbf{VP}_2$ we call $\mathbf{VNP}_1$ and $\mathbf{VNP}_2$. Using interpolation techniques we can deduce $\mathbf{VNP}_2 = \mathbf{VNP}$ from (7.1), provided the field is infinite. Using more sophisticated techniques we prove

$$\mathbf{VNP}_1 = \mathbf{VNP}. \tag{7.2}$$

From (7.2) easily follows $\mathbf{VP}_1 \subsetneq \mathbf{VNP}_1$. Also, from [AW16] we get $\mathbf{VP}_2 \subsetneq \mathbf{VNP}_2$. We have thus separated complexity classes from their nondeterministic closures.

Further related work

An excellent exposition on the history of small-width computation can be found in [AW16], along with an explicit polynomial that cannot be computed by width-2 abps, namely $x_1x_2 + x_3x_4 + \cdots + x_{15}x_{16}$. Saha, Saptharishi and Saxena in [SSS09, Cor. 14] showed that $x_1x_2 + x_3x_4 + x_5x_6$ cannot be computed by width-2 abps that correspond to the iterated matrix multiplication of upper triangular matrices. Bürgisser in [Bür04] studied approximations in the model of general algebraic circuits, finding general upper bounds on the error degree. For most algebraic complexity classes $\mathbf{C}$ the relation between $\mathbf{C}$ and $\overline{\mathbf{C}}$ has not been an active object of study. As pointed out recently by Forbes [For16], Nisan's result [Nis91] implies that $\mathbf{C} = \overline{\mathbf{C}}$ for $\mathbf{C}$ being the class of size-k algebraic branching programs on noncommuting variables. A structured study of $\overline{\mathbf{VP}}$ and $\overline{\mathbf{VP}_\mathsf{s}}$ was started in [GMQ16]. Much work in lower bounds for algebraic approximation algorithms has been done in the area of bilinear complexity, dating back to [BCRL79, Str83, Lic84] and more recently e.g. [Lan06, LO15, HIL13, Zui17, LM16a].

This chapter is organised as follows. In Section 7.2 we discuss definitions and basic results. In Section 7.3 we prove that the approximation closure of $\mathbf{VP}_2$ equals the approximation closure of $\mathbf{VP}_\mathsf{e}$, i.e. $\overline{\mathbf{VP}_2} = \overline{\mathbf{VP}_\mathsf{e}}$. In Section 7.4 we prove that the nondeterminism closure of $\mathbf{VP}_1$ equals $\mathbf{VNP}$.

7.2 Definitions and basic results

We briefly recall the definition of circuits, formulas and branching programs and we recall the definition of the corresponding complexity classes. Then we discuss some straightforward relationships among these classes and review the proof of a theorem by Ben-Or and Cleve, which inspired our work. Finally, we discuss the approximation closure and the nondeterminism closure for algebraic complexity classes.

7.2.1 Computational models

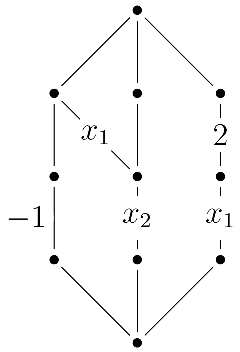
Let $x_1, x_2, \ldots$ be formal variables. By $\mathbb{F}[\mathbf{x}]$ we mean the ring of polynomials over $\mathbb{F}$ with variables $x_1, x_2, \ldots, x_k$ with k large enough.

A *circuit* is a directed acyclic graph G with one or more source vertices and one sink vertex. Each source vertex is labelled by a variable x_i or a constant $c \in \mathbb{F}$. The other vertices are labelled by either $+$ or $\times$ and have in-degree 2 (that is, fan-in 2). Each vertex computes an element in $\mathbb{F}[\mathbf{x}]$ by recursion over the graph. The element computed by the sink is the element computed by the circuit. The *size* of a circuit is the number of vertices.

A *formula* is a circuit whose graph is a tree.

An *algebraic branching program* (abp) is a directed acyclic graph with a source vertex s and a sink vertex t that has affine linear forms $\sum_i \alpha_i x_i + \beta$, $\alpha, \beta \in \mathbb{F}$ as edge labels. Moreover, we require that each vertex is labeled with an integer (its layer) and that edges in the abp only point from vertices in layer i to vertices in layer $i + 1$. The *width* of an abp is the cardinality of its largest layer. The *size* of an abp is the number of its vertices. The *value* of an abp is the sum of the values of all s–t-paths, where the value of an s–t-path is the product of its edge labels. We say that an abp computes its value.

For example, the following abp has depth 5, width 3 and computes the polynomial $x_1x_2 + x_2 + 2x_1 - 1$.



An abp G corresponds naturally to an iterated product of matrices: for any two consecutive layers L_i, L_{i+1} in G, let M_i be the matrix $(e_{v,w})_{v\in L_i, w\in L_{i+1}}$ with $e_{v,w}$ the label of the edge from v to w (or 0 if there is no edge from v to w). Then the value of G equals the product $M_d \cdots M_2M_1$.

For example, the above abp corresponds to the following iterated matrix product:

$$\begin{pmatrix} 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} -1 & 0 & 0 \\ 0 & x_2 & 0 \\ 0 & 0 & x_1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ x_1 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

7.2.2 Complexity classes VP, $\mathbf{VP}_{\mathsf{e}}$, $\mathbf{VP}_k$

The circuit size of a polynomial f is the size of the smallest circuit computing f. The formula size of a polynomial f is the size of the smallest formula computing f.

A *family* is a sequence $(f_n)_{n\in\mathbb{N}}$ of multivariate polynomials over $\mathbb{F}$. A *class* is a set of families. The class **VP** consists of all families (f_n) with circuit size, degree and number of variables in poly(n). The class $\mathbf{VP}_{\mathsf{e}}$ consists of all families (f_n) with formula size in poly(n). (The origin of the subscript e in $\mathbf{VP}_{\mathsf{e}}$ is the term "arithmetic expression".) Clearly, $\mathbf{VP}_{\mathsf{e}} \subseteq \mathbf{VP}$.

We introduce classes defined by abps. Let $k \geq 1$. The class $\mathbf{VP}_k$ consists of all families computed by polynomial-size width-k abps with edges labelled by affine linear forms $\sum_i \alpha_i x_i + \beta$ with coefficients $\alpha_i, \beta \in \mathbb{F}$.

We note that the above classes depend on the choice of the ground field $\mathbb{F}$.

In our paper [BIZ17] we make a distinction between three different types of edge labels for abps. The class $\mathbf{VP}_k$ in this chapter corresponds to the class $\mathbf{VP}_k^{\mathrm{g}}$ in [BIZ17].

7.2.3 The theorem of Ben-Or and Cleve

This subsection is about the relations among $\mathbf{VP}_k$ and $\mathbf{VP}_{\mathsf{e}}$.

Lemma 7.1. *$\mathbf{VP}_k \subseteq \mathbf{VP}_\ell$ when $k \leq \ell$.*

Proof. This is clearly true. □

Lemma 7.2. *$\mathbf{VP}_k \subseteq \mathbf{VP}_{\mathsf{e}}$ for any k.*

Proof. For the simple proof we refer to [BIZ17]. □

Ben-Or and Cleve [BOC92] showed that for $k \geq 3$, the classes $\mathbf{VP}_k$ and $\mathbf{VP}_{\mathsf{e}}$ are in fact equal.

Theorem 7.3 (Ben-Or and Cleve [BOC92])**.** *For $k \geq 3$, $\mathbf{VP}_k = \mathbf{VP}_{\mathsf{e}}$.*

We will review the construction of Ben-Or and Cleve here, because we will use it to prove Theorem 7.8 and Theorem 7.15. The following depth-reduction lemma for formulas by Brent is a crucial ingredient.

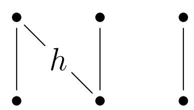
Lemma 7.4 (Brent [Bre74])**.** *Let f be an n-variate degree-d polynomial computed by a formula of size s. Then f can also be computed by a formula of size* $\mathrm{poly}(s, n, d)$ *and depth $\mathcal{O}(\log s)$.*

Proof. See the survey of Saptharishi [Sap16, Lemma 5.5] for a modern proof. □

Proof of Theorem 7.3. Lemma 7.2 says $\mathbf{VP}_k \subseteq \mathbf{VP}_{\mathsf{e}}$. We will prove the inlusion $\mathbf{VP}_{\mathsf{e}} \subseteq \mathbf{VP}_3$, from which follows $\mathbf{VP}_{\mathsf{e}} \subseteq \mathbf{VP}_k$ by Lemma 7.1 and thus $\mathbf{VP}_k = \mathbf{VP}_{\mathsf{e}}$. For a polynomial h, define the matrix

$$M(h) \coloneqq \begin{pmatrix} 1 & 0 & 0 \\ h & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

which, as part of an abp, looks like



We call the following matrices *primitive*:

- $M(h)$ with h any variable or any constant in $\mathbb{F}$
- the 3×3 permutation matrices, denoted by M_π with $\pi \in S_3$
- the diagonal matrices $M_{a,b,c} \coloneqq \mathrm{diag}(a, b, c)$ with $a, b, c \in \mathbb{F}$.

The entries of the primitives are variables or constants in $\mathbb{F}$, making them suitable to use in the construction of a width-3 abp.

Let $(f_n) \in \mathbf{VP_e}$. Then f_n can be computed by a formula of size $s(n) \in \text{poly}(n)$. By Brent's depth-reduction theorem for formulas (Lemma 7.4) f_n can be computed by a formula of size $\text{poly}(n)$ and depth $d(n) \in \mathcal{O}(\log s(n))$.

We will construct a sequence of primitives $A_1, \ldots, A_{m(n)}$ such that

$$A_1 \cdots A_{m(n)} = \begin{pmatrix} 1 & 0 & 0 \\ f_n & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

with $m(n) \in \mathcal{O}(4^{d(n)}) = \text{poly}(n)$. Then

$$f_n(\mathbf{x}) = \left(\begin{smallmatrix} 1 & 1 & 1 \end{smallmatrix}\right) M_{-1,1,0} A_1 \cdots A_m \left(\begin{smallmatrix} 1 \\ 1 \\ 1 \end{smallmatrix}\right),$$

so $f_n(\mathbf{x})$ can be computed by a width-3 abp of length $\text{poly}(n)$, proving the theorem.

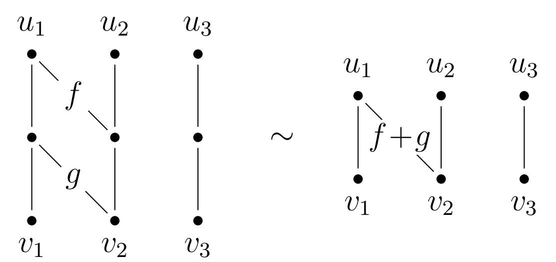
To explain the construction, let h be a polynomial and consider a formula computing h of depth d. The goal is to construct (recursively on the formula structure) primitives $A_1, \ldots, A_m$ such that

$$A_1 \cdots A_m = \begin{pmatrix} 1 & 0 & 0 \\ h & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

with $m \in \mathcal{O}(4^d)$.

Suppose h is a variable or a constant. Then $M(h)$ is itself a primitive matrix.

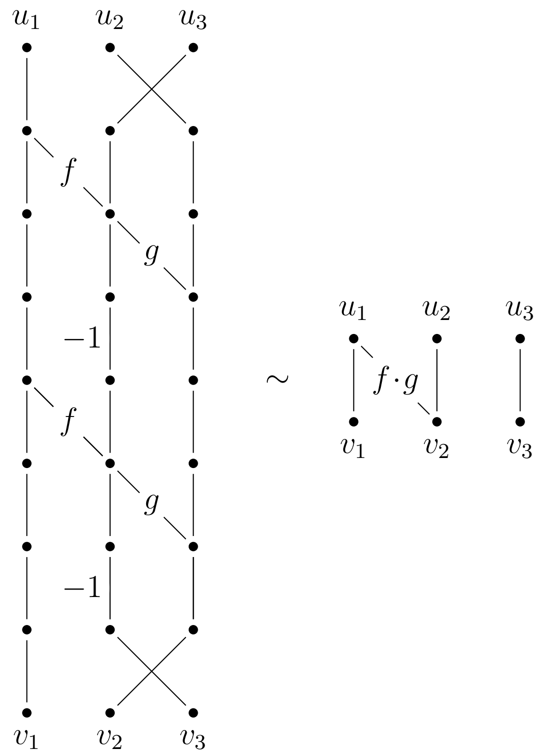
Suppose $h = f + g$ is a sum of two polynomials f, g and suppose $M(f)$ and $M(g)$ can be written as a product of primitives. Then $M(f+g)$ equals a product of primitives, because $M(f+g) = M(f)M(g)$. This can easily be verified directly, or by noting that in the corresponding partial abps the top-bottom paths (u_i-v_j paths) have the same value:



Suppose $h = fg$ is a product of two polynomials f, g and suppose $M(f)$ and $M(g)$ can be written as a product of primitives. Then $M(fg)$ equals a product of primitives, because

$$M(f \cdot g) = M_{(23)} \big(M_{1,-1,1} M_{(123)} M(g) M_{(132)} M(f) \big)^2 M_{(23)}$$

(here $(23) \in S_3$ denotes the transposition $1 \mapsto 1, 2 \mapsto 3, 3 \mapsto 2$ and $(123) \in S_3$ denotes the cyclic shift $1 \mapsto 2, 2 \mapsto 3, 3 \mapsto 1$) as can be verified either directly or by checking that in the corresponding partial abps the top-bottom paths (u_i-v_j paths) have the same value:



This completes the construction.

The length m of the construction is $m(h) = 1$ for h a variable or constant and recursively $m(f + g) = m(f) + m(g)$, $m(f \cdot g) = 2(m(f) + m(g))$, so $m \in \mathcal{O}(4^d)$ where d is the formula depth of h. $\square$

The above result of Ben-Or and Cleve (Theorem 7.3) raises the intriguing question whether the inclusion $\mathbf{VP}_2 \subseteq \mathbf{VP}_\mathsf{e}$ is strict. Allender and Wang [AW16] show that the inclusion is indeed strict; in fact, they show that some polynomials cannot be computed by any width-2 abp.

Theorem 7.5 (Allender and Wang [AW16]). *The polynomial*

$$x_1x_2 + x_3x_4 + \cdots + x_{15}x_{16}$$

cannot be computed by any width-2 abp. Therefore, we have the separation of classes $\mathbf{VP}_2 \subsetneq \mathbf{VP}_3 = \mathbf{VP}_\mathsf{e}$.

7.2.4 Approximation closure $\overline{\mathbf{C}}$

We define the norm of a complex multivariate polynomial as the sum of the absolute values of its coefficients. This defines a topology on the polynomial ring $\mathbb{C}[x_1, \ldots, x_m]$. Given a complexity measure L, say abp size or formula size, there is a natural notion of approximate complexity that is called *border complexity*. Namely, a polynomial $f \in \mathbb{C}[\mathbf{x}]$ has *border complexity* $\underline{L}^{\mathrm{top}}$ at most c if there is a sequence of polynomials $g_1, g_2, \ldots$ in $\mathbb{C}[\mathbf{x}]$ converging to f such that each g_i satisfies $L(g_i) \leq c$. It turns out that for reasonable classes over the field of complex numbers $\mathbb{C}$, this *topological* notion of approximation is equivalent to what we call *algebraic* approximation (see e.g. [Bür04]). Namely, a polynomial $f \in \mathbb{C}[\mathbf{x}]$ satisfies $\underline{L}(f)^{\mathrm{alg}} \leq c$ iff there are polynomials $f_1, \ldots, f_e \in \mathbb{C}[\mathbf{x}]$ such that the polynomial

$$h := f + \varepsilon f_1 + \varepsilon^2 f_2 + \cdots + \varepsilon^e f_e \in \mathbb{C}[\varepsilon, \mathbf{x}]$$

has complexity $L_{\mathbb{C}(\varepsilon)}(h) \leq c$, where ε is a formal variable and $L_{\mathbb{C}(\varepsilon)}(h)$ denotes the complexity of h over the field extension $\mathbb{C}(\varepsilon)$. This algebraic notion of approximation makes sense over any base field and we will use it in the statements and proofs of this chapter.

Definition 7.6. Let $\mathbf{C}(\mathbb{F})$ be a class over the field $\mathbb{F}$. We define the *approximation closure* $\overline{\mathbf{C}}(\mathbb{F})$ as follows: a family (f_n) over $\mathbb{F}$ is in $\overline{\mathbf{C}}(\mathbb{F})$ if there are polynomials $f_{n;i}(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ and a function $e : \mathbb{N} \to \mathbb{N}$ such that the family (g_n) defined by

$$g_n(\mathbf{x}) := f_n(\mathbf{x}) + \varepsilon f_{n;1}(\mathbf{x}) + \varepsilon^2 f_{n;2}(\mathbf{x}) + \cdots + \varepsilon^{e(n)} f_{n;e(n)}(\mathbf{x})$$

is in $\mathbf{C}(\mathbb{F}(\varepsilon))$. We define the *poly-approximation closure* $\overline{\mathbf{C}}^{\mathrm{poly}}(\mathbb{F})$ similarly, but with the additional requirement that $e(n) \in \mathrm{poly}(n)$. We call $e(n)$ the *error degree.*

7.2.5 Nondeterminism closure $\mathrm{N}(\mathbf{C})$

We introduce the nondeterminism closure for algebraic complexity classes.

Definition 7.7. Let $\mathbf{C}$ be a class. The class $\mathrm{N}(\mathbf{C})$ consists of families (f_n) with the following property: there is a family $(g_n) \in \mathbf{C}$ and $p(n), q(n) \in \mathrm{poly}(n)$ such that

$$f_n(\mathbf{x}) = \sum_{\mathbf{b} \in \{0,1\}^{p(n)}} g_{q(n)}(\mathbf{b}, \mathbf{x}),$$

where $\mathbf{x}$ and $\mathbf{b}$ denote sequences of variables $x_1, x_2, \ldots$ and $b_1, b_2, \ldots, b_{p(n)}$. We say that $f_n(\mathbf{x})$ is a hypercube sum over g and that $b_1, b_2, \ldots, b_{p(n)}$ are the hypercube variables. For any subscript x, we will use the notation $\mathbf{VNP}_{\mathsf{x}}$ to denote $\mathrm{N}(\mathbf{VP}_{\mathsf{x}})$. We remark that the map $\mathbf{C} \mapsto \mathrm{N}(\mathbf{C})$ trivially satisfies all properties of being a Kuratowski closure operator, i.e., $\mathrm{N}(\emptyset) = \emptyset$, $\mathbf{C} \subseteq \mathrm{N}(\mathbf{C})$, $\mathrm{N}(\mathbf{C} \cup \mathbf{D}) = \mathrm{N}(\mathbf{C}) \cup \mathrm{N}(\mathbf{D})$, and $\mathrm{N}(\mathrm{N}(\mathbf{C})) = \mathrm{N}(\mathbf{C})$.

7.3 Approximation closure of $\mathbf{VP}_2$

We show that every polynomial can be approximated by a width-2 abp. Even better we show that every polynomial can be approximated by a width-2 abp of size polynomial in the formula size, and with error degree polynomial in the formula size. This is the main result of the current chapter.

Theorem 7.8. $\mathbf{VP}_{\mathsf{e}} \subseteq \overline{\mathbf{VP}_2}^{\mathrm{poly}}$ *when* $\mathrm{char}(\mathbb{F}) \neq 2$.

Proof. For a polynomial h define the matrix $M(h) \coloneqq \left(\begin{smallmatrix} 1 & 0 \\ h & 1 \end{smallmatrix}\right)$. We call the following matrices *primitives*:

- $M(h)$ with h any variable or constant in $\mathbb{F}$
- $\begin{pmatrix} \frac{1}{2\varepsilon} & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} -2\varepsilon & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} -1 & \varepsilon \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} -1 & -\varepsilon \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$

The entries of the primitives are variables or constants in the base field $\mathbb{F}(\varepsilon)$, making them suitable to use in a width-2 abp over the base field $\mathbb{F}(\varepsilon)$.

Let $(f_n) \in \mathbf{VP}_{\mathsf{e}}$, so $f_n(\mathbf{x})$ can be computed by a formula of size $s(n) \in \mathrm{poly}(n)$. By Brent's depth reduction theorem for formulas (Lemma 7.4) f_n can be computed by a formula of size $\mathrm{poly}(n)$ and depth $d(n) \in \mathcal{O}(\log s(n))$.

We will construct a sequence of primitives $A_1, \ldots, A_{m(n)}$ such that

$$A_1 \cdots A_{m(n)} = \begin{pmatrix} 1 & 0 \\ f_n & 1 \end{pmatrix} + \varepsilon \begin{pmatrix} f_{n;111} & f_{n;112} \\ f_{n;121} & f_{n;122} \end{pmatrix} + \varepsilon^2 \begin{pmatrix} f_{n;211} & f_{n;212} \\ f_{n;221} & f_{n;222} \end{pmatrix} + \cdots + \varepsilon^e \begin{pmatrix} f_{n;e11} & f_{n;e12} \\ f_{n;e21} & f_{n;e22} \end{pmatrix}$$

for some $f_{n;ijk} \in \mathbb{F}[\mathbf{x}]$, with $m(n), e(n) \in \mathcal{O}(8^{d(n)}) = \mathrm{poly}(n)$. Then

$$\left(\begin{smallmatrix} 1 & 1 \end{smallmatrix}\right)\left(\begin{smallmatrix} -1 & 0 \\ 0 & 1 \end{smallmatrix}\right) A_1 \cdots A_{m(n)} \left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right) = f_n(\mathbf{x}) + \mathcal{O}(\varepsilon),$$

so $f_n(\mathbf{x})$ can be approximated by a width-2 abp of length $\mathrm{poly}(n)$ and with error degree $\mathrm{poly}(n)$, proving the theorem.

We begin with the construction. Let h be a polynomial and consider a formula computing h of depth d. The goal is to construct, recursively on the tree structure of the formula, a sequence of primitives $A_1, \ldots, A_m$ such that for some $h_{ijk} \in \mathbb{F}[\mathbf{x}]$

$$A_1 \cdots A_m = \begin{pmatrix} 1 & 0 \\ h & 1 \end{pmatrix} + \varepsilon \begin{pmatrix} 0 & 0 \\ h_{121} & 0 \end{pmatrix} + \varepsilon^2 \begin{pmatrix} h_{211} & h_{212} \\ h_{221} & h_{222} \end{pmatrix} + \cdots + \varepsilon^e \begin{pmatrix} h_{e11} & h_{e12} \\ h_{e21} & h_{e22} \end{pmatrix} \tag{7.3}$$

with $m, e \in \mathcal{O}(8^d)$. Notice the particular first-degree error pattern in (7.3), which our recursion will rely on.

Suppose h is a variable or a constant. Then $M(h)$ is itself a primitive satisfying (7.3).

Suppose $h = f + g$ is a sum of two polynomials f, g and suppose that

$$F = \begin{pmatrix} 1 & 0 \\ f & 1 \end{pmatrix} + \varepsilon \begin{pmatrix} 0 & 0 \\ f' & 0 \end{pmatrix} + \mathcal{O}(\varepsilon^2) \tag{7.4}$$

$$G = \begin{pmatrix} 1 & 0 \\ g & 1 \end{pmatrix} + \varepsilon \begin{pmatrix} 0 & 0 \\ g' & 0 \end{pmatrix} + \mathcal{O}(\varepsilon^2) \tag{7.5}$$

are products of primitives for some $f', g' \in \mathbb{F}[\mathbf{x}]$. Then

$$G \cdot F = \begin{pmatrix} 1 & 0 \\ f+g & 1 \end{pmatrix} + \varepsilon \begin{pmatrix} 0 & 0 \\ f'+g' & 0 \end{pmatrix} + \mathcal{O}(\varepsilon^2)$$

is a product of primitives satisfying (7.3).

Suppose $h = fg$ is a product of two polynomials and suppose that F and G are of the form (7.4) and (7.5) and are products of primitives. We will construct $M((f+g)^2)$, $M(-f^2)$, $M(-g^2)$ approximately in such a way that when we use the identity $(f+g)^2 - f^2 - g^2 = 2fg$, the error terms cancel properly. Here we will use that $\mathrm{char}(\mathbb{F}) \neq 2$. Define the expressions $\mathsf{sq}_+(A)$ and $\mathsf{sq}_-(A)$ by

$$\mathsf{sq}_\pm(A) \coloneqq \begin{pmatrix} -\varepsilon & 0 \\ 0 & 1 \end{pmatrix} \cdot A \cdot \begin{pmatrix} -1 & \pm\varepsilon \\ 0 & 1 \end{pmatrix} \cdot A \cdot \begin{pmatrix} \frac{1}{\varepsilon} & 0 \\ 0 & 1 \end{pmatrix}.$$

Then

$$\mathsf{sq}_\pm(F) = \begin{pmatrix} 1 \mp \varepsilon f & 0 \\ \pm f^2 + \mathcal{O}(\varepsilon) & 1 \pm \varepsilon f \end{pmatrix} + \mathcal{O}(\varepsilon^2).$$

We have

$$\begin{aligned} &\mathsf{sq}_-(F) \cdot \mathsf{sq}_-(G) \cdot \mathsf{sq}_+(G \cdot F) \\ &\quad = \begin{pmatrix} 1 + \varepsilon g & 0 \\ -g^2 + \mathcal{O}(\varepsilon) & 1 - \varepsilon g \end{pmatrix} \\ &\qquad \cdot \begin{pmatrix} 1 + \varepsilon f & 0 \\ -f^2 + \mathcal{O}(\varepsilon) & 1 - \varepsilon f \end{pmatrix} \\ &\qquad \cdot \begin{pmatrix} 1 - \varepsilon(f+g) & 0 \\ (f+g)^2 + \mathcal{O}(\varepsilon) & 1 + \varepsilon(f+g) \end{pmatrix} + \mathcal{O}(\varepsilon^2) \end{aligned}$$

which simplifies to

$$\mathsf{sq}_-(F) \cdot \mathsf{sq}_-(G) \cdot \mathsf{sq}_+(G \cdot F) = \begin{pmatrix} 1 & 0 \\ 2fg + \mathcal{O}(\varepsilon) & 1 \end{pmatrix} + \mathcal{O}(\varepsilon^2).$$

We conclude

$$\begin{aligned}
&\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \cdot \mathsf{sq}_{-}(G) \cdot \mathsf{sq}_{-}(F) \cdot \mathsf{sq}_{+}(G \cdot F) \cdot \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & 1 \end{pmatrix} \\
&= \begin{pmatrix} -2\varepsilon & 0 \\ 0 & 1 \end{pmatrix} \cdot G \cdot \begin{pmatrix} -1 & -\varepsilon \\ 0 & 1 \end{pmatrix} \cdot G \cdot \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \cdot F \cdot \begin{pmatrix} -1 & -\varepsilon \\ 0 & 1 \end{pmatrix} F \\
&\quad \cdot \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \cdot G \cdot F \cdot \begin{pmatrix} -1 & \varepsilon \\ 0 & 1 \end{pmatrix} \cdot G \cdot F \cdot \begin{pmatrix} \frac{1}{2\varepsilon} & 0 \\ 0 & 1 \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 \\ fg + \mathcal{O}(\varepsilon) & 1 \end{pmatrix} + \mathcal{O}(\varepsilon^2).
\end{aligned}$$

This completes the construction.

The length m of the construction is $m(h) = 1$ for h a variable or constant and recursively $m(f+g) = m(f) + m(g)$, $m(f \cdot g) = 4(m(f) + m(g)) + 7$. We conclude $m \in \mathcal{O}(8^d)$. The error degree e of the construction satisfies the same recursion, so $e \in \mathcal{O}(8^d)$. □

Remark 7.9. The construction in the above proof of Theorem 7.8 is different from the construction in our paper [BIZ17]. The recursion in the above proof is simpler, while the construction in [BIZ17] has a better error degree and has a special form which relates it to a family of polynomials called continuants.

Corollary 7.10. $\overline{\mathbf{VP}_2} = \overline{\mathbf{VP}_\mathsf{e}}$ *and* $\overline{\mathbf{VP}_2}^{\mathrm{poly}} = \overline{\mathbf{VP}_\mathsf{e}}^{\mathrm{poly}}$ *when* $\mathrm{char}(\mathbb{F}) \neq 2$.

Proof. We have $\mathbf{VP}_2 \subseteq \mathbf{VP}_\mathsf{e}$ by Lemma 7.2. Taking closures on both sides, we obtain $\overline{\mathbf{VP}_2} \subseteq \overline{\mathbf{VP}_\mathsf{e}}$ and $\overline{\mathbf{VP}_2}^{\mathrm{poly}} \subseteq \overline{\mathbf{VP}_\mathsf{e}}^{\mathrm{poly}}$.

When $\mathrm{char}(\mathbb{F}) \neq 2$, $\mathbf{VP}_\mathsf{e} \subseteq \overline{\mathbf{VP}_2}^{\mathrm{poly}}$ (Theorem 7.8). By taking closures follows $\overline{\mathbf{VP}_\mathsf{e}} \subseteq \overline{\mathbf{VP}_2}$ and $\overline{\mathbf{VP}_\mathsf{e}}^{\mathrm{poly}} \subseteq \overline{\mathbf{VP}_2}^{\mathrm{poly}}$. □

Corollary 7.11. $\overline{\mathbf{VP}_2}^{\mathrm{poly}} = \mathbf{VP}_\mathsf{e}$ *when* $\mathrm{char}(\mathbb{F}) \neq 2$ *and* $\mathbb{F}$ *is infinite.*

Proof. By Corollary 7.10 $\overline{\mathbf{VP}_2}^{\mathrm{poly}} = \overline{\mathbf{VP}_\mathsf{e}}^{\mathrm{poly}}$. We prove $\overline{\mathbf{VP}_\mathsf{e}}^{\mathrm{poly}} = \mathbf{VP}_\mathsf{e}$ in Lemma 7.12 below. □

Lemma 7.12. $\overline{\mathbf{VP}_\mathsf{e}}^{\mathrm{poly}} = \mathbf{VP}_\mathsf{e}$ *when* $\mathrm{char}(\mathbb{F}) \neq 2$ *and* $\mathbb{F}$ *is infinite.*

Proof. The inclusion $\mathbf{VP}_\mathsf{e} \subseteq \overline{\mathbf{VP}_\mathsf{e}}^{\mathrm{poly}}$ is trivially true. We prove the other direction. Let $(f_n) \in \overline{\mathbf{VP}_\mathsf{e}}^{\mathrm{poly}}$. Then there are polynomials $f_{n;i}(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ and $e(n) \in \mathrm{poly}(n)$ such that

$$f_n(\mathbf{x}) + \varepsilon f_{n;1}(\mathbf{x}) + \varepsilon^2 f_{n;2}(\mathbf{x}) + \cdots + \varepsilon^{e(n)} f_{n;e(n)}(\mathbf{x})$$

is computed by a poly-size formula Γ over $\mathbb{F}(\varepsilon)$. Let $\alpha_0, \alpha_1, \ldots, \alpha_{e(n)}$ be distinct elements in $\mathbb{F}$ such that replacing ε by α_j in Γ is a valid substitution, i.e. not causing division by zero. These α_j exist since our field is infinite by assumption. View

$$g_n(\varepsilon) \coloneqq f_n(\mathbf{x}) + \varepsilon f_{n;1}(\mathbf{x}) + \varepsilon^2 f_{n;2}(\mathbf{x}) + \cdots + \varepsilon^{e(n)} f_{n;e(n)}(\mathbf{x})$$

as a polynomial in ε. The polynomial $g_n(\varepsilon)$ has degree at most $e(n)$ so we can write $g_n(\varepsilon)$ as follows (Lagrange interpolation on $e(n)+1$ points)

$$g_n(\varepsilon) = \sum_{j=0}^{e(n)} g_n(\alpha_j) \prod_{\substack{0 \le m \le e(n): \\ m \ne j}} \frac{\varepsilon - \alpha_m}{\alpha_j - \alpha_m}. \tag{7.6}$$

Clearly, $f_n(\mathbf{x}) = g_n(0)$. However, replacing ε by 0 in Γ is not a valid substitution in general. From (7.6) we see directly how to write $g_n(0)$ as a linear combination of the values $g_n(\alpha_j)$, namely

$$g_n(0) = \sum_{j=0}^{e(n)} g_n(\alpha_j) \prod_{\substack{0 \le m \le e(n): \\ m \ne j}} \frac{-\alpha_m}{\alpha_j - \alpha_m},$$

that is,

$$g_n(0) = \sum_{j=0}^{e(n)} \beta_j \, g_n(\alpha_j) \quad \text{with} \quad \beta_j \coloneqq \prod_{\substack{0 \le m \le e(n): \\ m \ne j}} \frac{\alpha_m}{\alpha_m - \alpha_j}.$$

The value $g_n(\alpha_j)$ is computed by the formula Γ with ε replaced by α_j, which we denote by $\Gamma|_{\varepsilon=\alpha_j}$. Thus $f_n(\mathbf{x})$ is computed by the poly-size formula $\sum_{j=0}^{e(n)} \beta_j \, \Gamma|_{\varepsilon=\alpha_j}$. We conclude $(f_n) \in \mathbf{VP}_{\mathsf{e}}$. □

Remark 7.13. The statement of Lemma 7.12 also holds with $\mathbf{VP}_{\mathsf{e}}$ replaced with $\mathbf{VP}_{\mathsf{s}}$ or with $\mathbf{VP}$ by a similar proof.

7.4 Nondeterminism closure of $\mathbf{VP}_1$

Recall the definition of $\mathbf{VNP}_{\mathsf{x}} = \mathrm{N}(\mathbf{VP}_{\mathsf{x}})$ from Definition 7.7. Valiant proved the following characterisation of $\mathbf{VNP}$ in his seminal work [Val80]. See also [BCS97, Thm. 21.26], [Bür00, Thm. 2.13] and [MP08, Thm. 2].

Theorem 7.14 (Valiant [Val80])**.** $\mathbf{VNP}_{\mathsf{e}} = \mathbf{VNP}$.

We strengthen Valiant's characterisation of $\mathbf{VNP}$ from $\mathbf{VNP}_{\mathsf{e}}$ to $\mathbf{VNP}_1$.

Theorem 7.15. $\mathbf{VNP}_1 = \mathbf{VNP}$ *when* $\mathrm{char}(\mathbb{F}) \neq 2$.

The idea of the proof is "to simulate in $\mathbf{VNP}_1$" the primitives that we used in the proof of $\mathbf{VP}_\mathsf{e} \subseteq \mathbf{VP}_3$ (Theorem 7.3).

Proof of Theorem 7.15. Clearly, $\mathbf{VNP}_1 \subseteq \mathbf{VNP}$ by Lemma 7.2 and taking the nondeterminism closure N. We will prove that $\mathbf{VNP} \subseteq \mathbf{VNP}_1$. Recall that in the proof of $\mathbf{VP}_\mathsf{e} \subseteq \mathbf{VP}_3$ (Theorem 7.3), we defined for any polynomial h the matrix

$$M(h) := \begin{pmatrix} 1 & 0 & 0 \\ h & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

and we called the following matrices *primitives*:

- $M(h)$ with h any variable or any constant in $\mathbb{F}$
- the 3×3 permutation matrices, denoted by M_π for $\pi \in S_3$
- the diagonal matrices $M_{a,b,c} := \operatorname{diag}(a, b, c)$ with $a, b, c \in \mathbb{F}$.

In the proof of $\mathbf{VP}_\mathsf{e} \subseteq \mathbf{VP}_3$ we constructed, for any family $(f_n) \in \mathbf{VP}_\mathsf{e}$, a sequence of primitive matrices $A_{n,1}, \ldots, A_{n,t(n)}$ with $t(n) \in \operatorname{poly}(n)$ such that

$$f_n(\mathbf{x}) = \left(\begin{smallmatrix} 1 & 1 & 1 \end{smallmatrix}\right) M_{-1,1,0} A_{n,1} \cdots A_{n,t(n)} \left(\begin{smallmatrix} 1 \\ 1 \\ 1 \end{smallmatrix}\right). \tag{7.7}$$

We will show $\mathbf{VP}_\mathsf{e} \subseteq \mathbf{VNP}_1$ by constructing a hypercube sum over a width-1 abp that evaluates the right-hand side of (7.7). This implies $\mathbf{VNP}_\mathsf{e} \subseteq \mathbf{VNP}_1$ by taking the N-closure. Then by Valiant's Theorem 7.14, $\mathbf{VNP} \subseteq \mathbf{VNP}_1$.

Let $f(\mathbf{x})$ be a polynomial and let $A_1, \ldots, A_k$ be primitive matrices such that $f(\mathbf{x})$ is computed as

$$f(\mathbf{x}) = \left(\begin{smallmatrix} 1 & 1 & 1 \end{smallmatrix}\right) A_k \cdots A_1 \left(\begin{smallmatrix} 1 \\ 1 \\ 1 \end{smallmatrix}\right).$$

View this expression as a width-3 abp G, with vertex layers labeled as shown in the left-hand diagram in Fig. 7.1. Assume for simplicity that all edges between layers are present, possibly with label 0. The sum of the values of every s–t path in G equals $f(\mathbf{x})$,

$$f(\mathbf{x}) = \sum_{j \in [3]^k} A_k[j_k, j_{k-1}] \cdots A_1[j_2, j_1]. \tag{7.8}$$

We introduce some hypercube variables. To every vertex of G, except s and t, we associate a bit; the bits in the ith layer we call $b_1[i]$, $b_2[i]$, $b_3[i]$. To an s–t path in G we associate an assignment of the $b_j[i]$ by setting the bits of vertices visited by the path to 1 and the others to 0. For example, in the right-hand

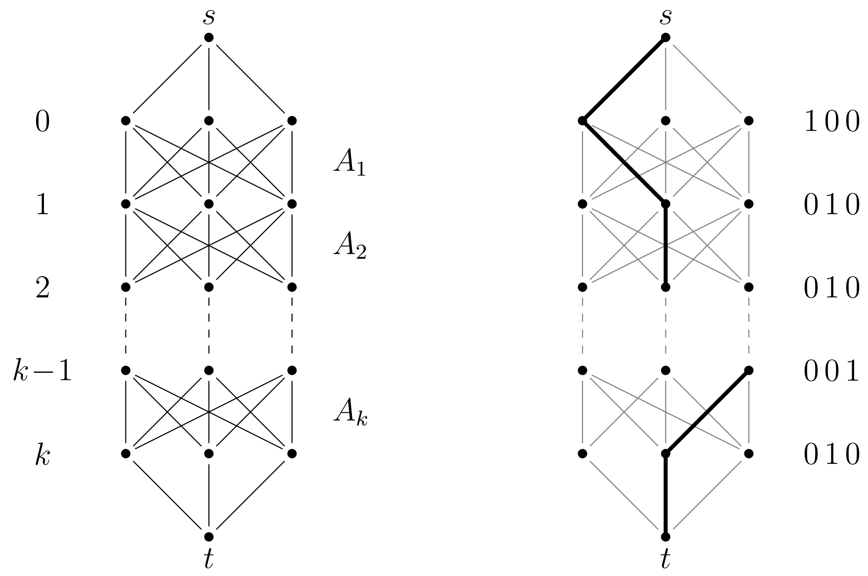


Figure 7.1: Illustration of the layer labelling and the path labelling used in the proof of Theorem 7.15.

diagram in Fig. 7.1 we show an s–t path with the corresponding assignment of the bits $b_j[i]$. The assignments of the $b_j[i]$ corresponding to s–t paths are precisely the assignments such that for every $i \in [k]$ exactly one of $b_1[i]$, $b_2[i]$, $b_3[i]$ equals 1. Let

$$V(b_1, b_2, b_3) \coloneqq \prod_{i\in[k]} \big(b_1[i] + b_2[i] + b_3[i]\big) \prod_{\substack{s,t\in[3]:\\ s\neq t}} \big(1 - b_s[i]b_t[i]\big). \tag{7.9}$$

Then the assignments of the $b_j[i]$ corresponding to s–t paths are precisely the assignments such that $V(b_1, b_2, b_3) = 1$. Otherwise, $V(b_1, b_2, b_3) = 0$.

We will write $f(\mathbf{x})$ as a hypercube sum by replacing each $A_i[j_i, j_{i-1}]$ in (7.8) by a product of affine linear forms $S_i(A_i)$ with variables $\mathbf{b}$ and $\mathbf{x}$,

$$\sum_{\mathbf{b}} V(b_1, b_2, b_3) S_k(A_k) \cdots S_1(A_1).$$

Define the expression $\mathsf{Eq}(\alpha, \beta) \coloneqq (1 - \alpha - \beta)(1 - \alpha - \beta)$ for $\alpha, \beta \in \{0, 1\}$. The expression $\mathsf{Eq}(\alpha, \beta)$ evaluates to 1 if α equals β and evaluates to 0 otherwise.

- For any variable or constant x define

$$\begin{aligned} S_i(M(x)) \coloneqq {} & \big(1 + (x-1)(b_1[i] - b_1[i-1])\big) \\ & \cdot \big(1 - (1 - b_2[i])b_2[i-1]\big) \\ & \cdot \mathsf{Eq}\big(b_3[i-1], b_3[i]\big). \end{aligned}$$

- For any permutation $\pi \in S_3$ define

$$\begin{aligned} S_i(M_\pi) \coloneqq{} & \mathsf{Eq}\big(b_1[i-1], b_{\pi(1)}[i]\big) \\ & \cdot \mathsf{Eq}\big(b_2[i-1], b_{\pi(2)}[i]\big) \\ & \cdot \mathsf{Eq}\big(b_3[i-1], b_{\pi(3)}[i]\big). \end{aligned}$$

- For any constants $a, b, c \in \mathbb{F}$ define

$$\begin{aligned} S_i(M_{a,b,c}) \coloneqq{} & \big(a \cdot b_1[i-1] + b \cdot b_2[i-1] + c \cdot b_3[i-1]\big) \\ & \cdot \mathsf{Eq}\big(b_1[i-1], b_1[i]\big) \\ & \cdot \mathsf{Eq}\big(b_2[i-1], b_2[i]\big) \\ & \cdot \mathsf{Eq}\big(b_3[i-1], b_3[i]\big). \end{aligned}$$

One verifies that

$$f(\mathbf{x}) = \sum_{\mathbf{b}} V(b_1, b_2, b_3) S_k(A_k) \cdots S_1(A_1).$$

Some of the factors in the expressions for the $S_i(A_i)$ are not affine linear. As a final step we apply the equality $1 + xy = \frac{1}{2} \sum_{c \in \{0,1\}} (x + 1 - 2c)(y + 1 - 2c)$ to write these factors as products of affine linear forms, introducing new hypercube variables. $\square$

7.5 Conclusion

We finish with an overview of inclusions, equalities and separations among the classes $\mathbf{VP}_k$, $\mathbf{VP}_\mathsf{e}$, $\mathbf{VP}$ and their approximation and nondeterminism closures (when $\mathrm{char}(\mathbb{F}) \neq 2$), see Fig. 7.2. The figure relies on the following two simple lemmas, proofs of which can be found in our paper [BIZ17].

Lemma 7.16 ([BIZ17, Prop. 5.10])**.** $\mathbf{VP}_1 = \overline{\mathbf{VP}_1}$.

Lemma 7.17 ([BIZ17, Prop. 5.11])**.** $\mathbf{VP}_1 \subsetneq \mathbf{VNP}_1$ *when* $\mathrm{char}(\mathbb{F}) \neq 2$.

$$\begin{array}{ccccccccc}
 & \overline{\mathbf{VP}_1} & \subsetneq & \overline{\mathbf{VP}_2} & \overset{7.10}{=} & \overline{\mathbf{VP}_\mathsf{e}} & \subseteq & \overline{\mathbf{VP}} \\
 \text{7.16} & \parallel & & \rotatebox{90}{$\subsetneq$} & & \rotatebox{90}{$\subseteq$} & & \rotatebox{90}{$\subseteq$} \\
 & \mathbf{VP}_1 & \subsetneq & \mathbf{VP}_2 & \overset{\text{[AW16]}}{\subsetneq} & \mathbf{VP}_\mathsf{e} & \overset{\text{[Val79]}}{\subseteq} & \mathbf{VP} \\
 \text{7.17} & \rotatebox{-90}{$\subsetneq$} & & \rotatebox{-90}{$\subsetneq$} & & \rotatebox{-90}{$\subseteq$} & & \rotatebox{-90}{$\subseteq$} \\
 & \mathbf{VNP}_1 & = & \mathbf{VNP}_2 & \overset{7.15}{=} & \mathbf{VNP}_\mathsf{e} & \overset{\text{[Val80]}}{=} & \mathbf{VNP}
\end{array}$$

Figure 7.2: Overview of relations among the algebraic complexity classes $\mathbf{VP}_k$, $\mathbf{VP}_\mathsf{e}$, $\mathbf{VP}$ and their approximation and nondeterminism closures (when $\mathrm{char}(\mathbb{F})$ is not 2). The relations without reference are either by definition or follow directly from the other relations.

Bibliography

[AJRS13] Elizabeth S. Allman, Peter D. Jarvis, John A. Rhodes, and Jeremy G. Sumner. Tensor rank, invariants, inequalities, and applications. *SIAM J. Matrix Anal. Appl.*, 34(3):1014–1045, 2013. doi:10.1137/120899066. p. 15

[Alo98] Noga Alon. The Shannon capacity of a union. *Combinatorica*, 18(3):301–310, 1998. doi:10.1007/PL00009824. p. 39

[ASU13] Noga Alon, Amir Shpilka, and Christopher Umans. On sunflowers and matrix multiplication. *Comput. Complexity*, 22(2):219–243, Jun 2013. doi:10.1007/s00037-013-0060-1. p. 50

[AW16] Eric Allender and Fengming Wang. On the power of algebraic branching programs of width two. *Comput. Complexity*, 25(1):217–253, 2016. doi:10.1007/s00037-015-0114-7. p. 17, 111, 116, 125

[AZ14] Martin Aigner and Günter M. Ziegler. *Proofs from The Book*. Springer-Verlag, Berlin, fifth edition, 2014. doi:10.1007/978-3-662-44205-0. p. 73

[BC18] Boris Bukh and Christopher Cox. On a fractional version of Haemers' bound. *arXiv*, 2018. arXiv:1802.00476. p. 43, 44

[BCC+17] Jonah Blasiak, Thomas Church, Henry Cohn, Joshua A. Grochow, Eric Naslund, William F. Sawin, and Chris Umans. On cap sets and the group-theoretic approach to matrix multiplication. *Discrete Anal.*, 2017. arXiv:1605.06702, doi:10.19086/da.1245. p. 50, 85, 86, 106

[BCRL79] Dario Bini, Milvio Capovani, Francesco Romani, and Grazia Lotti. $\mathrm{O}(n^{2.7799})$ complexity for $n \times n$ approximate matrix multiplication. *Inf. Process. Lett.*, 8(5):234–235, 1979. doi:10.1016/0020-0190(79)90113-3. p. 3, 112

[BCS97] Peter Bürgisser, Michael Clausen, and M. Amin Shokrollahi. *Algebraic complexity theory*, volume 315 of *Grundlehren Math. Wiss.* Springer-Verlag, Berlin, 1997. doi:10.1007/978-3-662-03338-8. p. 4, 5, 34, 50, 52, 68, 81, 121

[BCSX10] Arnab Bhattacharyya, Victor Chen, Madhu Sudan, and Ning Xie. *Testing Linear-Invariant Non-linear Properties: A Short Report*, pages 260–268. Springer Berlin Heidelberg, Berlin, Heidelberg, 2010. doi:10.1007/978-3-642-16367-8_18. p. 50

[BCZ17a] Markus Bläser, Matthias Christandl, and Jeroen Zuiddam. The border support rank of two-by-two matrix multiplication is seven. *arXiv*, 2017. arXiv:1705.09652. p. 15

[BCZ17b] Harry Buhrman, Matthias Christandl, and Jeroen Zuiddam. Nondeterministic Quantum Communication Complexity: the Cyclic Equality Game and Iterated Matrix Multiplication. In Christos H. Papadimitriou, editor, *8th Innovations in Theoretical Computer Science Conference (ITCS 2017)*, pages 24:1–24:18, 2017. arXiv:1603.03757, doi:10.4230/LIPIcs.ITCS.2017.24. p. 15

[Ber84] Stuart J. Berkowitz. On computing the determinant in small parallel time using a small number of processors. *Inform. Process. Lett.*, 18(3):147–150, 1984. doi:10.1016/0020-0190(84)90018-8. p. 110

[BI13] Peter Bürgisser and Christian Ikenmeyer. Explicit lower bounds via geometric complexity theory. *Proceedings 45th Annual ACM Symposium on Theory of Computing 2013*, pages 141–150, 2013. doi:10.1145/2488608.2488627. p. 110

[Bin80] Dario Bini. Relations between exact and approximate bilinear algorithms. Applications. *Calcolo*, 17(1):87–97, 1980. doi:10.1007/BF02575865. p. 3

[BIZ17] Karl Bringmann, Christian Ikenmeyer, and Jeroen Zuiddam. On Algebraic Branching Programs of Small Width. In Ryan O'Donnell, editor, *32nd Computational Complexity Conference (CCC 2017)*, pages 20:1–20:31, 2017. doi:10.4230/LIPIcs.CCC.2017.20. p. 1, 109, 113, 114, 120, 124

[Blä13a] Markus Bläser. *Fast Matrix Multiplication.* Number 5 in Graduate Surveys. Theory of Computing Library, 2013. doi:10.4086/toc.gs.2013.005. p. 34

[Bla13b] Anna Blasiak. *A graph-theoretic approach to network coding.* PhD thesis, Cornell University, 2013. URL: `https://ecommons.cornell.edu/bitstream/handle/1813/34147/ab675.pdf`. p. 44

[BLMW11] Peter Bürgisser, Joseph M. Landsberg, Laurent Manivel, and Jerzy Weyman. An overview of mathematical issues arising in the geometric complexity theory approach to VP $\neq$ VNP. *SIAM J. Comput.*, 40(4):1179–1209, 2011. `doi:10.1137/090765328`. p. 110

[BOC92] Michael Ben-Or and Richard Cleve. Computing algebraic formulas using a constant number of registers. *SIAM J. Comput.*, 21(1):54–58, 1992. `doi:10.1137/0221006`. p. 17, 111, 114

[BPR$^+$00] Charles H. Bennett, Sandu Popescu, Daniel Rohrlich, John A. Smolin, and Ashish V. Thapliyal. Exact and asymptotic measures of multipartite pure-state entanglement. *Phys. Rev. A*, 63(1):012307, 2000. `doi:10.1103/PhysRevA.63.012307`. p. 50

[Bre74] Richard P. Brent. The parallel evaluation of general arithmetic expressions. *J. ACM*, 21(2):201–206, April 1974. `doi:10.1145/321812.321815`. p. 114

[Bri87] Michel Brion. Sur l'image de l'application moment. In *Séminaire d'algèbre Paul Dubreil et Marie-Paule Malliavin (Paris, 1986)*, volume 1296 of *Lecture Notes in Math.*, pages 177–192. Springer, Berlin, 1987. `doi:10.1007/BFb0078526`. p. 9, 95, 96

[BS83] Eberhard Becker and Niels and Schwartz. Zum Darstellungssatz von Kadison-Dubois. *Arch. Math. (Basel)*, 40(5):421–428, 1983. `doi:10.1007/BF01192806`. p. 7, 12, 34

[Bür90] Peter Bürgisser. *Degenerationsordnung und Trägerfunktional bilinearer Abbildungen.* PhD thesis, Universität Konstanz, 1990. `http://nbn-resolving.de/urn:nbn:de:bsz:352-opus-20311`. p. 59, 103

[Bür00] Peter Bürgisser. *Completeness and reduction in algebraic complexity theory*, volume 7 of *Algorithms and Computation in Mathematics.* Springer-Verlag, Berlin, 2000. `doi:10.1007/978-3-662-04179-6`. p. 121

[Bür04] Peter Bürgisser. The complexity of factors of multivariate polynomials. *Found. Comput. Math.*, 4(4):369–396, 2004. `doi:10.1007/s10208-002-0059-5`. p. 112, 117

[BX15] Arnab Bhattacharyya and Ning Xie. Lower bounds for testing triangle-freeness in boolean functions. *Comput. Complexity*, 24(1):65–101, 2015. `doi:10.1007/s00037-014-0092-1`. p. 50

[CHM07] Matthias Christandl, Aram W. Harrow, and Graeme Mitchison. Nonzero Kronecker coefficients and what they tell us about spectra. *Comm. Math. Phys.*, 270(3):575–585, 2007. `doi:10.1007/s00220-006-0157-3`. p. 92

[CJZ18] Matthias Christandl, Asger Kjærulff Jensen, and Jeroen Zuiddam. Tensor rank is not multiplicative under the tensor product. *Linear Algebra Appl.*, 543:125–139, 2018. `doi:10.1016/j.laa.2017.12.020`. p. 15

[CKSV16] Suryajith Chillara, Mrinal Kumar, Ramprasad Saptharishi, and V. Vinay. The chasm at depth four, and tensor rank: Old results, new insights. *arXiv*, 2016. `arXiv:1606.04200`. p. 15

[CLP17] Ernie Croot, Vsevolod F. Lev, and Péter Pál Pach. Progression-free sets in $\mathbf{Z}_4^n$ are exponentially small. *Ann. of Math. (2)*, 185(1):331–337, 2017. `doi:10.4007/annals.2017.185.1.7`. p. 50, 83

[CM06] Matthias Christandl and Graeme Mitchison. The spectra of quantum states and the Kronecker coefficients of the symmetric group. *Comm. Math. Phys.*, 261(3):789–797, 2006. `doi:10.1007/s00220-005-1435-1`. p. 93

[CMR+14] Toby Cubitt, Laura Mančinska, David E. Roberson, Simone Severini, Dan Stahlke, and Andreas Winter. Bounds on entanglement-assisted source-channel coding via the Lovász theta number and its variants. *IEEE Trans. Inform. Theory*, 60(11):7330–7344, 2014. `arXiv:1310.7120`, `doi:10.1109/TIT.2014.2349502`. p. 44

[CT12] Thomas M. Cover and Joy A. Thomas. *Elements of information theory*. John Wiley & Sons, 2012. p. 62

[CU13] Henry Cohn and Christopher Umans. Fast matrix multiplication using coherent configurations. In *Proceedings of the Twenty-Fourth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1074–1086. SIAM, 2013. p. 15

[CVZ16] Matthias Christandl, Péter Vrana, and Jeroen Zuiddam. Asymptotic tensor rank of graph tensors: beyond matrix multiplication. *arXiv*, 2016. arXiv:1609.07476. p. 1, 67, 69, 81, 87

[CVZ18] Matthias Christandl, Péter Vrana, and Jeroen Zuiddam. Universal points in the asymptotic spectrum of tensors. In *Proceedings of 50th Annual ACM SIGACT Symposium on the Theory of Computing (STOC'18)*. ACM, New York, 2018. arXiv:1709.07851, doi:10.1145/3188745.3188766. p. 1, 49, 67, 89, 90, 98, 105, 107

[CW82] Don Coppersmith and Shmuel Winograd. On the asymptotic complexity of matrix multiplication. *SIAM J. Comput.*, 11(3):472–492, 1982. doi:10.1137/0211038. p. 3

[CW87] Don Coppersmith and Shmuel Winograd. Matrix multiplication via arithmetic progressions. In *Proceedings of the nineteenth annual ACM symposium on Theory of computing*, pages 1–6. ACM, 1987. p. 3

[CW90] Don Coppersmith and Shmuel Winograd. Matrix multiplication via arithmetic progressions. *J. Symbolic Comput.*, 9(3):251–280, 1990. doi:10.1016/S0747-7171(08)80013-2. p. 4, 6, 8, 10, 50, 69

[CZ18] Matthias Christandl and Jeroen Zuiddam. Tensor surgery and tensor rank. *Comput. Complexity*, Mar 2018. doi:10.1007/s00037-018-0164-8. p. 88

[Dra15] Jan Draisma. Multilinear Algebra and Applications (lecture notes), 2015. URL: https://mathsites.unibe.ch/jdraisma/publications/mlappl.pdf. p. 15

[DVC00] Wolfgang Dür, Guivre Vidal, and Juan Ignacio Cirac. Three qubits can be entangled in two inequivalent ways. *Phys. Rev. A (3)*, 62(6):062314, 12, 2000. doi:10.1103/PhysRevA.62.062314. p. 50

[Ede04] Yves Edel. Extensions of generalized product caps. *Des. Codes Cryptogr.*, 31(1):5–14, 2004. doi:10.1023/A:1027365901231. p. 50, 85

[EG17] Jordan S. Ellenberg and Dion Gijswijt. On large subsets of $\mathbf{F}_q^n$ with no three-term arithmetic progression. *Ann. of Math. (2)*, 185(1):339–343, 2017. doi:10.4007/annals.2017.185.1.8. p. 10, 50, 83, 85, 86

[FK14] Hu Fu and Robert Kleinberg. Improved lower bounds for testing triangle-freeness in boolean functions via fast matrix multiplication.

In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2014)*, pages 669–676, 2014. doi:10.4230/LIPIcs.APPROX-RANDOM.2014.669. p. 50

[For16] Michael Forbes. Some concrete questions on the border complexity of polynomials. Presentation given at the Workshop on Algebraic Complexity Theory WACT 2016 in Tel Aviv, https://www.youtube.com/watch?v=1HMogQIHT6Q, 2016. p. 112

[Fra02] Matthias Franz. Moment polytopes of projective G-varieties and tensor products of symmetric group representations. *J. Lie Theory*, 12(2):539–549, 2002. URL: http://emis.ams.org/journals/JLT/vol.12_no.2/16.html. p. 95, 96

[Fri17] Tobias Fritz. Resource convertibility and ordered commutative monoids. *Math. Structures Comput. Sci.*, 27(6):850–938, 2017. doi:10.1017/S0960129515000444. p. 39

[Ful97] William Fulton. *Young tableaux*, volume 35 of *London Mathematical Society Student Texts*. Cambridge University Press, Cambridge, 1997. With applications to representation theory and geometry. p. 90

[GKKS13] Ankit Gupta, Pritish Kamath, Neeraj Kayal, and Ramprasad Saptharishi. Approaching the chasm at depth four. In *2013 IEEE Conference on Computational Complexity—CCC 2013*, pages 65–73. IEEE Computer Soc., Los Alamitos, CA, 2013. doi:10.1109/CCC.2013.16. p. 110

[GMQ16] Joshua A. Grochow, Ketan D. Mulmuley, and Youming Qiao. Boundaries of VP and VNP. In Ioannis Chatzigiannakis, Michael Mitzenmacher, Yuval Rabani, and Davide Sangiorgi, editors, *43rd International Colloquium on Automata, Languages, and Programming (ICALP 2016)*, volume 55, pages 34:1–34:14, 2016. arXiv:1605.02815, doi:10.4230/LIPIcs.ICALP.2016.34. p. 112

[Gro13] Joshua A. Grochow. Unifying and generalizing known lower bounds via geometric complexity theory. *arXiv*, 2013. arXiv:1304.6333. p. 110

[GW09] Roe Goodman and Nolan R. Wallach. *Symmetry, representations, and invariants*, volume 255 of *Graduate Texts in Mathematics*. Springer, Dordrecht, 2009. doi:10.1007/978-0-387-79852-3. p. 90

[Hae79] Willem Haemers. On some problems of Lovász concerning the Shannon capacity of a graph. *IEEE Trans. Inform. Theory*, 25(2):231–232, 1979. doi:10.1109/TIT.1979.1056027. p. 39, 42, 44

[Hås90] Johan Håstad. Tensor rank is NP-complete. *J. Algorithms*, 11(4):644–654, 1990. doi:10.1016/0196-6774(90)90014-6. p. 49

[HHHH09] Ryszard Horodecki, Paweł Horodecki, Michał Horodecki, and Karol Horodecki. Quantum entanglement. *Rev. Modern Phys.*, 81(2):865–942, 2009. doi:10.1103/RevModPhys.81.865. p. 50

[HIL13] Jonathan D. Hauenstein, Christian Ikenmeyer, and Joseph M. Landsberg. Equations for lower bounds on border rank. *Exp. Math.*, 22(4):372–383, 2013. doi:10.1080/10586458.2013.825892. p. 15, 112

[Hum75] James E. Humphreys. *Linear algebraic groups*. Springer-Verlag, New York-Heidelberg, 1975. Graduate Texts in Mathematics, No. 21. p. 95

[HX17] Ishay Haviv and Ning Xie. Sunflowers and testing triangle-freeness of functions. *Comput. Complexity*, 26(2):497–530, Jun 2017. doi:10.1007/s00037-016-0138-7. p. 50

[Ike13] Christian Ikenmeyer. *Geometric complexity theory, tensor rank, and Littlewood–Richardson coefficients*. PhD thesis, Universität Paderborn, 2013. p. 14

[Kar72] Richard M. Karp. Reducibility among combinatorial problems. In *Complexity of computer computations (Proc. Sympos., IBM Thomas J. Watson Res. Center, Yorktown Heights, N.Y., 1972)*, pages 85–103. Plenum, New York, 1972. p. 38

[Knu94] Donald E. Knuth. The sandwich theorem. *Electron. J. Combin.*, 1, 1994. URL: http://www.combinatorics.org/Volume_1/Abstracts/v1i1a1.html. p. 43

[Kra84] Hanspeter Kraft. *Geometrische Methoden in der Invariantentheorie*. Springer, 1984. doi:10.1007/978-3-663-10143-7. p. 52, 90, 95

[KS08] Tali Kaufman and Madhu Sudan. Algebraic property testing: The role of invariance. In *Proceedings of the Fortieth Annual ACM Symposium on Theory of Computing*, STOC '08, pages 403–412, New York, NY, USA, 2008. ACM. doi:10.1145/1374376.1374434. p. 50

[KSS16] Robert Kleinberg, William F. Sawin, and David E. Speyer. The growth rate of tri-colored sum-free sets. *arXiv*, 2016. `arXiv:1607.00047`. p. 50, 81, 85

[Lan06] Joseph M. Landsberg. The border rank of the multiplication of 2×2 matrices is seven. *J. Amer. Math. Soc.*, 19(2):447–459, 2006. `doi:10.1090/S0894-0347-05-00506-0`. p. 112

[LG14] François Le Gall. Powers of tensors and fast matrix multiplication. In *ISSAC 2014—Proceedings of the 39th International Symposium on Symbolic and Algebraic Computation*, pages 296–303. ACM, New York, 2014. `doi:10.1145/2608628.2608664`. p. 4, 6, 8, 50, 87

[Lic84] Thomas Lickteig. A note on border rank. *Inf. Process. Lett.*, 18(3):173–178, 1984. `doi:10.1016/0020-0190(84)90023-1`. p. 112

[LM16a] Joseph M. Landsberg and Mateusz Michałek. A $2n^2 - \log(n) - 1$ lower bound for the border rank of matrix multiplication. *arXiv*, 2016. `arXiv:1608.07486`. p. 112

[LM16b] Joseph M. Landsberg and Mateusz Michałek. Abelian tensors. *J. Math. Pures Appl.*, 2016. `doi:10.1016/j.matpur.2016.11.004`. p. 15

[LMR13] Joseph M. Landsberg, Laurent Manivel, and Nicolas Ressayre. Hypersurfaces with degenerate duals and the geometric complexity theory program. *Comment. Math. Helv.*, 88(2):469–484, 2013. `doi:10.4171/CMH/292`. p. 110

[LO15] Joseph M. Landsberg and Giorgio Ottaviani. New lower bounds for the border rank of matrix multiplication. *Theory Comput.*, 11:285–298, 2015. `arXiv:1112.6007`, `doi:10.4086/toc.2015.v011a011`. p. 110, 112

[Lov79] László Lovász. On the Shannon capacity of a graph. *IEEE Trans. Inform. Theory*, 25(1):1–7, 1979. `doi:10.1109/TIT.1979.1055985`. p. 13, 37, 43

[Mar08] Murray Marshall. *Positive polynomials and sums of squares*, volume 146 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2008. `doi:10.1090/surv/146`. p. 35

[MP71] Robert J. McEliece and Edward C. Posner. Hide and seek, data storage, and entropy. *The Annals of Mathematical Statistics*, 42(5):1706–1716, 1971. `doi:10.1214/aoms/1177693169`. p. 43

[MP08] Guillaume Malod and Natacha Portier. Characterizing Valiant's algebraic complexity classes. *J. Complexity*, 24(1):16–38, 2008. `doi:10.1016/j.jco.2006.09.006`. p. 121

[MS01] Ketan D. Mulmuley and Milind Sohoni. Geometric complexity theory. I. An approach to the P vs. NP and related problems. *SIAM J. Comput.*, 31(2):496–526, 2001. `doi:10.1137/S009753970038715X`. p. 14, 110

[MS08] Ketan D. Mulmuley and Milind Sohoni. Geometric complexity theory II. Towards explicit obstructions for embeddings among class varieties. *SIAM J. Comput.*, 38(3):1175–1206, 2008. `doi:10.1137/080718115`. p. 110

[Nes84] Linda Ness. A stratification of the null cone via the moment map. *Amer. J. Math.*, 106(6):1281–1329, 1984. With an appendix by David Mumford. `doi:10.2307/2374395`. p. 9, 95, 96

[Nis91] Noam Nisan. Lower bounds for non-commutative computation. In *Proceedings of the twenty-third annual ACM symposium on Theory of computing*, pages 410–418. ACM, 1991. `doi:10.1145/103418.103462`. p. 112

[Nor16] Sergey Norin. A distribution on triples with maximum entropy marginal. *arXiv*, 2016. `arXiv:1608.00243`. p. 85

[NW97] Noam Nisan and Avi Wigderson. Lower bounds on arithmetic circuits via partial derivatives. *Comput. Complexity*, 6(3):217–234, 1996/97. `doi:10.1007/BF01294256`. p. 110

[Pan78] Victor Ya. Pan. Strassen's algorithm is not optimal. Trilinear technique of aggregating, uniting and canceling for constructing fast algorithms for matrix operations. In *19th Annual Symposium on Foundations of Computer Science (Ann Arbor, Mich., 1978)*, pages 166–176. IEEE, Long Beach, Calif., 1978. p. 3

[Pan80] Victor Ya. Pan. New fast algorithms for matrix operations. *SIAM J. Comput.*, 9(2):321–342, 1980. `doi:10.1137/0209027`. p. 3

[Pan81] Victor Ya. Pan. New combinations of methods for the acceleration of matrix multiplication. *Comput. Math. Appl.*, 7(1):73–125, 1981. `doi:10.1016/0898-1221(81)90009-2`. p. 3

[Pan84] Victor Ya. Pan. *How to multiply matrices faster*, volume 179 of *Lecture Notes in Computer Science*. Springer-Verlag, Berlin, 1984. `doi:10.1007/3-540-13866-8`. p. 3

[Pan18] Victor Ya. Pan. Fast feasible and unfeasible matrix multiplication. *arXiv*, 2018. arXiv:1804.04102. p. 6

[PD01] Alexander Prestel and Charles N. Delzell. *Positive polynomials*. Springer Monographs in Mathematics. Springer-Verlag, Berlin, 2001. From Hilbert's 17th problem to real algebra. doi:10.1007/978-3-662-04648-7. p. 35

[Peb16] Luke Pebody. Proof of a conjecture of Kleinberg-Sawin-Speyer. *arXiv*, 2016. arXiv:1608.05740. p. 85

[PS98] George Pólya and Gabor Szegő. *Problems and theorems in analysis. I.* Classics in Mathematics. Springer-Verlag, Berlin, 1998. Series, integral calculus, theory of functions, Translated from the German by Dorothee Aeppli, Reprint of the 1978 English translation. doi:10.1007/978-3-642-61905-2. p. 21

[Raz09] Ran Raz. Multi-linear formulas for permanent and determinant are of super-polynomial size. *J. ACM*, 56(2):Art. 8, 17, 2009. doi:10.1145/1502793.1502797. p. 110

[Raz13] Ran Raz. Tensor-rank and lower bounds for arithmetic formulas. *J. ACM*, 60(6):Art. 40, 15, 2013. doi:10.1145/2535928. p. 15

[Rom82] Francesco Romani. Some properties of disjoint sums of tensors related to matrix multiplication. *SIAM J. Comput.*, 11(2):263–267, 1982. doi:10.1137/0211020. p. 3

[Sap16] Ramprasad Saptharishi. A survey of lower bounds in arithmetic circuit complexity 3.0.2, 2016. Online survey. URL: https://github.com/dasarpmar/lowerbounds-survey. p. 5, 17, 111, 114

[Sch81] Arnold Schönhage. Partial and total matrix multiplication. *SIAM J. Comput.*, 10(3):434–455, 1981. p. 3

[Sch03] Alexander Schrijver. *Combinatorial optimization: polyhedra and efficiency*, volume 24. Springer Science & Business Media, 2003. p. 39, 43

[Sha56] Claude E. Shannon. The zero error capacity of a noisy channel. *Institute of Radio Engineers, Transactions on Information Theory*, IT-2(September):8–19, 1956. doi:10.1109/TIT.1956.1056798. p. 13, 37

[Sha09] Asaf Shapira. Green's conjecture and testing linear-invariant properties. In *Proceedings of the Forty-first Annual ACM Symposium on Theory of Computing*, STOC '09, pages 159–166, New York, NY, USA, 2009. ACM. `doi:10.1145/1536414.1536438`. p. 50

[Shi16] Yaroslav Shitov. How hard is the tensor rank? *arXiv*, 2016. `arXiv:1611.01559`. p. 49

[Sin64] Richard C. Singleton. Maximum distance q-nary codes. *IEEE Trans. Information Theory*, IT-10:116–118, 1964. `doi:10.1109/TIT.1964.1053661`. p. 103

[SOK14] Adam Sawicki, Michał Oszmaniec, and Marek Kuś. Convexity of momentum map, Morse index, and quantum entanglement. *Rev. Math. Phys.*, 26(3):1450004, 39, 2014. `doi:10.1142/S0129055X14500044`. p. 9

[SSS09] Chandan Saha, Ramprasad Saptharishi, and Nitin Saxena. The power of depth 2 circuits over algebras. In *IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science*, volume 4, pages 371–382, 2009. `arXiv:0904.2058`, `doi:10.4230/LIPIcs.FSTTCS.2009.2333`. p. 111

[Sto10] Andrew James Stothers. *On the complexity of matrix multiplication*. PhD thesis, University of Edinburgh, 2010. `http://hdl.handle.net/1842/4734`. p. 4, 6, 8, 50

[Str69] Volker Strassen. Gaussian elimination is not optimal. *Numer. Math.*, 13(4):354–356, 1969. `doi:10.1007/BF02165411`. p. 3, 5

[Str83] Volker Strassen. Rank and optimal computation of generic tensors. *Linear Algebra Appl.*, 52/53:645–685, 1983. `doi:10.1016/0024-3795(83)80041-X`. p. 112

[Str86] Volker Strassen. The asymptotic spectrum of tensors and the exponent of matrix multiplication. In *Proceedings of the 27th Annual Symposium on Foundations of Computer Science*, SFCS '86, pages 49–54, Washington, DC, USA, 1986. IEEE Computer Society. `doi:10.1109/SFCS.1986.52`. p. 4, 7

[Str87] Volker Strassen. Relative bilinear complexity and matrix multiplication. *J. Reine Angew. Math.*, 375/376:406–443, 1987. `doi:10.1515/crll.1987.375-376.406`. p. 3, 4, 51, 69

[Str88] Volker Strassen. The asymptotic spectrum of tensors. *J. Reine Angew. Math.*, 384:102–152, 1988. `doi:10.1515/crll.1988.384.102`. p. 4, 7, 12, 19, 27, 28, 30, 31, 33, 34, 51, 52, 53

[Str91] Volker Strassen. Degeneration and complexity of bilinear maps: some asymptotic spectra. *J. Reine Angew. Math.*, 413:127–180, 1991. `doi:10.1515/crll.1991.413.127`. p. 3, 4, 10, 11, 50, 51, 54, 57, 58, 59, 68, 69, 83, 84

[Str94] Volker Strassen. Algebra and complexity. In *First European Congress of Mathematics, Vol. II (Paris, 1992)*, volume 120 of *Progr. Math.*, pages 429–446. Birkhäuser, Basel, 1994. `doi:10.1007/s10107-008-0221-1`. p. 69

[Str05] Volker Strassen. Komplexität und Geometrie bilinearer Abbildungen. *Jahresber. Deutsch. Math.-Verein.*, 107(1):3–31, 2005. p. 4, 90, 96, 97, 102, 103

[Tao08] Terence Tao. *Structure and randomness: pages from year one of a mathematical blog.* American Mathematical Soc., 2008. p. 50

[Tao16] Terence Tao. A symmetric formulation of the Croot–Lev–Pach–Ellenberg–Gijswijt capset bound. `https://terrytao.wordpress.com`, 2016. p. 50, 60, 83, 86

[Tob91] Verena Tobler. *Spezialisierung und Degeneration von Tensoren.* PhD thesis, Universität Konstanz, 1991. `http://nbn-resolving.de/urn:nbn:de:bsz:352-opus-20324`. p. 59

[TS16] Terence Tao and Will Sawin. Notes on the “slice rank” of tensors. `https://terrytao.wordpress.com`, 2016. p. 50, 60

[Val79] Leslie G. Valiant. Completeness classes in algebra. In *Conference Record of the Eleventh Annual ACM Symposium on Theory of Computing (Atlanta, Ga., 1979)*, pages 249–261. ACM, New York, 1979. `doi:10.1145/800135.804419`. p. 109, 110, 125

[Val80] Leslie G. Valiant. *Reducibility by algebraic projections.* University of Edinburgh, Department of Computer Science, 1980. Internal Report. p. 111, 121, 125

[VC15] Péter Vrana and Matthias Christandl. Asymptotic entanglement transformation between W and GHZ states. *J. Math. Phys.*, 56(2):022204, 12, 2015. `arXiv:1310.3244`, `doi:10.1063/1.4908106`. p. 71

[VDDMV02] F. Verstraete, J. Dehaene, B. De Moor, and H. Verschelde. Four qubits can be entangled in nine different ways. *Phys. Rev. A (3)*, 65(5, part A):052112, 5, 2002. `doi:10.1103/PhysRevA.65.052112`. p. 50

[Wal14] Michael Walter. *Multipartite quantum states and their marginals*. PhD thesis, ETH Zurich, 2014. `arXiv:1410.6820`. p. 95

[WDGC13] Michael Walter, Brent Doran, David Gross, and Matthias Christandl. Entanglement polytopes: multiparticle entanglement from single-particle information. *Science*, 340(6137):1205–1208, 2013. `arXiv:1208.0365`, `doi:10.1126/science.1232957`. p. 9, 97

[Wil12] Virginia Vassilevska Williams. Multiplying matrices faster than Coppersmith-Winograd. Extended abstract. In *STOC'12—Proceedings of the 2012 ACM Symposium on Theory of Computing*, pages 887–898. ACM, New York, 2012. `doi:10.1145/2213977.2214056`. p. 4, 6, 8, 50

[Zui17] Jeroen Zuiddam. A note on the gap between rank and border rank. *Linear Algebra Appl.*, 525:33–44, 2017. `doi:10.1016/j.laa.2017.03.015`. p. 14, 112

[Zui18] Jeroen Zuiddam. The asymptotic spectrum of graphs and the Shannon capacity. *arXiv*, 2018. `arXiv:1807.00169`. p. 37

Glossary

$\langle n \rangle$ $n \times \cdots \times n$ diagonal tensor. 47

$\langle a, b, c \rangle$ matrix multiplication tensor. 48

$G * H$ or-product. 42

$G \boxtimes H$ strong graph product, and-product. 35

$\alpha(G)$ stability number. 35

$\overline{\chi}(G)$ clique cover number. 40

K_k complete graph on k vertices. 36

$F^{\theta}(t)$ quantum functional. 96

$G(t)$ $\mathrm{GL}_{n_1} \times \cdots \times \mathrm{GL}_{n_k}$ for $t \in \mathbb{F}^{n_1} \otimes \cdots \otimes \mathbb{F}^{n_k}$. 52

$H(P)$ Shannon entropy of probability distribution P. 52

$h(p)$ binary entropy of probability $p \in [0, 1]$. 53

$\tau(\Phi)$ hitting set number. 59

$\underset{\sim}{\tau}(\Phi)$ asymptotic hitting set number. 60

ω matrix multiplication exponent. 47

$\mathbf{P}$ moment polytope. 94

$\mathcal{P}(X)$ the set of probability distributions on X. 52

R rank. 27

$\underset{\sim}{\mathrm{R}}$ asymptotic rank. 27

$\underline{\mathrm{R}}(t)$ border rank. 50

$\mathrm{R}(G)$ rank of a graph, clique cover number. 40

$\mathrm{R}(t)$ tensor rank. 47

$\mathrm{SR}(t)$ slice rank. 58

Q subrank. 27

$\underset{\sim}{\mathrm{Q}}$ asymptotic subrank. 27

$\underline{\mathrm{Q}}(t)$ border subrank. 50

$\mathrm{Q}(\Phi)$ combinatorial subrank. 10

$\mathrm{Q}(G)$ subrank of a graph, stability number. 40

$\mathrm{supp}(t)$ support. 52

$\Theta(G)$ Shannon capacity. 35

$\vartheta(G)$ Lovász theta number. 41

$G \sqcup H$ disjoint union. 36

$W(t)$ $S_{n_1} \times \cdots \times S_{n_k}$ for $t \in \mathbb{F}^{n_1} \otimes \cdots \otimes \mathbb{F}^{n_k}$. 53

$\mathbf{X}(S, \leqslant)$ asymptotic spectrum of semiring S with Strassen preorder $\leqslant$. 25

$\zeta^{(S)}(t)$ gauge point. 51

$\zeta^{\theta}(t)$ support functional. 52

Summary

Algebraic complexity, asymptotic spectra and entanglement polytopes

Matrix rank is well-known to be multiplicative under the Kronecker product, additive under the direct sum, normalised on identity matrices and non-increasing under multiplying from the left and from the right by any matrices. In fact, matrix rank is the only real matrix parameter with these four properties. In 1986 Strassen proposed to study the extension to tensors: find all maps from k-tensors to the reals that are multiplicative under the tensor Kronecker product, additive under the direct sum, normalised on "identity tensors", and non-increasing under acting with linear maps on the k tensor factors. Strassen called the collection of these maps the "asymptotic spectrum of k-tensors". He proved that understanding the asymptotic spectrum implies understanding the asymptotic relations among tensors, including the asymptotic subrank and the asymptotic rank. In particular, knowing the asymptotic spectrum means knowing the arithmetic complexity of matrix multiplication, a central problem in algebraic complexity theory.

One of the main results in this dissertation is the first explicit construction of an infinite family of elements in the asymptotic spectrum of complex k-tensors, called the quantum functionals. Our construction is based on information theory and moment polytopes i.e. entanglement polytopes. Moreover, among other things, we study the relation of the recently introduced slice rank to the quantum functionals and find that "asymptotic" slice rank equals the minimum over the quantum functionals. Besides studying the above tensor parameters, we extend the Coppersmith–Winograd method (for obtaining asymptotic combinatorial subrank lower bounds) to higher-order tensors, i.e. order at least 4. We apply this generalisation to obtain new upper bounds on the asymptotic tensor rank of complete graph tensors via the laser method. (Joint work with Christandl and Vrana; QIP 2018, STOC 2018.)

In graph theory, as a new instantiation of the abstract theory of asymptotic spectra we introduce the asymptotic spectrum of graphs. Analogous to the situation for tensors, understanding the asymptotic spectrum of graphs means

understanding the Shannon capacity, a graph parameter capturing the zero-error communication complexity of communication channels. In different words: we prove a new duality theorem for Shannon capacity. Some known elements in the asymptotic spectrum of graphs are the Lovász theta number and the fractional Haemers bounds.

Finally, we study an algebraic model of computation called algebraic branching programs. An algebraic branching program (abp) is the trace of a product of matrices with affine linear forms as matrix entries. The maximum size of the matrices is called the width of the abp. In 1992 Ben-Or and Cleve proved that width-3 abps can compute any polynomial efficiently in the formula size. On the other hand, in 2011 Allender and Wang proved that some polynomials cannot be computed by any width-2 abp. We prove that any polynomial can be efficiently approximated by a width-2 abp, where approximation is defined in the sense of degeneration. (Joint work with Ikenmeyer and Bringmann; CCC 2017, JACM 2018.)

Titles in the ILLC Dissertation Series:

ILLC DS-2009-01: **Jakub Szymanik**
Quantifiers in TIME and SPACE. Computational Complexity of Generalized Quantifiers in Natural Language

ILLC DS-2009-02: **Hartmut Fitz**
Neural Syntax

ILLC DS-2009-03: **Brian Thomas Semmes**
A Game for the Borel Functions

ILLC DS-2009-04: **Sara L. Uckelman**
Modalities in Medieval Logic

ILLC DS-2009-05: **Andreas Witzel**
Knowledge and Games: Theory and Implementation

ILLC DS-2009-06: **Chantal Bax**
Subjectivity after Wittgenstein. Wittgenstein's embodied and embedded subject and the debate about the death of man.

ILLC DS-2009-07: **Kata Balogh**
Theme with Variations. A Context-based Analysis of Focus

ILLC DS-2009-08: **Tomohiro Hoshi**
Epistemic Dynamics and Protocol Information

ILLC DS-2009-09: **Olivia Ladinig**
Temporal expectations and their violations

ILLC DS-2009-10: **Tikitu de Jager**
"Now that you mention it, I wonder...": Awareness, Attention, Assumption

ILLC DS-2009-11: **Michael Franke**
Signal to Act: Game Theory in Pragmatics

ILLC DS-2009-12: **Joel Uckelman**
More Than the Sum of Its Parts: Compact Preference Representation Over Combinatorial Domains

ILLC DS-2009-13: **Stefan Bold**
Cardinals as Ultrapowers. A Canonical Measure Analysis under the Axiom of Determinacy.

ILLC DS-2010-01: **Reut Tsarfaty**
Relational-Realizational Parsing

ILLC DS-2010-02: **Jonathan Zvesper**
Playing with Information

ILLC DS-2010-03: **Cédric Dégremont**
The Temporal Mind. Observations on the logic of belief change in interactive systems

ILLC DS-2010-04: **Daisuke Ikegami**
Games in Set Theory and Logic

ILLC DS-2010-05: **Jarmo Kontinen**
Coherence and Complexity in Fragments of Dependence Logic

ILLC DS-2010-06: **Yanjing Wang**
Epistemic Modelling and Protocol Dynamics

ILLC DS-2010-07: **Marc Staudacher**
Use theories of meaning between conventions and social norms

ILLC DS-2010-08: **Amélie Gheerbrant**
Fixed-Point Logics on Trees

ILLC DS-2010-09: **Gaëlle Fontaine**
Modal Fixpoint Logic: Some Model Theoretic Questions

ILLC DS-2010-10: **Jacob Vosmaer**
Logic, Algebra and Topology. Investigations into canonical extensions, duality theory and point-free topology.

ILLC DS-2010-11: **Nina Gierasimczuk**
Knowing One's Limits. Logical Analysis of Inductive Inference

ILLC DS-2010-12: **Martin Mose Bentzen**
Stit, Iit, and Deontic Logic for Action Types

ILLC DS-2011-01: **Wouter M. Koolen**
Combining Strategies Efficiently: High-Quality Decisions from Conflicting Advice

ILLC DS-2011-02: **Fernando Raymundo Velazquez-Quesada**
Small steps in dynamics of information

ILLC DS-2011-03: **Marijn Koolen**
The Meaning of Structure: the Value of Link Evidence for Information Retrieval

ILLC DS-2011-04: **Junte Zhang**
System Evaluation of Archival Description and Access

ILLC DS-2011-05: **Lauri Keskinen**
Characterizing All Models in Infinite Cardinalities

ILLC DS-2011-06: **Rianne Kaptein**
Effective Focused Retrieval by Exploiting Query Context and Document Structure

ILLC DS-2011-07: **Jop Briët**
Grothendieck Inequalities, Nonlocal Games and Optimization

ILLC DS-2011-08: **Stefan Minica**
Dynamic Logic of Questions

ILLC DS-2011-09: **Raul Andres Leal**
Modalities Through the Looking Glass: A study on coalgebraic modal logic and their applications

ILLC DS-2011-10: **Lena Kurzen**
Complexity in Interaction

ILLC DS-2011-11: **Gideon Borensztajn**
The neural basis of structure in language

ILLC DS-2012-01: **Federico Sangati**
Decomposing and Regenerating Syntactic Trees

ILLC DS-2012-02: **Markos Mylonakis**
Learning the Latent Structure of Translation

ILLC DS-2012-03: **Edgar José Andrade Lotero**
Models of Language: Towards a practice-based account of information in natural language

ILLC DS-2012-04: **Yurii Khomskii**
Regularity Properties and Definability in the Real Number Continuum: idealized forcing, polarized partitions, Hausdorff gaps and mad families in the projective hierarchy.

ILLC DS-2012-05: **David García Soriano**
Query-Efficient Computation in Property Testing and Learning Theory

ILLC DS-2012-06: **Dimitris Gakis**
Contextual Metaphilosophy - The Case of Wittgenstein

ILLC DS-2012-07: **Pietro Galliani**
The Dynamics of Imperfect Information

ILLC DS-2012-08: **Umberto Grandi**
Binary Aggregation with Integrity Constraints

ILLC DS-2012-09: **Wesley Halcrow Holliday**
Knowing What Follows: Epistemic Closure and Epistemic Logic

ILLC DS-2012-10: **Jeremy Meyers**
Locations, Bodies, and Sets: A model theoretic investigation into nominalistic mereologies

ILLC DS-2012-11: **Floor Sietsma**
Logics of Communication and Knowledge

ILLC DS-2012-12: **Joris Dormans**
Engineering emergence: applied theory for game design

ILLC DS-2013-01: **Simon Pauw**
Size Matters: Grounding Quantifiers in Spatial Perception

ILLC DS-2013-02: **Virginie Fiutek**
Playing with Knowledge and Belief

ILLC DS-2013-03: **Giannicola Scarpa**
Quantum entanglement in non-local games, graph parameters and zero-error information theory

ILLC DS-2014-01: **Machiel Keestra**
Sculpting the Space of Actions. Explaining Human Action by Integrating Intentions and Mechanisms

ILLC DS-2014-02: **Thomas Icard**
The Algorithmic Mind: A Study of Inference in Action

ILLC DS-2014-03: **Harald A. Bastiaanse**
Very, Many, Small, Penguins

ILLC DS-2014-04: **Ben Rodenhäuser**
A Matter of Trust: Dynamic Attitudes in Epistemic Logic

ILLC DS-2015-01: **María Inés Crespo**
Affecting Meaning. Subjectivity and evaluativity in gradable adjectives.

ILLC DS-2015-02: **Mathias Winther Madsen**
The Kid, the Clerk, and the Gambler - Critical Studies in Statistics and Cognitive Science

ILLC DS-2015-03: **Shengyang Zhong**
Orthogonality and Quantum Geometry: Towards a Relational Reconstruction of Quantum Theory

ILLC DS-2015-04: **Sumit Sourabh**
Correspondence and Canonicity in Non-Classical Logic

ILLC DS-2015-05: **Facundo Carreiro**
Fragments of Fixpoint Logics: Automata and Expressiveness

ILLC DS-2016-01: **Ivano A. Ciardelli**
Questions in Logic

ILLC DS-2016-02: **Zoé Christoff**
Dynamic Logics of Networks: Information Flow and the Spread of Opinion

ILLC DS-2016-03: **Fleur Leonie Bouwer**
What do we need to hear a beat? The influence of attention, musical abilities, and accents on the perception of metrical rhythm

ILLC DS-2016-04: **Johannes Marti**
Interpreting Linguistic Behavior with Possible World Models

ILLC DS-2016-05: **Phong Lê**
Learning Vector Representations for Sentences - The Recursive Deep Learning Approach

ILLC DS-2016-06: **Gideon Maillette de Buy Wenniger**
Aligning the Foundations of Hierarchical Statistical Machine Translation

ILLC DS-2016-07: **Andreas van Cranenburgh**
Rich Statistical Parsing and Literary Language

ILLC DS-2016-08: **Florian Speelman**
Position-based Quantum Cryptography and Catalytic Computation

ILLC DS-2016-09: **Teresa Piovesan**
Quantum entanglement: insights via graph parameters and conic optimization

ILLC DS-2016-10: **Paula Henk**
Nonstandard Provability for Peano Arithmetic. A Modal Perspective

ILLC DS-2017-01: **Paolo Galeazzi**
Play Without Regret

ILLC DS-2017-02: **Riccardo Pinosio**
The Logic of Kant's Temporal Continuum

ILLC DS-2017-03: **Matthijs Westera**
Exhaustivity and intonation: a unified theory

ILLC DS-2017-04: **Giovanni Cinà**
Categories for the working modal logician

ILLC DS-2017-05: **Shane Noah Steinert-Threlkeld**
Communication and Computation: New Questions About Compositionality

ILLC DS-2017-06: **Peter Hawke**
The Problem of Epistemic Relevance

ILLC DS-2017-07: **Aybüke Özgün**
Evidence in Epistemic Logic: A Topological Perspective

ILLC DS-2017-08: **Raquel Garrido Alhama**
Computational Modelling of Artificial Language Learning: Retention, Recognition & Recurrence

ILLC DS-2017-09: **Miloš Stanojević**
Permutation Forests for Modeling Word Order in Machine Translation

ILLC DS-2018-01: **Berit Janssen**
Retained or Lost in Transmission? Analyzing and Predicting Stability in Dutch Folk Songs

ILLC DS-2018-02: **Hugo Huurdeman**
Supporting the Complex Dynamics of the Information Seeking Process

ILLC DS-2018-03: **Corina Koolen**
Reading beyond the female: The relationship between perception of author gender and literary quality

ILLC DS-2018-04: **Jelle Bruineberg**
Anticipating Affordances: Intentionality in self-organizing brain-body-environment systems

ILLC DS-2018-05: **Joachim Daiber**
Typologically Robust Statistical Machine Translation: Understanding and Exploiting Differences and Similarities Between Languages in Machine Translation

ILLC DS-2018-06: **Thomas Brochhagen**
Signaling under Uncertainty

ILLC DS-2018-07: **Julian Schlöder**
Assertion and Rejection

ILLC DS-2018-08: **Srinivasan Arunachalam**
Quantum Algorithms and Learning Theory

ILLC DS-2018-09: **Hugo de Holanda Cunha Nobrega**
Games for functions: Baire classes, Weihrauch degrees, transfinite computations, and ranks

ILLC DS-2018-10: **Chenwei Shi**
Reason to Believe

ILLC DS-2018-11: **Malvin Gattinger**
New Directions in Model Checking Dynamic Epistemic Logic

ILLC DS-2018-12: **Julia Ilin**
Filtration Revisited: Lattices of Stable Non-Classical Logics