Asymptotic properties of hypergraphs and channels

Duy Hoang Ta

Numéro National de Thèse : 2022LYSEN024

THÈSE DE DOCTORAT DE L'UNIVERSITÉ DE LYON
opérée par
l'École Normale Supérieure de Lyon

École Doctorale N°512
École Doctorale en Informatique et Mathématiques de Lyon

Discipline : Informatique

Soutenue publiquement le 11/07/2022, par :
Duy-Hoang Ta

Asymptotic properties of hypergraphs and channels

Propriétés asymptotiques des hypergraphes et des canaux

Devant le jury composé de :

Michael Walter, Professeur, Ruhr-Universität Bochum	Rapporteur
Péter Vrana, Professeur, Budapest University of Technology and Economics	Rapporteur
Sophie Laplante, Professeure des universités, Université Paris Diderot	Examinatrice
Stéphan Thomassé, Professeur des universités, ENS de Lyon	Examinateur
Omar Fawzi, Directeur de recherche INRIA, ENS de Lyon	Directeur de thèse

Contents

Abstract

Understanding the properties of objects under a natural product operation is a central theme in mathematics, computer science and physics. Examples of such basic objects include noisy communication channels in information theory, computational problems in algebraic complexity, and graphs in discrete mathematics. In this PhD thesis, we study the asymptotic growth of relevant properties for the powers of such objects.

The first objects we consider are hypergraphs equipped with the strong product operation and the property of interest is the independence number. The asymptotic growth of the independence number of a hypergraph is known as the Shannon capacity. We introduce a general method for lower bounding the Shannon capacity of hypergraphs via combinatorial degenerations, a notion which originates from the study of matrix multiplication in algebraic complexity theory. This allows us to obtain the best-known lower bounds for multiple combinatorial problems, including the corner problem and its application in communication complexity.

Tensors are the second considered objects. We can equip them with the tensor product and the property of interest is the symmetric subrank. The symmetric subrank is a notion we introduce motivated by limitations of current tensor methods to bound the Shannon capacity of hypergraphs. We prove precise relations and separations between subrank and symmetric subrank. We also prove that for symmetric tensors, the subrank and the symmetric subrank are asymptotically equal. This proves the asymptotic subrank analogon of a conjecture known as Comon's conjecture in the theory of tensors.

Finally, we study the growth of the divergence between tensor powers of quantum channels. By exploiting symmetries, we propose efficient algorithms to approximate the asymptotic channel divergence between channels. As an application, we obtain improved bounds on quantum channel capacities.

Résumé

Comprendre les propriétés des objets dans le cadre d'une opération de produit naturel est un thème central en mathématiques, en informatique et en physique. Des exemples de tels objets de base incluent les canaux de communication bruyants en théorie de l'information, les problèmes de calcul en complexité algébrique et les graphes en mathématiques discrètes. Dans cette thèse, nous étudions la croissance asymptotique de propriétés pertinentes pour les puissances de tels objets.

Les premiers objets que nous considérons sont des hypergraphes munis de l'opération de produit fort et la propriété d'intérêt est le nombre d'indépendance. La croissance asymptotique du nombre d'indépendance d'un hypergraphe est connue sous le nom de capacité de Shannon. Nous introduisons une méthode générale pour minorer la capacité de Shannon des hypergraphes via des dégénérescences combinatoires, une notion issue de l'étude de la multiplication matricielle en théorie de la complexité algébrique. Cela nous permet d'obtenir les bornes inférieures les plus connues pour de multiples problèmes combinatoires, y compris le problème du coin et son application dans la complexité de la communication.

Les tenseurs sont les seconds objets considérés. Nous pouvons les équiper du produit tensoriel et la propriété d'intérêt est le sous-rang symétrique. Le sous-rang symétrique est une notion que nous introduisons motivée par les limitations des méthodes de tenseur actuelles pour borner la capacité de Shannon des hypergraphes. Nous prouvons des relations et séparations précises entre sous-rang et sous-rang symétrique. Nous prouvons également que pour les tenseurs symétriques, le sous-rang et le sous-rang symétrique sont asymptotiquement égaux. Cela prouve l'analogue de sous-rang asymptotique d'une conjecture connue sous le nom de conjecture de Comon dans la théorie des tenseurs.

Enfin, nous étudions la croissance de la divergence entre les puissances tensorielles des canaux quantiques. En exploitant les symétries, nous proposons des algorithmes efficaces pour approximer la divergence asymptotique des canaux entre canaux. Comme application, nous obtenons des bornes améliorées sur les capacités des canaux quantiques.

Contents of the thesis

This thesis is mainly based on three papers. The first one is joint work with Matthias Christandl, Omar Fawzi, and Jeroen Zuiddam [CFTZ22] and is presented in Chapter 3. The second paper is presented in Chapter 4 and is joint work with Matthias Christandl, Omar Fawzi, and Jeroen Zuiddam [CFTZ21]. The third paper is joint work with Omar Fawzi and Ala Shayeghi [FST21] and is presented in Chapter 5.

Notation

Common	
log	Binary logarithm.
ln	Natural logarithm.
$\mathbb{N}$	Natural numbers.
$\mathbb{Z}$	Integer numbers.
$\mathbb{R}$	Real numbers.
$\mathbb{C}$	Complex numbers.
$\mathbb{Z}_m$	Integers modulo m.
$A^\dagger$	Conjugate transpose of the matrix A.
$[k]$	Set $\{1, \ldots, k\}$.
$\Pr(E)$	Probability of the event E.
$\mathbb{E}[X]$	Expectation of a random variable X.
$\mathcal{P}([k])$	Set of all probability distributions on $[k]$.
Spaces	
$X, Y, \ldots$	Hilbert spaces associated with the systems $X, Y, \ldots$
d_X	Dimension of the space X.
XY	Tensor product $X \otimes Y$ or composite system XY.
$\mathscr{L}(X, Y)$	Space of linear operators from X to Y.
$\mathscr{L}(X)$	$\mathscr{L}(X, X)$.
Vectors	
$\lvert\psi\rangle^X, \lvert\phi\rangle^X, \ldots$	Vectors belonging to X.
$\langle\psi\rvert^X$	Dual vectors in $\mathcal{L}(X, \mathbb{C})$.
$\langle\psi\vert\phi\rangle$	Inner product of the vectors $\lvert\psi\rangle$ and $\lvert\phi\rangle$.
Operators	
$\mathrm{D}(X)$	Set of density operators on X.
ρ^X	Density operator on X.
id_X	Identity map on X or $\mathcal{L}(X)$.
$\lVert . \rVert_\infty$	Infinity norm .

Chapter 1

Introduction

In this Chapter, we give a high-level introduction to this thesis. After this introduction and the Preliminaries in Chapter 2, we get into the technical components of the thesis.

1.1 Asymptotics, regularization and amortization

The following question arises in several parts of mathematics, physics and computer science. Let $f : \mathcal{R} \to \mathbb{R}$ be a real-valued function on $\mathcal{R}$, where $\mathcal{R}$ is a universe. These objects can have a variety of meanings depending on the context. In graph theory, $\mathcal{R}$ may be the set of (un)directed graphs, and for each graph $r \in \mathcal{R}$, the number $f(r)$ can possibly be a graph parameter, such as the independence number, the clique number or the matching number of r. In complexity theory, $\mathcal{R}$ may be the set of computational tasks, and $f(r)$ may be minimum cost in time or space which is needed to perform r. In information theory, $\mathcal{R}$ may be a set of noisy communication channels, and $f(r)$ may be some entropic measures of correlations. To fix notation, throughout the thesis we will use "task" to refer to elements of $\mathcal{R}$ and "cost" to refer to the function f. In a general setting, computing the cost $f(r)$ of a task r may be very complex. For example, given an undirected graph G, computing the independence number of G was known as NP-Hard [Kar72].

Let's assume that $\mathcal{R}$ can be equipped with a binary multiplication operation $(\cdot)$ that is closed in $\mathcal{R}$. Given $n \geq 1$ an integer and for a task $r \in \mathcal{R}$, let r^n denote the n-fold product of r with itself. Intuitively, r^n can be understood as "n copies of r". More precisely, r^n is a task of performing the task r in n times. The asymptotic behavior of $f(r^n)$ for any fixed r and as n grows is a main object of study in this thesis. Namely, we consider the limit[1] $\tilde{f}(r) = \lim_{n\to\infty}(f(r^n))^{1/n}$ where r is fixed. This notion captures some ideas such as the amortized (in complexity theory) or the regularized (in information theory) cost of performing r.

In this thesis, we study the asymptotic cost $\tilde{f}(r)$ for a certain family of functions f as well as their applications. The most appealing aspect of this study is that solving one such general problem results in many computational and economic problems being solved in batches. For example, computing the same function on multiple inputs, simultaneously communicating many messages through parallel channels, buying or selling many identical items, etc. In many areas of mathematics and physics, some mathematical operations such

[1] Assuming a limit exists, which will be considered in all cases throughout this thesis

as "direct sum" and "tensor product" are often applied to different types of objects, and it is natural to investigate how sum and product affect the various parameters of these objects. Here are some concrete examples where we elucidate the above mentioned concept.

Matrix multiplication A well-known problem in computer science concerning tensors is the number of arithmetic operations required to perform matrix multiplication between two $n \times n$ matrices. The exponent of matrix multiplication is defined as the smallest real number ω such that multiplying $n \times n$ matrices can be performed in $\mathcal{O}(n^\omega)$ arithmetic operations. It is known to be between 2 and $2.37\ldots$ [LG14, AW21]. Let $t \in \mathbb{F}^d \otimes \mathbb{F}^d \otimes \mathbb{F}^d$ be a 3-order tensor over field $\mathbb{F}$, the tensor rank $\mathrm{R}(t)$ of t is defined as the smallest number k such that t can be written as $\sum_{i=1}^k u_i \otimes v_i \otimes w_i$ where $u_i, v_i, w_i \in \mathbb{F}^d$ for all $i \in [k]$. Consider $\mathcal{R}$ the set of 3-order tensors equipped with the tensor product operation $\otimes$ and the cost function, in this context, the tensor rank R, i.e., $f = \mathrm{R}$, the complexity of matrix multiplication can be determined by the tensor rank of the matrix multiplication tensors $\langle m, m, m\rangle := \sum_{i,j,k=1}^m e_{ij} \otimes e_{jk} \otimes e_{ki} \in \mathbb{F}^{m^2} \otimes \mathbb{F}^{m^2} \otimes \mathbb{F}^{m^2}$, where $\{e_{ij}\}_{i,j=1}^m$ are standard basis of $\mathbb{F}^{m^2} \cong \mathbb{F}^m \otimes \mathbb{F}^m$. In fact, the exponent ω is exactly determined by the asymptotic rank of $\langle 2, 2, 2\rangle$ [Str88], that is, $\omega = \log\left(\tilde{\mathrm{R}}(\langle 2, 2, 2\rangle)\right)$, where $\tilde{\mathrm{R}}(\langle 2, 2, 2\rangle := \lim_{n\to\infty}(\mathrm{R}(\langle 2, 2, 2\rangle^{\otimes n}))^{1/n}$.

Zero-error communication In this context, we consider the universe $\mathcal{R}$ consisting of all undirected graphs. Let $G = (V(G), E(G)), H = (V(H), E(H))$ be two undirected graphs in $\mathcal{R}$. The *strong product* $G \boxtimes H$ is the graph with vertex set given by $V(G) \times V(H)$ and edge set given by all pairs $\{(u, v), (u', v')\}$ such that $(\{u, u'\} \in E(G)$ and $v = v')$ or $(u = u'$ and $\{v, v'\} \in E(H))$ or $(\{u, u'\} \in E(G)$ and $\{v, v'\} \in E(H))$. We equip $\mathcal{R}$ with the strong product operation $\boxtimes$ and the cost function, in this context, is the independence number. For any graph $G \in \mathcal{R}$, let $\alpha(G)$ be the independence number of G. Then, the respective asymptotic quantity is $\Theta(G) := \lim_{n\to\infty} \alpha(G^{\boxtimes n})^{1/n}$, also known as the *Shannon capacity* of G which was introduced by Shannon [Sha56]. Since α has supermultiplicative property, we can write $\Theta(G) = \sup_n \alpha(G^{\boxtimes n})^{1/n}$. The Shannon capacity is an important and widely studied parameter in information theory. Rather indeed, its motivation comes from studying the zero-error capacity of a discrete memoryless noisy channel [Sha56]. In this model, a transmitter wants to send a message over the channel to a receiver, and the receiver must decode the message without any error. The zero-error capacity is the supremum over all achievable communication rates under this constraint, in the limit of multiple channel uses. This problem can be modeled by a *confusion graph* G associated to the channel. The vertices of G are labeled by the input symbols, and there is an edge between two vertices if and only if the corresponding inputs can result in the same output. By definition, for any $n \in \mathbb{N}$, $\alpha(G^{\boxtimes n})$ is the maximum number of messages that can be transmitted in n uses of the channel with no confusion. Thus, the zero-error capacity of a channel is often referred to as Shannon capacity of the confusion graph of the channel.

Finally, in many cases, the asymptotic cost $\tilde{f}$ may be easier to compute or study analytically, than the actual function f, and thus, sometimes $\tilde{f}(r)$ can be used to bound $f(r)$. The following problem illustrates such a situation.

Nondeterministic communication complexity In the two-party communication complexity model, we have two players (usually referred to as Alice and Bob) and a boolean function $g : X \times Y \to \{0, 1\}$. Alice is given $x \in X$ and Bob is given $y \in Y$. Both know the function g and their goal is to collaboratively compute $g(x, y)$. It is well known that the *nondeterministic communication complexity* of g, denoted $N^1(g)$, is characterized by the number of monochromatic rectangles needed to cover the matrix associated with the function, which is defined as follows. Let M_g be a $|X| \times |Y|$ boolean matrix with the rows labelled by elements of X and columns labelled by the elements of Y, and $M_g(x, y) = 1$ iff $g(x, y) = 1$ for all $x \in X, y \in Y$. A rectangle of M_g is a submatrix of the form $A \times B$, where $A \subseteq X, B \subseteq Y$. A rectangle $A \times B$ is called 1-monochromatic if $M_g(x, y) = 1$ for all $(x, y) \in A \times B$. The nondeterministic cover number of the function g, denoted by $C^1(g)$, is the minimum number of 1-monochromatic rectangles (allowing overlaps) that can cover all the 1's in M_g. In fact, we have $N^1(g) = \log(C^1(g))$ for all boolean functions g. Now, let $\mathcal{R}$ be the set of boolean matrices. We equip $\mathcal{R}$ with tensor product operation and the cost function, in this context, is the cover number, i.e., $f = C^1$. Then, the asymptotic cover number of g, denoted $\tilde{C}^1(g)$, can be used to characterize the *amortized nondeterministic communication complexity* [KKN95] of g, denoted $\tilde{N}^1(g)$. Namely, we have $\tilde{N}^1(g) = \log(\tilde{C}^1(g))$. Moreover, $\tilde{C}^1(g)$ is exactly the Fractional cover number (see [KKN95] for the definition of Fractional cover number) of function g which can be computed efficiently by a linear program [KKN95]. Finally, the $\tilde{C}^1$ can also be used as a lower bound for C^1.

1.2 Asymptotic properties of hypergraphs: Shannon capacity

1.2.1 Context

The first objects we consider are hypergraphs equipped with the strong product operation and the cost of interest is the independence number, i.e., $\mathcal{R}$ is set of hypergraphs and f is independence number. The asymptotic growth of the independence number in powers of a hypergraph is known as the Shannon capacity. Many Ramsey type problems can be expressed as the Shannon capacity of some fixed hypergraph, such as the cap set problem that saw a recent breakthrough in [CLP17, EG17], and the USP problems that arise in the study of matrix multiplication [CKSU05, ASU13]. Another important instance of these problems is the corner problem studied in the context of multiparty communication complexity in the Number On the Forehead (NOF) model [Shk06a, Shk06b, LM07, CFL83, LPS18]. More precisely, let $\mathbb{F}_p$ be a finite field, the corner problem asks to determine the size of largest subset of $\mathbb{F}_p^n \times \mathbb{F}_p^n$ that does not contain a configuration of the form

$$(x, y), (x + \lambda, y), (x, y + \lambda),$$

where $x, y, \lambda \in \mathbb{F}_p^n$ and $\lambda \neq 0^n$.

To study this problem, we consider a more general one named generalized multidimensional Szemerédi problem[2] over $\mathbb{F}_p^n \times \mathbb{F}_p^n$. Let $S \subseteq \mathbb{F}_p \times \mathbb{F}_p$ be an ordered nonempty

[2]Which is a generalization of the multidimensional Szemerédi question [FK79] over $\mathbb{F}_p^n \times \mathbb{F}_p^n$ that will be presented in more detail in Section 3.2

set of size k. A subset $A \subseteq \mathbb{F}_p^n \times \mathbb{F}_p^n$ is called *S-free* if for any k ordered distinct points $[(x^1, y^1), \ldots, (x^k, y^k)]$ of A, there is a coordinate $i \in [n]$ such that the ordered tuple $[(x_i^1, y_i^1), \ldots, (x_i^k, y_i^k)]$ does not have a form

$$\{(a,b) + (\overline{\lambda}u, \overline{\lambda}v) : (u,v) \in S\}, \text{ for some } (a,b) \in \mathbb{F}_p \times \mathbb{F}_p \text{ and } \overline{\lambda} \in \{1, \ldots, p-1\}.$$

Given an ordered nonempty set $S \subseteq \mathbb{F}_p \times \mathbb{F}_p$, the generalized multidimensional Szemerédi problem asks to determine the size of the largest S-free subset of $\mathbb{F}_p^n \times \mathbb{F}_p^n$. It is easy to verify that, the corner problem over $\mathbb{F}_p^n \times \mathbb{F}_p^n$ as a special case of this problem by choosing $S = \{(0,0), (1,0), (0,1)\}$. More notably, it can be rephrased as determining the Shannon capacity of a fixed hypergraph given a fixed $S \subseteq \mathbb{F}_p \times \mathbb{F}_p$. The generalized multidimensional Szemerédi problem over $\mathbb{F}_p^n \times \mathbb{F}_p^n$ will present in detail in Section 3.2.

1.2.2 Summary of the contributions

Chapter 3 We introduce and study a general algebraic method for lower bounding the Shannon capacity of directed hypergraphs via *combinatorial degenerations*. It is a combinatorial kind of "approximation" of subgraphs that originates from the study of matrix multiplication in algebraic complexity theory (and plays an important role there) but is used by us in a novel way. Using the combinatorial degeneration method, we make progress on some special cases of the generalized multidimensional Szemerédi problem. Especially, on the corner problem, our method gives an explicit construction of a corner-free subset in $\mathbb{F}_2^n \times \mathbb{F}_2^n$ of size $\Omega(3.39^n/\mathrm{poly}(n))$. Our result improves the previous lower bound $\Omega(2.82^n)$ of Linial, Pitassi and Shraibman [LPS18] and gets us closer to the best upper bound $4^{n-o(n)}$. Our new construction of corner-free sets also implies an improved NOF protocol for the *Eval problem*. In the Eval problem over a group G, three players need to determine whether their inputs $x_1, x_2, x_3 \in G$ sum to zero. We find that the NOF communication complexity of the Eval problem over $\mathbb{F}_2^n$ is at most $0.24n + \mathcal{O}(\log n)$, which improves the previous upper bound $0.5n + \mathcal{O}(\log n)$ from [ACFN15]. More specifically, we show several first lower bounds for some other special cases of generalized multidimensional Szemerédi problem over $\mathbb{F}_p^n$ for some small value p. Finally, we investigate the existing tensor methods for upper bounding the Shannon capacity (including slice rank, subrank, analytic rank, geometric rank, and G-stable rank). We find that these methods have strong limitations caused by the existence of large induced matchings. In particular, this implies a strong barrier for these methods to be used to prove nontrivial upper bounds for the corner problem.

1.3 Asymptotic symmetric subrank of tensors

1.3.1 Context

Tensors are the second objects we considered. We can equip them with the tensor product and the cost of interest is the symmetric subrank. The symmetric subrank is a notion of rank that we introduce that is motivated by limitations of current tensor methods to bound the Shannon capacity of hypergraphs. The symmetric subrank upper bounds the independence number of hypergraphs but not the induced matching number and we propose this method as a route to circumvent the induced matching barrier.

Our symmetric subrank is inspired by the notion of subrank, which was introduced by Strassen [Str87] in the study of matrix multiplication and measures the size of the largest "identity tensor" that can be obtained from a k-tensor by applying linear operations to the k tensor legs. In the symmetric subrank we require these linear maps to be all equal. This definition makes sense for symmetric as well as non-symmetric tensors, as long as we think explicitly of our tensors as having equal side-lengths, so that applying one linear map to all tensors legs makes sense. The *symmetric subrank* of $f = (f_{i_1,\ldots,i_k})_{i_1,\ldots,i_k\in[d]} \in (\mathbb{F}^d)^{\otimes k}$ is defined as

$$\mathrm{Q_s}(f) = \max\{r \in \mathbb{N} : \langle r\rangle \leq_s f\} \ ,$$

where $\langle r\rangle \leq_s f$ means that there exists a matrix $(A_{i,j})_{i\in[r],j\in[d]} \in \mathbb{F}^{r\times d}$ such that for any $i_1,\ldots,i_k \in [r]$,

$$\sum_{j_1,\ldots,j_k\in[d]} A_{i_1,j_1}\cdots A_{i_k,j_k} f_{j_1,\ldots,j_k} = \begin{cases} 1 \text{ if } i_1 = i_2 = \cdots = i_k \\ 0 \text{ otherwise.} \end{cases} \tag{1.1}$$

As we mentioned, the symmetric subrank is the symmetric variation on Strassen's subrank [Str87]. Namely, in the definition of the subrank of tensor f (denoted by $\mathrm{Q}(f)$), instead of using k times the same matrix A in equation (1.1), we may choose k possibly different matrices $A^{(1)},\ldots,A^{(k)}$. The relation between the symmetric subrank and the subrank is analogous to the relation between the symmetric rank and the rank [CGLM08]. Note though that unlike the symmetric rank which only makes sense for symmetric tensors, the symmetric subrank can be defined for any tensor. Another simple observation about the symmetric subrank is that it can never be larger than the other relevant notions of rank: $\mathrm{Q_s}(f) \leq \mathrm{Q}(f) \leq d$ for any tensor f in dimension d.

An important component of our analysis is the study of the asymptotic behaviour of the subrank and symmetric subrank. This is captured by the asymptotic subrank

$$\tilde{\mathrm{Q}}(f) = \lim_{n\to\infty} \mathrm{Q}(f^{\otimes n})^{1/n}$$

and the asymptotic symmetric subrank[3]

$$\tilde{\mathrm{Q}}_{\mathrm{s}}(f) = \lim_{n\to\infty} \mathrm{Q_s}(f^{\otimes n})^{1/n}.$$

These notions are relevant for instance for the study of Shannon capacity problems of hypergraphs. More precisely, the important property of $\tilde{\mathrm{Q}}_{\mathrm{s}}$ is that it directly gives an upper bound on the Shannon capacity of hypergraphs. Strassen [Str86, Str88, Str91, CVZ18] proved a duality theorem for $\tilde{\mathrm{Q}}$ of the form $\tilde{\mathrm{Q}}(f) = \min_{\phi\in X} \phi(f)$ where the dual region X is the "asymptotic spectrum of tensors", a set of very special well-behaved tensor parameters. We are naturally lead to a theoretical study of symmetry in the asymptotic theory of tensors of Strassen, the introduction of the asymptotic spectrum of symmetric tensors X_{s} and the analogous duality theorem for $\tilde{\mathrm{Q}}_{\mathrm{s}}(f)$ (and more general related parameters).

[3]For this defition to make sense (i.e. the limit to exist) we need to put mild conditions on the tensor, see Propoposition 4.3.1. However, we can give a more general definition by replacing the lim by a limsup which is always valid.

1.3.2 Summary of the contributions

Chapter 4 We prove precise relations and separations between subrank and symmetric subrank. We prove that for symmetric tensors the subrank and the symmetric subrank are asymptotically equal over appropriate fields[4], i.e., $\tilde{\mathrm{Q}}(f) = \tilde{\mathrm{Q}}_{\mathrm{s}}(f)$ for any symmetric tensor f. This proves the asymptotic subrank analogon of a conjecture known as Comon's conjecture in the theory of tensors. This result allows us to prove a strong connection between the general and symmetric version of an asymptotic duality theorem of Strassen. We introduce a representation-theoretic method to asymptotically bound the symmetric subrank called the symmetric quantum functional in analogy with the quantum functionals [CVZ18], and we study the relations between these functionals.

1.4 Asymptotic divergence of quantum channels

1.4.1 Context

Lastly, the objects we considered are quantum channels whose operation and cost are the tensor product of channels and the divergence, respectively. Namely, we use the recently introduced $\mathrm{D}^{\#}$ Rényi channel divergences [FF21b] as the cost of interest. Our objective is to study the asymptotic divergence between the channels. In this context, the asymptotic quantity is also called *regularized* channel divergence.

Quantum channel discrimination is a fundamental information processing task in quantum information theory and has been studied in various aspects [CPR00, Aci01, Sac05, CMW16, FFRS20, WW19, WBHK20, FF21b, FGW22]. The problem asks to distinguish between two quantum channels $\mathcal{N}$ and $\mathcal{M}$ having black box access to n uses for each of them. The strategy involves the choice of n input quantum states and the observation of n output quantum states. Choosing the inputs can be done either beforehand or adaptively according to the previous outputs. The word "adaptive" refers to the fact that the input to the used channel, at a fixed step, can depend on the previous outputs received. In contrast, a strategy is called parallel (or nonadaptive) if the n black boxes are used in parallel on a fixed input state. Following the usual notation in hypothesis testing, we denote by α_n the type I error probability, which is the probability that the channel is actually $\mathcal{N}$ but our procedure says $\mathcal{M}$; and β_n is the type II error probability which is the probability that our procedure outputs $\mathcal{N}$ when the actual channel is $\mathcal{M}$. In general, the goal is to determine the trade-off between these two errors.

Although several regimes can be considered, the most studied is the asymmetric hypothesis testing setting (Stein's setting). This setting aims to consider the asymptotic behavior of the optimal type II error probability $(-\frac{1}{n}\log\beta_n)$, under the condition that the type I error probability α_n does not exceed a constant $\epsilon \in (0,1)$. The works [FFRS20, WW19, WBHK20] showed that if we take $\epsilon \to 0$ this is given by the regularized Umegaki channel divergence $\mathrm{D}^{\mathrm{reg}}(\mathcal{N}\|\mathcal{M}) := \lim_{n\to\infty}\frac{1}{n}\mathrm{D}(\mathcal{N}^{\otimes n}\|\mathcal{M}^{\otimes n})$, where $\mathrm{D}(\mathcal{N}\|\mathcal{M})$ is the Umegaki channel divergence [5] between $\mathcal{N}$ and $\mathcal{M}$.

[4] Algebraically closed of characteristic at least $k+1$, where k is the order of the tensor

[5] For a divergence $\mathbf{D}$ defined on quantum states, the corresponding channel divergence is defined by maximizing the divergence between the channel outputs over the set of possible inputs [LKDW18]. The Umegaki divergence is defined as $\mathrm{D}(\rho\|\sigma) = \mathrm{tr}(\rho(\log\rho - \log\sigma))$ for positive semidefinite operators ρ and σ.

In the strong converse regime, we consider the behavior of α_n under the condition $\beta_n \leq 2^{-rn}$ with $r > \mathrm{D}^{\mathrm{reg}}(\mathcal{N}\|\mathcal{M})$. From [FF21b], we know that this is determined by the *regularized sandwiched Rényi channel divergence* $\widetilde{\mathrm{D}}_\alpha^{\mathrm{reg}}(\mathcal{N}\|\mathcal{M})$ (defined as $\lim_{n\to\infty} \frac{1}{n}\widetilde{\mathrm{D}}_\alpha(\mathcal{N}^{\otimes n}\|\mathcal{M}^{\otimes n})$, where $\widetilde{\mathrm{D}}_\alpha(\mathcal{N}\|\mathcal{M})$ is the sandwiched Rényi channel divergence[6] between channels $\mathcal{N}$ and $\mathcal{M}$). In addition, in a recent work [FGW22], the authors conjectured that any strategy that makes the type II error decay with an exponent larger than the regularized Umegaki channel divergence will unavoidably result in the type I error converging to 1 exponentially fast in the asymptotic limit. This conjecture will imply the continuity of the regularized Sandwiched Rényi channel divergence at $\alpha = 1$. More precisely, the conjecture states that if the probability of type II error as $n \to \infty$ i.e. $(\liminf_{n\to\infty} -\frac{1}{n}\log\beta_n)$ is greater than $\mathrm{D}^{\mathrm{reg}}(\mathcal{N}\|\mathcal{M})$, then there exists $c > 0$ such that the probability of type I error $(1-\alpha_n)$ is lesser than 2^{-cn} for sufficiently large n. If this conjecture holds, then $\lim_{\alpha\to 1^+}\widetilde{\mathrm{D}}_\alpha^{\mathrm{reg}}(\mathcal{N}\|\mathcal{M}) = \inf_{\alpha>1}\widetilde{\mathrm{D}}_\alpha^{\mathrm{reg}}(\mathcal{N}\|\mathcal{M}) = \mathrm{D}^{\mathrm{reg}}(\mathcal{N}\|\mathcal{M})$.

As we have seen so far, the regularized sandwiched Rényi divergence between channels plays an importance role in this context. Given two quantum channels $\mathcal{N}$ and $\mathcal{M}$, we want to compute the quantity $\widetilde{\mathrm{D}}_\alpha^{\mathrm{reg}}(\mathcal{N}\|\mathcal{M})$. Since the sandwiched Rényi divergence between channels is non-additive in general [FFRS20], it is unclear whether its regularization can be computed efficiently. In [FF21b], the authors provided a converging hierarchy of upper bounds on the regularized divergence between channels. It allows us to show that $\widetilde{\mathrm{D}}_\alpha^{\mathrm{reg}}(\mathcal{N}\|\mathcal{M})$ can be approximated by $\frac{1}{n}\mathrm{D}_\alpha^{\#}(\mathcal{N}^{\otimes n}\|\mathcal{M}^{\otimes n})$ with arbitrary accuracy for sufficiently large n in finite time. Moreover, $\mathrm{D}_\alpha^{\#}(\mathcal{N}^{\otimes n}\|\mathcal{M}^{\otimes n})$ can be written in terms of a convex program as in [FF21b]. However, the size of this program grows exponentially with n.

1.4.2 Summary of the contributions

Chapter 5 We exploit the symmetries in the $\mathrm{D}^{\#}$ in order to obtain a hierarchy of semidefinite programming bounds on various regularized quantities. Specifically, for quantum channels $\mathcal{N}, \mathcal{M}$, we show that the optimization program defining $\mathrm{D}^{\#}(\mathcal{N}^{\otimes n}\|\mathcal{M}^{\otimes n})$ can be computed in $\mathrm{poly}(n)$ time, for fixed input and output dimensions. This result allows us to prove that for fixed input and output dimensions, the regularized sandwiched Rényi divergence between any two quantum channels can be approximated up to an ϵ accuracy in time that is polynomial in $1/\epsilon$. As applications, we give a general procedure to give efficient bounds on the regularized Umegaki channel divergence as well as the classical capacity and two-way assisted quantum capacity of quantum channels. In particular, we obtain slight improvements for the capacity of the amplitude damping channel.

[6] We recall that the sandwiched Rényi divergence [MDS+13, WWY14] of order $\alpha > 1$ is defined as $\widetilde{\mathrm{D}}_\alpha(\rho\|\sigma) := \frac{1}{\alpha-1}\log\mathrm{tr}\left[\left(\sigma^{\frac{1-\alpha}{2\alpha}}\rho\sigma^{\frac{1-\alpha}{2\alpha}}\right)^\alpha\right]$ for positive semidefinite operators ρ and σ.

Chapter 2

Preliminaries

The objective of this chapter is to introduce some notations and results that will be used throughout this thesis.

2.1 Basic notations

Asymptotics We use the standard symbols from asymptotic analysis: $\mathcal{O}, o, \Omega$. We write $f = \text{poly}(n)$ if there is a constant $c \geq 0$ such that $f(n) = \mathcal{O}(n^c)$.

Groups and Fields We will typically use multiplicative notation for the group operation of groups. Two particular groups which will arise frequently are, for $n \in \mathbb{N}$, the cyclic group C_n, and the symmetric group $\mathfrak{S}_n$ of permutations on n elements.

For $p \in \mathbb{N}$, if p is a power of a prime, we write $\mathbb{F}_p$ for the finite field of order p.

Polynomial on a vector space. For a finite dimensional complex vector space $\mathcal{H}$, the *dual vector space* $\mathcal{H}^*$ of $\mathcal{H}$ is the vector space of all linear transformations $\varphi : \mathcal{H} \to \mathbb{C}$. The *coordinate ring* of $\mathcal{H}$, denoted $\mathcal{O}(\mathcal{H})$, is the algebra consisting of all $\mathbb{C}$-linear combinations of products of elements from $\mathcal{H}^*$. An element of $\mathcal{O}(\mathcal{H})$ is called a polynomial on $\mathcal{H}$. A polynomial $p \in \mathcal{O}(\mathcal{H})$ is called *homogeneous* if it is a $\mathbb{C}$-linear combination of a product of k non-constant elements of $\mathcal{H}^*$ (for a fixed non-negative integer k). We denote by $\mathcal{O}_k(\mathcal{H})$ the set all homogeneous polynomials of degree k.

2.2 Tensor ranks

Tensors

Let $f \in \mathbb{F}^{d_1} \otimes \cdots \otimes \mathbb{F}^{d_k}$ be a k-tensor over a field $\mathbb{F}$. Let $\{e_1, \ldots, e_{d_j}\}$ denote the standard basis of $\mathbb{F}^{d_j}$. We may then write f as

$$f = \sum f_{i_1,\ldots,i_k}\, e_{i_1} \otimes \ldots e_{i_k},$$

where the sum goes over $i \in [d_1] \times \cdots \times [d_k]$. In this way f corresponds to a k-way array $f \in \mathbb{F}^{d_1 \times \cdots \times d_k}$. For $f \in \mathbb{F}^{d_1} \otimes \cdots \otimes \mathbb{F}^{d_k}$ and $f' \in \mathbb{F}^{d'_1} \otimes \cdots \otimes \mathbb{F}^{d'_k}$, we define the tensor

product as $(f \otimes f')_{(i_1,j_1),\ldots,(i_k,j_k)} = f_{i_1,\ldots,i_k} \cdot f'_{j_1,\ldots,j'_k}$. We define the *support* of f as the set

$$\operatorname{supp}(f) \coloneqq \{(i_1,\ldots,i_k) : f_{i_1,\ldots,i_k} \neq 0\} \subseteq [d_1] \times \cdots \times [d_k].$$

For $r \in \mathbb{N}$, we call $\langle r \rangle \coloneqq \sum_{i=1}^r e_i^{\otimes k}$ the *unit tensor* of size r.

For $f \in \mathbb{F}^{d_1} \otimes \cdots \otimes \mathbb{F}^{d_k}$ and $i \in [k]$, we denote by $\operatorname{flatten}_i(f)$ the image of f under the grouping $\mathbb{F}^{d_1} \otimes \cdots \otimes \mathbb{F}^{d_k} \to \mathbb{F}^{d_i} \otimes \left(\bigotimes_{j \neq i} \mathbb{F}^{d_j}\right)$, which we call a flattening. We can think of $\operatorname{flatten}_i(f)$ as a d_i by $\prod_{j \neq i} d_j$ matrix.

A k-tensor $f \in (\mathbb{F}^d)^{\otimes k}$ is said to be *symmetric* if $f_{i_1,\ldots,i_k} = f_{i_{\sigma(1)},\ldots,i_{\sigma(k)}}$ for any $i_1,\ldots,i_k \in [d]$ and any permutation $\sigma \in \mathfrak{S}_k$. For example, a tensor $f \in (\mathbb{F}^d)^{\otimes 3}$ is symmetric if $f_{ijk} = f_{ikj} = f_{jik} = f_{jki} = f_{kij} = f_{kji}$, for all $i,j,k \in [d]$.

Overview: notions of tensor rank

In this section, we give an introduction to some of the existing notions of the rank of tensors and their usefulness.

Let $f \in \mathbb{F}^{d_1} \otimes \cdots \otimes \mathbb{F}^{d_k}$ be a tensor. The tensor rank $\mathrm{R}(f)$ of f is defined as the smallest number r such that f can be written as $\sum_{i=1}^r u_i^1 \otimes \cdots \otimes u_i^k$ with $u_i^t \in \mathbb{F}^{d_t}$ for all $t \in [k]$.

We say that the tensor $f \in \mathbb{F}^{d_1} \otimes \cdots \otimes \mathbb{F}^{d_k}$ *restricts to* $f' \in \mathbb{F}^{d'_1} \otimes \cdots \otimes \mathbb{F}^{d'_k}$, and write $f' \leq f$ if there exist linear maps $A^{(i)} : \mathbb{F}^{d_i} \to \mathbb{F}^{d'_i}$ such that $f' = (A^{(1)} \otimes \cdots \otimes A^{(k)}) \cdot f$. Written in the standard basis, this corresponds to having for all $i_1 \in [d'_1],\ldots,i_k \in [d'_k]$ that

$$f'_{i_1,\ldots,i_k} = \sum_{j_1 \in [d_1],\ldots,j_k \in [d_k]} A^{(1)}_{i_1,j_1} \ldots A^{(k)}_{i_k,j_k} f_{j_1,\ldots,j_k}.$$

Example 2.2.1. Here we see restriction in action in a small example. For the tensors

$$f = e_0 \otimes e_0 \otimes e_0 + e_1 \otimes e_1 \otimes e_1,$$
$$f' = e_0 \otimes (e_0 \otimes e_0 + e_1 \otimes e_1),$$

we have $f' \leq f$ by letting $A^{(1)} : e_0 \mapsto e_0, e_1 \mapsto e_0$ and letting $A^{(2)}$ and $A^{(3)}$ be the identity map.

Strassen [Str87] defined the *subrank* of f as

$$\mathrm{Q}(f) \coloneqq \max\{r \in \mathbb{N} : \langle r \rangle \leq f\}.$$

The subrank has been considered further in [CVZ18] for studying the capset problem and has also been used to study the barrier for Coppersmith-Winograd method for matrix multiplication problem in [CVZ19b, Alm19].

Similarly, one may define the "opposite" of the subrank as $\mathrm{R}(f) \coloneqq \min\{r \in \mathbb{N} : f \leq \langle r \rangle\}$, which redefines the notion of tensor rank. For $k = 2$, the subrank and rank of f are the usual matrix rank: $\mathrm{Q}(f) = \mathrm{R}(f) = \operatorname{rank}(f)$. When $k \geq 3$, however, there are f for which $\mathrm{Q}(f) < \mathrm{R}(f)$. In fact, the tensor rank can be larger than the dimensions $d_1,\ldots,d_k$, whereas the subrank cannot exceed $\min_i d_i$.

There are many applications which need us to understand the rate of growth of the rank and subrank of tensor product powers of a fixed tensor. Strassen [Str87] defined the *asymptotic subrank* of $f \in \mathbb{F}^{d_1} \otimes \cdots \otimes \mathbb{F}^{d_k}$ as

$$\tilde{\mathrm{Q}}(f) := \lim_{n\to\infty} \mathrm{Q}(f^{\otimes n})^{1/n}\,.$$

Since the subrank is super-multiplicative, we can, by Fekete's lemma (see Lemma A.3.1), replace the limit by a supremum. Similar to asymptotic subrank we can define the *asymptotic rank* of $f \in \mathbb{F}^{d_1} \otimes \cdots \otimes \mathbb{F}^{d_k}$ as

$$\tilde{\mathrm{R}}(f) = \lim_{n\to\infty} \mathrm{R}(f^{\otimes n})^{1/n}\,.$$

A well-known problem in computer science concerning tensors is about the number of arithmetic operations required for multiplying two $n \times n$ matrices. The answer is known to be between n^2 and $\mathcal{O}(n^{2.37\cdots})$, or in other words, the exponent of matrix multiplication ω is known to be between 2 and $2.37\ldots$ [LG14, AW21]. The complexity of matrix multiplication turns out to be determined by the tensor rank of matrix multiplication tensors $\langle m, m, m\rangle$ corresponding to taking the trace of the product of three $m \times m$ matrices. More formally, we can write $\langle m, m, m\rangle$ as $\sum_{i,j,k=1}^{m} e_{ij} \otimes e_{jk} \otimes e_{ki} \in (\mathbb{F}^m)^{\otimes 2} \otimes (\mathbb{F}^m)^{\otimes 2} \otimes (\mathbb{F}^m)^{\otimes 2}$, where $\{e_{ij}\}_{i,j=1}^{m}$ are standard basis of $(\mathbb{F}^m)^{\otimes 2}$. In fact, the exponent ω is exactly determined by the asymptotic rank of $\langle 2, 2, 2\rangle$, that is, $\omega = \tilde{\mathrm{R}}(\langle 2, 2, 2\rangle)$.

The second tool we focus on is the slice rank. The notion of slice rank was introduced by Tao [Tao16] and was further developed in [TS16] and [BCC+17] as a variation on tensor rank to study cap sets and approaches to fast matrix multiplication algorithms. A tensor in $\mathbb{F}^{d_1} \otimes \cdots \otimes \mathbb{F}^{d_k}$ has *slice rank one* if it has the form $u \otimes v$ for $u \in \mathbb{F}^{d_i}$ and $v \in \bigotimes_{j\neq i} \mathbb{F}^{d_j}$ for some $i \in [k]$. The *slice rank* of f, denoted by $\mathrm{SR}(f)$, is the smallest number r such that f can be written as sum of r slice rank one tensors. Since slice rank is not sub-multiplicative and not super-multiplicative, the limit $\lim_{n\to\infty} \mathrm{SR}(f^{\otimes n})^{1/n}$ might not always exist [CVZ18]. We define

$$\tilde{\mathrm{SR}}(f) = \limsup_{n\to\infty} \mathrm{SR}(f^{\otimes n})^{1/n}\,.$$

Since slice rank is monotonic under the restriction order and is normalized on $\langle r\rangle$ [Tao16], i.e., $\mathrm{SR}(\langle r\rangle) = r$, it follows that $\mathrm{Q}(f) \leq \mathrm{SR}(f)$ and $\tilde{\mathrm{Q}}(f) \leq \tilde{\mathrm{SR}}(f)$.

2.3 Shannon capacity

We first recall some notations and basic concepts in graphs and hypergraphs. Then, in Section 2.3.1 we recall the notion of Shannon capacity of graphs.

Graphs

Definition 2.3.1. An undirected graph is a pair (V, E), where V is a set of objects called vertices and E is a set of two element subsets of V called edges.

It is often common to add a notion of direction to the edges of a graph. This gives us the concept of a directed graph (*or digraph*).

Definition 2.3.2. A directed graph, also called a digraph, is a pair (V, E), where V is a set of objects called vertices and $E \subseteq V \times V$ is a collection of ordered pairs of elements of V, called directed edges (or sometimes just edges).

Let $G = (V, E)$ be an (un)directed graph. For each edge $e = (u, v)$ of G, we call u, v are endpoints of the edge e. An *independent set* or *stable set* in G is a subset of V such that no edge has all its endpoints covered by vertices in the corresponding subset. The *independence number* or *stability number* $\alpha(G)$ is the cardinality of the largest independent set in G.

A *directed walk* in a directed graph $G = (V, E)$ is a sequence of vertices $v_0, v_1, \ldots, v_k$ and edges $(v_0, v_1), (v_1, v_2), \ldots, (v_{k-1}, v_k)$. A *directed path* (or *path*) in a directed graph G is a walk where the vertices in the walk are all different. A *directed closed walk* (or *closed walk*) in a directed graph is a walk where $v_0 = v_k$. A *directed cycle* (or *cycle*) in a directed graph is a closed walk where all the vertices v_i are different for $0 \leq i \leq k$.

Definition 2.3.3. A directed graph is called a *directed acyclic graph* (or *acyclic graph*) if it does not contain any directed cycles.

Hypergraphs

Definition 2.3.4. A *directed k-uniform hypergraph* H is a pair $H = (V, E)$ where V is a finite set of elements called vertices, and E is a set of k-tuples of elements of V which are called hyperedges or edges. If the set of edges E is invariant under permuting the k coefficients of its elements, then we may also think of H as an *undirected k-uniform hypergraph*.

Let $H = (V, E)$ be a directed k-uniform hypergraph with n vertices. The *adjacency tensor* A of H is defined as

$$A_{i_1,\ldots,i_k} = \begin{cases} 1 \text{ if } i_1 = i_2 = \cdots = i_k \text{ or } (i_1, \ldots, i_k) \in E, \\ 0 \text{ otherwise.} \end{cases}$$

Definition 2.3.5. An *independent set* in a directed k-uniform hypergraph $H = (V, E)$ is a subset S of the vertices V that induces no edges, meaning for every $(e_1, \ldots, e_k) \in E$ there is an $i \in [k]$ such that $e_i \notin S$. The *independence number* of H, denoted by $\alpha(H)$, is the maximal size of an independent set in H.

2.3.1 Shannon capacity of graphs

Let G, H be two (un)directed graphs. The *strong product* of G and H, denoted $G \boxtimes H$, is defined as

1. The vertex set $V(G \boxtimes H) = V(G) \times V(H)$.

2. Any two distinct vertices (u, u') and (v, v') form an edge in $G \boxtimes H$ if $u = v$ and $(v, v') \in E(H)$ or $(u, u') \in E(G)$ and $v = v'$ or $(u, u') \in E(G)$ and $(v, v') \in E(H)$.

One can observe that if S and T are independent sets in two directed graphs G and H, respectively, then $S \times T$ is an independent set in the strong product $G \boxtimes H$. Therefore, we have $\alpha(G)\alpha(H) \leq \alpha(G \boxtimes H)$.

For any (un)directed graph G, let $G^{\boxtimes n}$ denote the n-fold product of G with itself. The *Shannon capacity* $\Theta(G)$ is defined as the limit

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

The limit exists and equals the supremum $\sup_n \alpha(G^{\boxtimes n})^{1/n}$ by Fekete's lemma (Lemma A.3.1).

The Shannon capacity was introduced by Shannon [Sha56] and is an important and widely studied parameter in information theory (see e.g., [Alo98, Boh03, Hae79, Lov79, Zui19, PS19]). It is the effective size of an alphabet in an information channel (for more formal definitions, refer to [Sha56]) represented by the graph $G = (V, E)$. The input is a set of letters $V = \{1, \ldots, d\}$ and two letters are confusable when transmitted over the channel if and only if there is an edge between them in G. Then $\alpha(G)$ is the maximum size of a set of pairwise non-confusable single letters. Moreover, for any $n \in \mathbb{N}$, $\alpha(G^{\boxtimes n})$ is the maximum size of a set of pairwise non-confusable n-letter words. So the effective size of the alphabet in the information channel is given by $\lim_{n\to\infty}(\alpha(G^{\boxtimes n}))^{1/n}$ (Note that this is the Shannon capacity of the graph G).

Computing the Shannon capacity is nontrivial already for small graphs. In [Lov79], Lovász computed the value $\Theta(C_5) = \sqrt{5}$, where C_k denotes the undirected cycle on k vertices. In fact, even the value of $\Theta(C_7)$ is currently not known. The algorithmic problem of computing the Shannon capacity Θ is not even known to be decidable. On the other hand, deciding whether $\alpha(G) \leq k$, given a graph G and some $k \in \mathbb{N}$, is NP-complete [Kar72].

2.4 Asymptotic spectra theory

In this section, we recall the theory of asymptotic spectra of Strassen which introduced in series of papers [Str86, Str87, Str88, Str91] and have been further developed in [Zui18, Vra21, WZ21]. A lot of the following definitions are standard and we have used the notations from [Zui18].

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

1. $(a + b) + c = a + (b + c)$, $a + b = b + a$
2. $0 + a = a$, $0 \cdot a = 0$, $1 \cdot a = a$
3. $(a \cdot b) \cdot c = a \cdot (b \cdot c)$
4. $a \cdot (b + c) = a \cdot b + a \cdot c$

A semiring $(S, +, \cdot, 0, 1)$ is *commutative* if for all $a, b \in S$ holds $a \cdot b = b \cdot a$. For any natural number $n \in \mathbb{N}$, let $n \in S$ denote the sum of n times the element $1 \in S$.

A *preorder* $\leq$ on S is a relation such that for any $a, b, c \in S$ holds that $a \leq a$, and that if $a \leq b$ and $b \leq c$, then $a \leq c$. A preorder $\leq$ on S is a *Strassen preorder* if for all $a, b, c \in S$, $n, m \in \mathbb{N}$ holds

1. $n \leq m$ in $\mathbb{N}$ if and only if $n \leq m$ in S
2. if $a \leq b$, then $a + c \leq b + c$ and $a \cdot c \leq b \cdot c$
3. if $b \neq 0$, then there exists an $r \in \mathbb{N}$ such that $a \leq r \cdot b$.

Let $S = (S, +, \cdot, 0, 1)$ and $S' = (S', +, \cdot, 0, 1)$ be semirings. A *semiring homomorphism* from S to S' is a map $\phi : S \to S'$ such that $\phi(a+b) = \phi(a) + \phi(b)$, $\phi(a \cdot b) = \phi(a) \cdot \phi(b)$ for all $a, b \in S$, and $\phi(1) = 1$. Let $\mathbb{R}_{\geq 0} = (\mathbb{R}_{\geq 0}, +, \cdot, 0, 1)$ be the semiring of non-negative real numbers with the usual addition and multiplication operations. The *asymptotic spectrum* $X(S, \leq)$ of the semiring $S = (S, +, \cdot, 0, 1)$ with respect to the preorder $\leq$ is the set of $\leq$-monotone semiring homomorphisms from S to $\mathbb{R}_{\geq 0}$, that is,

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

Let $a \in S$. The *rank* of a is defined as $\mathrm{R}(a) \coloneqq \min\{r \in \mathbb{N} : a \leq r\}$ and the *subrank* of a is defined as $\mathrm{Q}(a) \coloneqq \max\{n \in \mathbb{N} : n \leq a\}$. Similarly, the *asymptotic rank* and *asymptotic subrank* of a are defined as

$$\tilde{\mathrm{R}}(a) \coloneqq \lim_{N \to \infty} \sqrt[N]{\mathrm{R}(a^N)}$$
$$\tilde{\mathrm{Q}}(a) \coloneqq \lim_{N \to \infty} \sqrt[N]{\mathrm{Q}(a^N)}.$$

Since Q is supermultiplicative, and R is submultiplicative, Fekete's lemma (Lemma A.3.1) implies that these limits indeed exist and can be replaced by

$$\tilde{\mathrm{R}}(a) = \inf_N \sqrt[N]{\mathrm{R}(a^N)}$$
$$\tilde{\mathrm{Q}}(a) = \sup_N \sqrt[N]{\mathrm{Q}(a^N)}.$$

In [Str88], the author proved the following dual characterization of $\tilde{\mathrm{R}}(a)$ and $\tilde{\mathrm{Q}}(a)$ in terms of the asymptotic spectrum.

Theorem 2.4.1 ([Str88], Theorem 3.8)**.** *Let S be a commutative semiring and let $\leq$ be a Strassen preorder on S. For any $a \in S$ such that $1 \leq a$ and $2 \leq a^k$ for some $k \in \mathbb{N}$, holds*

$$\tilde{\mathrm{R}}(a) = \max_{\phi \in X(S, \leq)} \phi(a)$$
$$\tilde{\mathrm{Q}}(a) = \min_{\phi \in X(S, \leq)} \phi(a).$$

The asymptotic preorder $\lesssim$ associated to $\leq$ is defined by $a \lesssim b$ if there is a sequence if natural numbers $(x_n)_{n \in \mathbb{N}}$ such that $\inf_n (x_n)^{1/n} = 1$ and such that for all $n \in \mathbb{N}$ we have $a^n \leq x_n \cdot b^n$. The asymptotic spectrum of a commutative semiring with respect to a Strassen preoder $\leq$ also characterizes the asymptotic preorder $\lesssim$ associated to $\leq$. The dual characterization is that $a \lesssim b$ if and only if for all $\phi \in X(S, \leq)$ it holds that $\phi(a) \leq \phi(b)$. See [Str88, Corolary 2.6] and see also [Zui18, Theorem 2.12].

Asymptotic spectrum of graphs

We present a semiring with a Strassen preorder on graphs which was introduced by Zuiddam [Zui19] to gain a better understanding of the Shannon capacity of graphs.

Let G and H be (un)directed graphs. The *disjoint union* $G \sqcup H$ is defined by

$$V(G \sqcup H) = V(G) \sqcup V(H)$$
$$E(G \sqcup H) = E(G) \sqcup E(H)\,.$$

For $n \in \mathbb{N}$, the complete graph K_n is the graph with $V(K_n) = \{1, \ldots, n\}$ and $E(K_n) = \{(i,j) : i,j \in \{1,\ldots,n\}, i \neq j\}$. Thus K_0 is an empty graph and K_1 is the graph consisting of a single vertex and no edges. We denote $\overline{K}_n$ is a graph on n vertices and no edges. Thus $K_0 = \overline{K}_0$ and $K_1 = \overline{K}_1$.

We define the *cohomomorphism preoder* $\leq$ on graphs by $G \leq H$ if and only if there is an injective map $f : V(G) \to V(H)$ such that $(u,v) \notin E(G)$ then $(f(u), f(v)) \notin E(H)$.

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 $(f(u), f(v)) \in E(H)$. A graph homomorphism $f : G \to H$ is a *graph isomorphism* if the map f is bijective between $V(G)$ and $V(H)$ and bijective as a map between $E(G)$ and $E(H)$. We write $G \sim H$ if there is a graph isomorphism $f : G \to H$. Let $\mathcal{G}$ be the set of isomorphism classes of graphs (which is basically the set of equivalence classes with respect to $\sim$). In [Zui19], the author proved that $\mathcal{G} = (\mathcal{G}, \sqcup, \boxtimes, K_0, K_1)$ is a commutative semiring and that the cohomomorphism preorder $\leq$ is a Strassen preorder on $\mathcal{G}$. By definition, the asymptotic spectrum of graphs $X(\mathcal{G}, \leq)$ consists of all maps $\phi : \mathcal{G} \to \mathbb{R}_{\geq 0}$ such that, for all $G, H \in \mathcal{G}$, the following conditions hold:

1. $\phi(G \sqcup H) = \phi(G) + \phi(H)$
2. $\phi(G \boxtimes H) = \phi(G)\phi(H)$
3. $\phi(K_1) = 1$
4. $G \leq H$ then $\phi(G) \leq \phi(H)$.

Note that the subrank of a graph G equals the independence number of G, that is

$$\alpha(G) = \max\{n \in \mathbb{N} : \overline{K}_n \leq G\}.$$

From Theorem 2.4.1, the Shannon capacity of G can be characterized by using asymptotic spectrum of graphs as

$$\Theta(G) = \min_{\phi \in X(\mathcal{G}, \leq)} \phi(G)\,.$$

2.5 Preliminaries on representation theory

In this section we give the definitions and notation from representation theory used throughout the thesis. For further details, we refer the reader to Refs. [Ser77] and [FH13]. Let $\mathcal{H}$ be a finite dimensional complex Hilbert space and G be a finite group. A *linear*

representation of G on $\mathcal{H}$ is a group homomorphism $\varrho : G \to \mathrm{GL}(\mathcal{H})$, where $\mathrm{GL}(\mathcal{H})$ is the general linear group on $\mathcal{H}$. The space $\mathcal{H}$ is called a *G-module*. For $v \in \mathcal{H}$ and $g \in G$, we write $g \cdot v$ as shorthand for $\varrho(g)v$ and denote by $\mathscr{L}(\mathcal{H})$ the set of linear operators on $\mathcal{H}$. For $X \in \mathscr{L}(\mathcal{H})$, the action of $g \in G$ on X is given by $\varrho(g)X\varrho(g)^*$.

A representation $\varrho : G \to \mathrm{GL}(\mathcal{H})$ of G is called *irreducible* if it contains no proper submodule $\mathcal{H}'$ of $\mathcal{H}$ such that $g\mathcal{H}' \subseteq \mathcal{H}'$. Let $\mathcal{H}$ and $\mathcal{H}'$ be G-modules, a linear map $\psi : \mathcal{H} \to \mathcal{H}'$ is called a *G-equivariant map* if $g \cdot \psi(v) = \psi(g \cdot v)$ for all $g \in G, v \in \mathcal{H}$. Two G-modules $\mathcal{H}$ and $\mathcal{H}'$ are called *G-isomorphic*, that is $\mathcal{H} \cong \mathcal{H}'$, if there is a bijective equivariant map from $\mathcal{H}$ to $\mathcal{H}'$. We denote by $\mathrm{End}^G(\mathcal{H})$, the set of all G-equivariant maps from $\mathcal{H}$ to $\mathcal{H}$, i.e.,

$$\mathrm{End}^G(\mathcal{H}) = \{T \in \mathscr{L}(\mathcal{H}) : T(g \cdot v) = g \cdot T(v), \forall v \in \mathcal{H}, g \in G\}.$$

We recall the well-known Schur's lemma, which characterizes all G-equivariant maps between irreducible G-modules.

Lemma 2.5.1 (Schur's lemma)**.** *Let V and W be irreducible G-modules, and let $f : V \to W$ be a G-equivariant map.*

1. *If V and W are nonisomorphic, then $f = 0$.*
2. *If there exists a G-isomorphism $\phi : V \to W$, then $f = \lambda\phi$ for some $\lambda \in \mathbb{C}$.*

Let G be a finite group acting on a finite dimensional complex vector space $\mathcal{H}$. Then the space $\mathcal{H}$ can be decomposed as $\mathcal{H} = \mathcal{H}_1 \oplus \cdots \oplus \mathcal{H}_t$ such that each $\mathcal{H}_i$ is a direct sum $\mathcal{H}_{i,1} \oplus \cdots \oplus \mathcal{H}_{i,m_i}$ of irreducible G-modules with the property that $\mathcal{H}_{i,j} \cong \mathcal{H}_{i',j'}$ if and only if $i = i'$. The G-modules $\mathcal{H}_1, \ldots, \mathcal{H}_t$ are called the *G-isotypical components* and $(m_1, \ldots, m_t)$ are called the *multiplicities* of the corresponding irreducible representations.

Let $d = \dim(\mathcal{H})$. For $n \in \mathbb{N}$, an integer partition is a sequence $\lambda = (\lambda_1, \ldots, \lambda_d)$ of nonnegative integers satisfying $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_d$. We say that λ is a partition of n denoted by $\lambda \vdash n$, if $\lambda_1 + \cdots + \lambda_d = n$. We denote the number of nonzero parts of λ by $\ell(\lambda)$. Similarly we denote by $\lambda \vdash_d n$, if $\lambda \vdash n$ and $\ell(\lambda) \leq d$. Consider the action of the symmetric group $\mathfrak{S}_n$ on the tensor power $\mathcal{H}^{\otimes n}$ by permuting the tensor legs (i.e. the coordinates of the rank-1 tensor),

$$\pi \cdot v_1 \otimes \cdots \otimes v_n = v_{\pi^{-1}(1)} \otimes \cdots \otimes v_{\pi^{-1}(n)}\,, \pi \in \mathfrak{S}_n\,.$$

Let the general linear group $\mathrm{GL}(\mathcal{H})$ act on $\mathcal{H}^{\otimes n}$ via the diagonal embedding $g \mapsto (g, \ldots, g)$,

$$g \cdot v_1 \otimes \cdots \otimes v_n = (gv_1) \otimes \cdots \otimes (gv_n),\ g \in \mathrm{GL}(\mathcal{H}).$$

The action of $\mathfrak{S}_n$ and $\mathrm{GL}(\mathcal{H})$ commute, so we have a well-defined action of the product group $\mathfrak{S}_n \times \mathrm{GL}(\mathcal{H})$ on $\mathcal{H}^{\otimes n}$. Schur-Weyl duality describes the decomposition of the space $\mathcal{H}^{\otimes n}$ into direct sum of irreducible $\mathfrak{S}_n \times \mathrm{GL}(\mathcal{H})$ representations. This decomposition is

$$\mathcal{H}^{\otimes n} \cong \bigoplus_{\lambda \vdash_d n} [\lambda] \otimes \mathbb{S}_\lambda(\mathcal{H})\,,$$

where $[\lambda]$ is an irreducible $\mathfrak{S}_n$-module of type λ and $\mathbb{S}_\lambda(\mathcal{H})$ is an irreducible of $\mathrm{GL}(\mathcal{H})$-module of type λ when $\ell(\lambda) \leq d$ and 0 when $\ell(\lambda) > d$.

2.6 Semidefinite programming

A matrix $A \in \mathbb{C}^{n\times n}$ is called *Hermitian* if $A^* = A$, where A^* is the conjugate transpose of A. The eigenvalues of a Hermitian matrix are real. Let $A \in \mathbb{C}^{n\times n}$. Then A is called *positive semidefinite*, denoted by $A \geq 0$, if A is Hermitian and all eigenvalues of A are non-negative. It is known that the following are equivalent:

1. $A \geq 0$,
2. $A^* = A$ and $v^*Av \geq 0$ for all $v \in \mathbb{C}^n$,
3. $A = L^*L$ for some $L \in \mathbb{C}^{n\times n}$.

The space of complex matrices is equipped with a complex inner product denoted by $\langle X, Y\rangle = \mathrm{tr}(Y^*X)$, which is linear in the first entry. One can also observe that the inner product of two Hermitian matrices is always real.

We now describe the notion of semidefinite programming. Let $m, n \in \mathbb{Z}_{\geq 0}, b_1, \ldots, b_m \in \mathbb{R}$ and let $C, B_1, \ldots, B_m \in \mathbb{C}^{n\times n}$. A (complex) semidefinite program is an optimization problem of the form

$$\begin{aligned} \text{Minimize: } & \langle C, X\rangle \\ \text{subject to: } & \langle B_i, X\rangle = b_i \text{ for all } i = 1, \ldots, m, \\ & X \geq 0. \end{aligned}$$

Here $X \in \mathbb{C}^{n\times n}$ is matrix variables. A Hermitian matrix $X \in \mathbb{C}^{n\times n}$ is called a *feasible solution* of the above program if it is positive semidefinite and fullfills all m linear constraints. It is called an *optimal solution* if it feasible and if for every feasible solution Y we have $\langle C, X\rangle \leq \langle C, Y\rangle$.

Semidefinite programs can be solved approximately up to any fixed precision in polynomial time by the ellipsoid method [GLS12, Chapter 2]. In practice though, interior point methods [NN94] are preferred, which also run in polynomial time. For further details about semidefinite programming, refer to [Tod01].

2.7 Quantum information

Basic notation

Let $\mathcal{H}$ be a finite dimensional complex Hilbert space; we denote by $\mathscr{L}(\mathcal{H})$ the set of linear operators on $\mathcal{H}$, $\mathscr{P}(\mathcal{H})$ denotes the set of positive semidefinite operators on $\mathcal{H}$, and $\mathrm{D}(\mathcal{H}) := \{\rho \in \mathscr{P}(\mathcal{H}) : \mathrm{tr}(\rho) = 1\}$ is the set of density operators on $\mathcal{H}$. For any two Hermitian operators $\rho, \sigma \in \mathscr{L}(\mathcal{H})$, we write $\rho \leq \sigma$ if $\sigma - \rho \in \mathscr{P}(\mathcal{H})$. Given $\rho \in \mathscr{L}(\mathcal{H})$, the support of ρ, denoted $\mathrm{supp}(\rho)$, is the orthogonal complement of its kernel. For $\rho, \sigma \in \mathscr{L}(\mathcal{H})$, we write $\rho \ll \sigma$, if $\mathrm{supp}(\rho) \subseteq \mathrm{supp}(\sigma)$.

Let X, Y be finite dimensional complex Hilbert spaces. For $A \in \mathscr{P}(X \otimes Y)$, we often explicitly indicate the quantum systems as a subscript by writing A_{XY}. The marginal on the subsystem X is denoted by $A_X = \mathrm{tr}_Y(A_{XY})$.

Let $\{|x\rangle\}_x$ and $\{|y\rangle\}_y$ be the standard bases for X and Y, respectively. We will use a correspondence between linear operator in $\mathscr{L}(Y, X)$ and vectors in $X \otimes Y$, given by the linear map $\mathrm{vec} : \mathscr{L}(Y, X) \to X \otimes Y$, defined as $\mathrm{vec}\,(|x\rangle\langle y|) = |x\rangle|y\rangle$.

Quantum channels

The evolution of a quantum system is described mathematically by a *quantum channel*, which is a linear, *completely positive*, and *trace-preserving* (CPTP) map acting on the quantum states of the underlying Hilbert space of the system. In detail, let $\mathcal{H}, \mathcal{H}'$ be two finite dimensional complex Hilbert spaces. Cosider a map $\mathcal{N} : \mathscr{L}(\mathcal{H}) \to \mathscr{L}(\mathcal{H}')$.

- $\mathcal{N}$ is called *positive* if it maps positive semidefinite operators to positive semidefinite operators, i.e., $\mathcal{N}(A) \geq 0$ for all $A \geq 0$. A map $\mathcal{N}$ is called *completely positive* if the map $\mathcal{I}_k \otimes \mathcal{N}$ is positive for all integers $k \geq 1$, where $\mathcal{I}_k$ denotes the identity map on $\mathscr{L}(\mathbb{C}^k)$.
- $\mathcal{N}$ is called *trace preserving* if $\mathrm{tr}(\mathcal{N}(A)) = \mathrm{tr}(A)$ for all linear operators A.

Throughout this thesis, we write $\mathcal{N}_{X \to Y}$ to denote a map $\mathcal{N} : \mathscr{L}(X) \to \mathscr{L}(Y)$ taking a quantum system X to a quantum system Y.

Choi representation

The Choi representation of a quantum channel gives a way to represent a quantum channel as a bipartite operator and is an essential concept in quantum information theory. Let X' be isomorphic to X and $|\Phi\rangle_{XX'} = \sum_x |x\rangle_X |x\rangle_{X'}$ be the unnormalized maximally entangled state. For a linear map $\mathcal{N}_{X' \to Y}$, we denote by $J^{\mathcal{N}}_{XY} \in \mathscr{P}(X \otimes Y)$ the corresponding *Choi matrix* defined as $J^{\mathcal{N}}_{XY} = (\mathcal{I}_X \otimes \mathcal{N})(|\Phi\rangle\langle\Phi|_{XX'})$, where $\mathcal{I}_X$ denotes the identity map on $\mathscr{L}(X)$.

Chapter 3

Shannon capacity of hypergraphs

This chapter is based on joint work with Matthias Christandl, Omar Fawzi, and Jeroen Zuiddam [CFTZ22].

3.1 Introduction

In this chapter, we discuss the Shannon capacity of hypergraphs, which is a natural generalization of the Shannon capacity of graphs introduced in Section 2.3.1.

We first generalize the strong product operation from graphs to directed k-uniform hypergraphs. The *strong product* of a pair of directed k-uniform hypergraphs $G = (V_G, E_G)$ and $H = (V_H, E_H)$ is denoted by $G \boxtimes H$ and defined as follows. It is a directed k-uniform hypergraph with vertex set $V_G \times V_H$ and the following edge set: Any k vertices $(g_1, h_1), \ldots, (g_k, h_k) \in V_G \times V_H$ form an edge $((g_1, h_1), \ldots, (g_k, h_k))$ if one of the following three conditions holds:

1. $g_1 = \cdots = g_k$ and $(h_1, \ldots, h_k) \in E_H$
2. $(g_1, \ldots, g_k) \in E_G$ and $h_1 = \cdots = h_k$
3. $(g_1, \ldots, g_k) \in E_G$ and $(h_1, \ldots, h_k) \in E_H$

It is easy to verify that, if S and T are independent sets in two directed k-uniform hypergraphs G and H, respectively, then $S \times T$ is an independent set in the strong product $G \boxtimes H$. Therefore, we have $\alpha(G)\alpha(H) \leq \alpha(G \boxtimes H)$. For any directed k-uniform hypergraph H, let $H^{\boxtimes n}$ denote the n-fold product of H with itself. The *Shannon capacity*[1] of a directed k-uniform hypergraph H is defined as the limit

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

By Fekete's lemma (see Lemma A.3.1) we can write $\Theta(H) = \sup_n (\alpha(H^{\boxtimes n}))^{1/n}$.

There is a large and important collection of Ramsey-type combinatorial problems such as the *capset problem* which are closely related to central problems in complexity theory. These problems can be formulated in terms of the asymptotic growth of the size of

[1] In the setting of directed graphs, also the term *Sperner capacity* (typically applied to the complement graph) [GKV92, GKV93] is used for what we call Shannon capacity.

the maximum independent sets in powers of a fixed small (directed or undirected) hypergraph. An important instance of these problems is the corner problem studied in the context of multiparty communication complexity in the Number On the Forehead (NOF) model [Shk06a, Shk06b, LM07, CFL83, LPS18, LS21]. In the following, we briefly summarize the connection between the NOF communication complexity and corner problem as well as how to rephrase the corner problem to Shannon capacity of hypergraph.

NOF communication complexity

The NOF model is very rich in terms of connections to Ramsey theory and additive combinatorics [BGG06, Shr18, LPS18, LS21], as well as applications to boolean models of compution such as branching programs and boolean circuits [CFL83, BT94]. The goal in the NOF model is for k players to compute a fixed given function $F : \mathcal{X}_1 \times \cdots \times \mathcal{X}_k \to \{0,1\}$ on inputs $(x_1, \ldots, x_k) \in \mathcal{X}_1 \times \cdots \times \mathcal{X}_k$ where player i has access to input x_j for all $j \neq i$ but no access to input x_i. For $k = 2$, this model coincides with the standard two-party communication model of Yao [Yao79], but when $k \geq 3$, the shared information between the players makes this model surprisingly powerful [Gro94, BGKL04, ACFN15, CS14], and fundamental problems remain open. For instance, a sufficiently strong lower bound for an explicit function F for $k \geq \mathrm{polylog}(n)$ players with $n = \log|\mathcal{X}_i|$ would imply a breakthrough result in complexity theory, namely a lower bound on the complexity class ACC^0.

NOF complexity of the Eval problem

A central open problem in the theory of NOF communication is to construct an explicit function for which randomized protocols are significantly more efficient than deterministic ones [BDPW07]. A well-studied candidate for this separation (for $k = 3$) is the function $\mathsf{Eval}_{\mathbb{F}_2^n}$, which is defined by $\mathsf{Eval}_{\mathbb{F}_2^n}(x_1, x_2, x_3) = 1$ if and only if $x_1 + x_2 + x_3 = 0$, where the additions are all in $\mathbb{F}_2^n$. Thus the Eval problem naturally generalizes the equality problem for $k = 2$. It is known that in the *randomized* setting, the standard protocol for the two-party equality problem that uses $\mathcal{O}(1)$ bits of communication works in the same way for three parties for the Eval problem. However, in the *deterministic* setting, the communication complexity $\mathrm{D}_3(\mathsf{Eval}_{\mathbb{F}_2^n})$ remains wide open: the best known lower bound $\Omega(\log\log n)$ follows from the work of Lacey and McClain [LM07] and, before this work, the best upper bound was $0.5n + \mathcal{O}(\log n)$ [ACFN15].

Corner problem in combinatorics, and connection to the Eval problem

Chandra, Furst and Lipton [CFL83] found that the deterministic communication complexity of many problems in the NOF model can be recast as Ramsey theory problems. In particular, and this leads to the problem of interest in this chapter, the (deterministic) communication complexity of $\mathsf{Eval}_{\mathbb{F}_2^n}$ can be characterized in terms of corner-free subsets of $\mathbb{F}_2^n \times \mathbb{F}_2^n$, as follows. Recall that any triple of elements $(x, y), (x + \lambda, y), (x, \lambda + y)$ for $x, y, \lambda \in \mathbb{F}_2^n$ a *corner*. A subset $S \subseteq \mathbb{F}_2^n \times \mathbb{F}_2^n$ is called *corner-free* if it does not contain any nontrivial corners (where nontrivial means that $\lambda \neq 0$). Denoting by $r_\angle(\mathbb{F}_2^n)$ the size of the largest corner-free set in $\mathbb{F}_2^n \times \mathbb{F}_2^n$, the communication complexity of $\mathsf{Eval}_{\mathbb{F}_2^n}$ equals

$\log(4^n/r_\angle(\mathbb{F}_2^n))$ up to a $\mathcal{O}(\log n)$ additive term, which provides the close connection between the Eval problem in NOF communication and the corner problem in combinatorics. In particular, large corner-free sets in $\mathbb{F}_2^n \times \mathbb{F}_2^n$ correspond to efficient protocols for $\mathsf{Eval}_{\mathbb{F}_2^n}$.

General paradigm: Shannon capacity of hypergraphs

The point of view we will take (and the general setting in which the methods we introduce will apply) is to regard the corner problem as a Shannon capacity problem of directed hypergraphs. Namely, the size $r_\angle(\mathbb{F}_2^n)$ of the largest corner-free set in $\mathbb{F}_2^n \times \mathbb{F}_2^n$ can be characterized as the independence number of a (naturally defined) directed 3-uniform hypergraph with 4^n vertices. This hypergraph has a recursive form: it is obtained by taking the n-th power of a fixed (directed) hypergraph $H_{\mathrm{cor},\mathbb{F}_2}$ on 4 vertices. (We discuss this in more detail in Section 3.2.1.) The asymptotic growth of $r_\angle(\mathbb{F}_2^n)$ as $n \to \infty$ is characterized by the Shannon capacity $\Theta(H_{\mathrm{cor},\mathbb{F}_2})$ of the corner hypergraph $H_{\mathrm{cor},\mathbb{F}_2}$. That is, we have $r_\angle(\mathbb{F}_2^n) = \Theta(H_{\mathrm{cor},\mathbb{F}_2})^{n-o(1)}$. In this way, proving the strict upper bound $\Theta(H_{\mathrm{cor},\mathbb{F}_2}) < 4$ is equivalent to proving a linear lower bound on the communication complexity of $\mathsf{Eval}_{\mathbb{F}_2^n}$.

A generalization of the multidimensional Szemerédi problem

Let $S \subseteq \mathbb{F}_p \times \mathbb{F}_p$ be an ordered nonempty set of size k, we can generalize the multidimensional Szemerédi question [FK79] over $\mathbb{F}_p^n \times \mathbb{F}_p^n$ as follows. A subset $A \subseteq \mathbb{F}_p^n \times \mathbb{F}_p^n$ is called *S-free* if for any k ordered distinct points $[(x^1, y^1), \ldots, (x^k, y^k)]$ of A, there is a coordinate $i \in [n]$ such that the ordered tuple $[(x_i^1, y_i^1), \ldots, (x_i^k, y_i^k)]$ does not have the form

$$\{(a,b) + (\overline{\lambda}u, \overline{\lambda}v) : (u,v) \in S\}, \text{ for some } (a,b) \in \mathbb{F}_p \times \mathbb{F}_p \text{ and } \overline{\lambda} \in \{1, \ldots, p-1\}.$$

Given an ordered nonempty set $S \subseteq \mathbb{F}_p \times \mathbb{F}_p$, the *generalized multidimensional Szemerédi* problem asks to determine the size of the largest S-free subset of $\mathbb{F}_p^n \times \mathbb{F}_p^n$. Many combinatorial problems can be considered as special cases of this problem including corner, cap set, square, Lshape (for explicit definitions, see Sections 3.2.2 and 3.2.3). For instance, it is easy to verify that, the corner problem over $\mathbb{F}_p^n \times \mathbb{F}_p^n$ (for $p = 2$) as a special case of this problem by choosing $S = \{(0,0), (1,0), (0,1)\}$. Moreover, for a fixed $S \subseteq \mathbb{F}_p \times \mathbb{F}_p$, the problem can be rephrased as determining the Shannon capacity of a fixed hypergraph. We will present this problem and its special cases in more detail in Section 3.2.

New results in this chapter

Improved lower bounds for some combinatorial problems

Our first result consists of new lower bounds for multiple combinatorial problems, including the corner, square, Lshape problems (for explicit definitions, see Sections 3.2.2 and 3.2.3) via a new technique to give lower bounds for the Shannon capacity of hypergraphs. Especially, from the new lower bounds for the corner problem over the groups $\mathbb{F}_2$ and $\mathbb{F}_3$, we obtain improved protocols for the Eval problem.

For a hypergraph H and any $m \in \mathbb{N}$, if the m-th power $H^{\boxtimes m}$ of a hypergraph H contains an independent set of size s, then the capacity $\Theta(H)$ is at least $s^{1/m}$. This was used for example in [LPS18] with $m = 2$ on $H_{\mathrm{cor},\mathbb{F}_2}$ and they found an independent set of size $s = 8$. We improve on this simple bound by observing that it is actually sufficient to

construct a set of size s which does not contain "cycles". In the context of graphs, the notion of cycle is clear but for hypergraphs there are many possible definitions. Here, to get new bounds we use the notion of *combinatorial degeneration* to model such a "cycle" (see Theorem 3.3.3). Combinatorial degeneration is a method from algebraic complexity theory [Str91], where it is used to construct fast matrix multiplication algorithms.

Using the combinatorial degeneration method on corner hypergraphs we find new bounds for corner problem. There are follows (in the three equivalent forms):

Theorem 3.1.1 (Thm. 3.3.6)**.** *For the corner and Eval problem over $\mathbb{F}_2^n$ we have:*

- $\mathrm{D}_3(\mathsf{Eval}_{\mathbb{F}_2^n}) \leq 0.24n + \mathcal{O}(\log n)$
- $r_{\angle}(\mathbb{F}_2^n) \geq \frac{3.39^n}{\mathrm{poly}(n)}$
- $\Theta(H_{\mathrm{cor},\mathbb{F}_2}) \geq 3.39$

Theorem 3.1.2 (Thm. 3.3.5)**.** *For the corner and Eval problem over $\mathbb{F}_3^n$ we have:*

- $\mathrm{D}_3(\mathsf{Eval}_{\mathbb{F}_3^n}) \leq 0.37n + \mathcal{O}(\log n)$
- $r_{\angle}(\mathbb{F}_3^n) \geq \frac{7^n}{\mathrm{poly}(n)}$
- $\Theta(H_{\mathrm{cor},\mathbb{F}_3}) \geq 7$.

In addition, we also obtain the best-known lower bounds for square and Lshape (see Section 3.2.3 for definitions) problems (see Table 3.1). We also introduce the notion of an acyclic set of a hypergraph (Section 3.3.2) which puts a stronger requirement (it implies a combinatorial degeneration) but might be simpler to check than combinatorial degeneration.

Limitations of current upper bound methods for Shannon capacity

Our second result is a strong limitation of current methods to effectively upper bound the Shannon capacity of hypergraphs. This limitation is caused by induced matchings and applies to various combinatorial problems including the corner problem. We use a method of Strassen to show that these limitations are indeed very strong for the corner problem.

In order to elaborate on these results let us first give an overview of upper bound methods. The general question of upper bounds on the Shannon capacity of hypergraphs is particularly well-studied in the special setting of undirected graphs. Even for undirected graphs, it is not clear how to compute the Shannon capacity in general, but some tools were developed to give upper bounds. The difficulty is to find a good upper bound on the largest independent set that behaves well under the product $\boxtimes$. For undirected graphs, the best known methods are the Lovász theta function [Lov79] and the Haemers bound which is based on the matrix rank [Hae79]. For hypergraphs, we only know of algebraic methods that are based on various notions of tensor rank, and in particular the slice rank [TS16], and similar notions like the analytic rank [GW11, Lov19], the geometric rank [KMZ20], and the G-stable rank [Der20]. Even though the slice rank is not multiplicative under $\boxtimes$ it is possible to give good upper bounds on the asymptotic slice rank via an asymptotic

analysis [TS16], which is closely related to the Strassen support functionals [Str91] or the more recent quantum functionals [CVZ18].

Most of the rank-based bounds actually give upper bounds on the size of induced matchings and not only on the size of independent sets. It is simple and instructive to see this argument in the setting of undirected graphs. For a given graph $H = (V, E)$, let A be the adjacency matrix in which we set all the diagonal coefficients to 1. Then for any independent set $I \subseteq V$, the submatrix $(A_{i,j})_{i,j\in I}$ of A is the identity matrix and as a result $|I| \leq \mathrm{rank}(A)$. As the matrix rank is multiplicative under tensor product, we get $\Theta(H) \leq \mathrm{rank}(A)$. Observe that this argument works equally well if we consider an induced matching instead of an independent set. An induced matching of size s of the graph $H = (V, E)$ can be defined by two lists of vertices $I_1(1), \ldots, I_1(s)$ and $I_2(1), \ldots, I_2(s)$ of size s such that for any $\alpha, \beta \in \{1, \ldots, s\}$ we have

$$((I_1(\alpha), I_2(\beta)) \in E \text{ or } I_1(\alpha) = I_2(\beta)) \quad \Longleftrightarrow \quad \alpha = \beta \ .$$

In other words, the submatrix $(A_{i,j})_{i\in I_1, j\in I_2}$ is an identity matrix, which also implies that $s \leq \mathrm{rank}(A)$. As such, the matrix rank is an upper bound on the asymptotic maximum induced matching. Tensor rank methods such as the subrank, slice rank, analytic rank, geometric rank and G-stable rank also provide upper bounds on the asymptotic maximum induced matching.

Using a result of Strassen [Str91], we show that there is an induced matching of the n-th power of $H_{\mathrm{cor},\mathbb{F}_2}$ of size $4^{n-o(1)}$. This establishes a *barrier* on many existing tensor methods (such as slice rank, subrank, analytic rank, etc.) to make progress on corner problem.

Theorem 3.1.3. *The hypergraph $H_{\mathrm{cor},\mathbb{F}_2}^{\boxtimes n}$ has an induced matching of size $4^{n-o(n)}$. In other words, for any $n \geq 1$, there exist lists $I_1, I_2, I_3 \subseteq \mathbb{F}_2^n \times \mathbb{F}_2^n$ of size $s(n) = 4^{n-o(n)}$ such that the following holds. For any $\alpha, \beta, \gamma \in \{1, \ldots, s(n)\}$*

$$(I_1(\alpha), I_2(\beta), I_3(\gamma)) \ \textit{forms a corner} \quad \Longleftrightarrow \quad \alpha = \beta = \gamma \ . \tag{3.2}$$

We prove this result by establishing in Theorem 3.4.3 that the adjacency tensor of the hypergraph $H_{\mathrm{cor},\mathbb{F}_2}$ is *tight* (see Definition 3.4.2). Strassen showed in [Str91] that for tight sets, the asymptotic induced matching is characterized by the support functionals. By computing the support functionals for the relevant tensors, we establish the claimed result in Corollary 3.4.6.

Related work. There have been many recent works on the rich connections between NOF communication complexity and problems in combinatorics, including the works by Shraibman [Shr18], Linial, Pitassi and Shraibman [LPS18], Viola [Vio19], Alon and Shraibman [AS20], and Linial and Shraibman [LS21]. This most recent result [LS21] constructs large corner-free sets in $[N] \times [N]$ by improving the best known NOF communication protocols for the $\textsc{exact}_T$ problem by a constant factor. We note that, as far as we know, these constructions do not carry over to the groups of the form G^n that we consider in this chapter. Regarding the study of the generalization of multidimensional Szemerédi theorem with polynomials in a finite field, there are several recent works by Peluse [Pel18, Pel19] and Kuca [Kuc19, Kuc21]. As for the topic of Shannon capacity of hypergraphs, the notably interesting paper [KM90] introduces and studies slightly different notion of Shannon

capacity than ours. Among the tensor methods that suffer from the induced matching barrier are: the slice rank [Tao16], analytic rank [GW11, Lov19, Bri19], geometric rank [KMZ20], and G-stable rank [Der20]. Slice rank was used and studied extensively in combinatorics, in the context of cap sets [Tao16, KSS16], sunflowers [NS17] and right-corners [Nas20].

Outline of the chapter. Section 3.2 briefly introduces the multidimensional Szemerédi theorem and the generalized multidimensional Szemerédi problem over the finite field, as well as some well-known special cases of this problem in additive combinatorics. In Section 3.3, we introduce three methods in for constructing lower bounds for the Shannon capacity of directed k-unifrom hypergraphs with emphasis on the combinatorial degeneration method. Lastly, in Section 3.4, we discuss the limitations of some current tensor rank methods to get upper bounds for the Shannon capacity of hypergraphs.

3.2 Multidimensional generalization of Szemerédi's theorem in the finite field

We start by recalling the following famous theorem of Szemerédi:

Theorem 3.2.1 (Szemerédi's theorem, [Sze75])**.** *Let $k \geq 1$ be a natural number, and let $\delta > 0$. Then if N is sufficiently large, every subset A of $[N]$ of cardinality $|A| \geq \delta N$ contains an arithmetic progression $a, a+r, \ldots, a+(k-1)r$ of length k, where $a \in \mathbb{Z}$ and r is a positive integer.*

Szemerédi's theorem was a major landmark in additive number theory for several reasons. Not only did it solve a well-known conjecture in the subject, the powerful methods introduced in order to prove this theorem have turned out to be extremely beneficial in a variety of other problems, as well as stimulating development and progress in several fields of mathematics. Also, the theorem itself has been applied to prove many other results; for example, it was a significant component in the argument of Green and Tao [GT08](a celebrated theorem in additive combinatorics) for demonstrating that the prime numbers contain arbitrarily long arithmetic progressions. Namely, they proved that if A is a subset of the primes of positive relative upper density, thus $\limsup_{N\to\infty} \pi(N)|A \cap [N]| > 0$, where $\pi(N)$ is the number of primes less than or equal to N, then A necessarily contains infinitely many arithmetic progressions of length k for all k. Additionally, generalizing Szemerédi's theorem, Bergelson and Leibman [BL96] states that if $A \subset [N]$ contains no progression of the form $x, x+P_1(y), \ldots, x+P_k(y)$, where $y \neq 0$ and $P_1, \ldots, P_k \in \mathbb{Z}[y]$ are polynomials with integer coefficients satisfying $P_1(0) = \cdots = P_k(0) = 0$. Then $|A| = o(N)$.

In [FK79], Furstenberg and Katznelson established a multidimensional generalization of Szemerédi's theorem and proved the question by using ergodic methods; after that, a combinatorial approach was developed in [Gow07] and also independently in [NRS06].

Theorem 3.2.2 (Multidimensional Szemerédi theorem, Theorem B in [FK79])**.** *Let $d \geq 1$ be a natural number, let $v_1, \ldots, v_k$ be elements of $\mathbb{Z}^d$, and let $\delta > 0$. Then if N is sufficiently large, every subset A of $[N]^d$ of cardinality $|A| \geq \delta N^d$ contains a set of the form $a + rv_1, \ldots, a + rv_k$, where $a \in \mathbb{Z}^d$ and r is a positive integer.*

Similar to Szemerédi theorem, the multidimensional Szemerédi theorem has had many extensive generalizations as well as variants. For instance, Cook, Magyar, and Titichetrakun in [CMT18] and Tao, Ziegler in [TZ15] independently proved a prime version of the multidimensional Szemerédi theorem. Shortly after, Fox-Zhao [FZ15] came up with a very short proof of the same result. In the finite field setting, Kuca [Kuc21] proved the following bound which can be seen as the finite field version of the multidimensional polynomial Szemerédi theorem of Bergelson and Leibman [BL96].

Theorem 3.2.3 ([Kuc21])**.** *Let $n, k \in \mathbb{N}_{\geq 1}$ and $v_1, \ldots, v_k \in \mathbb{Z}^n$ be nonzero vectors and $P_1, \ldots, P_k \in \mathbb{Z}[y]$ be polynomials satisfying $0 < \deg P_1 < \cdots < \deg P_k$. There exists constants $c, C > 0$ and a threshold $p_0 \in \mathbb{N}$ such that for all primes $p > p_0$, each subset $A \subseteq \mathbb{F}_p^n$ of size at least Cp^{n-c} contains*

$$x, x + v_1 P_1(y), \ldots, x + v_k P_k(y)\,,$$

for some $x \in \mathbb{F}_p^n$ and nonzero $y \in \mathbb{F}_p$.

A special case of the Theorem 3.2.3 is that each subset of $\mathbb{F}_p^2$ of size $\Omega(p^{2-c})$ contains a nontrivial configuration of the form

$$(x, y), (x + \lambda, y), (x, y + \lambda^2)\,,$$

previously proved in [HLY21] (nontrivial here mean $\lambda \neq 0$).

For the one-dimensional ($n = 1$) case, there are some works in [Pel18, Pel19, Kuc19], which provided upper bounds for some variants of the Szemerédi theorem in the finite field setting.

As we have seen so far, in the some works in [Kuc21], or in [DLS20, HLY21] for the generalization of the multidimensional Szemerédi question or in [Pel18, Pel19] for the one-dimensional over finite field $\mathbb{F}_p^n$, the authors consider the case when n is fixed and for p large enough. In this chapter, we work in a different manner. Namely, we consider the generalization of the multidimensional Szemerédi theorem over the finite field $\mathbb{F}_p^n$ when p is fixed and for n arbitrary large. Before introducing the problem, we first need to define some concepts.

Let $S \subseteq \mathbb{F}_p \times \mathbb{F}_p$ be an ordered nonempty set[2]. For any $\lambda \in \{0, 1, \ldots, p-1\}$, we denote $\lambda S = \{(\lambda x, \lambda y) : (x, y) \in S\}$, where λx is the sum of λ times the element x. A nonempty subset $S \subseteq \mathbb{F}_p \times \mathbb{F}_p$ is called *non-degenerate* if for all distinct pairs $(x, y), (x', y') \in S$, the following condition holds

$$(\lambda x, \lambda y) \neq (\lambda x', \lambda y') \text{ for all } \lambda \in \mathbb{F}_p \setminus \{0\}\,.$$

For $n \in \mathbb{N}_{\geq 1}$, let $S \subseteq \mathbb{F}_p \times \mathbb{F}_p$ be an ordered non-degenerate set of size k. We call a set of k ordered pairs $[(x^1, y^1), \ldots, (x^k, y^k)]$, where $(x^t, y^t) \in \mathbb{F}_p^n \times \mathbb{F}_p^n$ for all $t \in [k]$, is a *S-configuration* on $\mathbb{F}_p^n \times \mathbb{F}_p^n$ if there is a $(\lambda_1, \ldots, \lambda_n) \in \mathbb{F}_p^n \setminus \{0^n\}$ such that for all $i \in [n]$ each component $[(x_i^1, y_i^1), \ldots, (x_i^k, y_i^k)]$ (in ordered) form a $\{(a, b) + \lambda_i S\}$ for some $(a, b) \in \mathbb{F}_p \times \mathbb{F}_p$. A subset $A \subseteq \mathbb{F}_p^n \times \mathbb{F}_p^n$ is called *S-free* if it contain no k distinct points that form a S-configuaration. Let $r_S(\mathbb{F}_p^n)$ be a size of largest S-free set in $\mathbb{F}_p^n \times \mathbb{F}_p^n$. The *generalized multidimensional Szemerédi* problem asks to determine $r_S(\mathbb{F}_p^n)$ given S.

[2]The following construction can be generalized for $S \subseteq (\mathbb{F}_p)^{\times d}$ with arbitrary $d \in \mathbb{N}_{\geq 1}$.

The $r_S(\mathbb{F}_p^n)$ can be characterized as the independence number of a k-uniform hypergraph with p^{2n} vertices. Namely, consider a directed k-uniform hypergraph $H_{S,\mathbb{F}_p} = (V, E)$ with vertex set $V = \{(x, y) : x, y \in \mathbb{F}_p\}$ and $E = \{(x, y) + \lambda S : x, y, \lambda \in \mathbb{F}_p, \lambda \neq 0\}$. Then by construction:

Lemma 3.2.4. *$r_S(\mathbb{F}_p^n) = \alpha(H_{S,\mathbb{F}_p}^{\boxtimes n})$. As a consequence, $r_S(\mathbb{F}_p^n) = \Theta(H_{S,\mathbb{F}_p})^{n-o(n)}$, where $\Theta(H)$ is the Shannon capacity of the hypergraph H.*

Proof. By the construction of $H_{S,\mathbb{F}_p}$, we have an edge in $H_{S,\mathbb{F}_p}^{\boxtimes n}$ corresponding to a S-configuration on $\mathbb{F}_p^n \times \mathbb{F}_p^n$. Therefore each independent set in $H_{S,\mathbb{F}_p}^{\boxtimes n}$ corresponds to an S-free set in $\mathbb{F}_p^n \times \mathbb{F}_p^n$. This prove the claim. □

There are many well-known problems in combinatorics which are special cases of the generalized multidimensional Szemerédi problem over $\mathbb{F}_p^n \times \mathbb{F}_p^n$ and can be solved by taking specific non-degenerate sets in each case. Next, we will present some of these problems and their applications. However, in the following subsections, we will consider these problems over any finite Abelian group $(G, +)$ instead of restricting them to just $\mathbb{F}_p$.

3.2.1 The corner problem and number on the forehead communication

Corner problem

A *corner* in $G \times G$ is a three-element set of the form $\{(x, y), (x + \lambda, y), (x, y + \lambda)\}$ for some $x, y, \lambda \in G$ and $\lambda \neq 0$. The element (x, y) is called the center of this corner. Let $r_\angle(G)$ be the size of the largest subset $S \subseteq G \times G$ such that no three elements in S form a corner, and the set S is called *corner-free*. The corner problem asks to determine $r_\angle(G)$ given G.

Trivially, we have the upper bound $r_\angle(G) \leq |G|^2$. The best-known general upper bound on $r_\angle(G)$ comes from [Shk06a, Shk06b], and reads

$$r_\angle(G) \leq \frac{|G|^2}{(\log\log|G|)^c},$$

where $0 < c < \frac{1}{73}$ is an absolute constant. In the finite field setting, [LM07] obtained a better upper bound for $r_\angle(G)$ with $G = \mathbb{F}_2^n$ as follows:

$$r_\angle(\mathbb{F}_2^n) \leq \mathcal{O}\Big(|G|^2 \frac{\log\log\log|G|}{\log\log|G|}\Big).$$

We may phrase the corner problem as a hypergraph independence problem. We define $H_{\mathrm{cor},G} = (V, E)$ to be the directed 3-uniform hypergraph with $V = \{(g_1, g_2) : g_1, g_2 \in G\}$ and $E = \{((g_1, g_2), (g_1 + \lambda, g_2), (g_1, g_2 + \lambda)) : g_1, g_2, \lambda \in G,\ \lambda \neq 0\}$. Then by construction:

Lemma 3.2.5. *$r_\angle(G^n) = \alpha(H_{\mathrm{cor},G}^{\boxtimes n})$.*

As a consequence, $r_\angle(G^n) = \Theta(H_{\mathrm{cor},G})^{n-o(n)}$.

Remark 3.2.6. For $G = \mathbb{F}_p$, the corner problem is a special case of the generalized multidimensional Szemerédi problem by taking $S = \{(0, 0), (1, 0), (0, 1)\}$, and it is easy to verify that S is a non-degenerate set. In other words, $r_S(\mathbb{F}_p^n) = r_\angle(G^n) = \alpha(H_{\mathrm{cor},G}^{\boxtimes n})$ in this setting.

Example 3.2.7. Let G correspond to addition in $\mathbb{F}_2$. Then $H_{\mathrm{cor},G} = (V, E)$ with

$$E = \{((0,0),(1,0),(0,1)),((0,1),(1,1),(0,0)),((1,0),(0,0),(1,1)),((1,1),(0,1),(1,0))\}.$$

Under the labeling $(0,0) = 0, (0,1) = 1, (1,0) = 2$ and $(1,1) = 3$, we will think of $H_{\mathrm{cor},\mathbb{F}_2}$ as the hypergraph $H_{\mathrm{cor},\mathbb{F}_2} = (V, E)$ with $V = (0,1,2,3)$ and $E = \{(0,2,1),(1,3,0),(2,0,3),(3,1,2)\}$.

Closely related to $r_\angle(G)$ is the minimum number of colors needed to color $G \times G$ so that no corner is monochromatic, which we denote by $c_\angle(G)$. Then:

Proposition 3.2.8 ([CFL83, LPS18])**.** *Let $(G,+)$ be a finite Abelian group. There is a constant c, such that for every $n \in \mathbb{N}$,*

$$\frac{|G|^{2n}}{r_\angle(G^n)} \le c_\angle(G^n) \le c\frac{n|G|^{2n}\log|G|}{r_\angle(G^n)}\,.$$

For $G = \mathbb{F}_2$, the current upper bound in the literature is $c_\angle(\mathbb{F}_2^n) \le \mathcal{O}(n2^{n/2})$ [LPS18], which we will improve later on.

Number on the forehead communication

The number on the forehead (NOF) model of communication [CFL83] is very rich both in terms of connections to Ramsey theory and additive combinatorics [BGG06, Shr18, LS21], as well as applications to boolean models of compution such as branching programs and boolean circuits [CFL83, BT94]. In this model, k players wish to evaluate a function $F : \mathcal{X}_1 \times \cdots \times \mathcal{X}_k \to \{0,1\}$ on a given input $x_1, \ldots, x_k$. The input is distributed among the players in a way that player i sees every x_j for $j \neq i$. This scenario is visualized as x_i being written on the forehead of Player i. The computational power of everyone is unlimited, but the number of exchanged bits has to be minimized. Let $\mathrm{D}_k(F)$ be the minimum number of bits they need to communicate to compute the function F in the NOF model with k players. For $k = 2$, this model corresponds to the standard two-party communication model [Yao79], but when $k \ge 3$, the shared information between the players makes this model surprisingly powerful [Gro94, BGKL04, ACFN15, CS14]. Fundamental problems remain open. For instance, a sufficiently strong lower bound for an explicit function F for $k \ge \mathrm{polylog}(n)$ players with $n = \log|\mathcal{X}_i|$ implies a breakthrough result in complexity theory, namely a lower bound on the complexity class ACC^0.

The Eval problem over Abelian group. A central open problem in NOF model is to construct an explicit function for which randomized protocols are significantly more efficient than deterministic ones [BDPW07]. A candidate for this separation (for $k = 3$) is the function Eval_{G^n}, a natural generalization of the equality problem, defined by $\mathsf{Eval}_{G^n}(x_1, x_2, x_3) = 1$ if and only if $x_1 + x_2 + x_3 = 0$. In the *randomized* setting, the standard protocol for the two-party equality problem that uses $\mathcal{O}(1)$ (for fixed group G) bits of communication works in the same way for three parties for the Eval problem. However, in the *deterministic* setting, the communication complexity $\mathrm{D}_3(\mathsf{Eval}_{G^n})$ remains wide open. For $G = \mathbb{F}_2$, the best known lower bound $\Omega(\log\log n)$ follows from [LM07] and the best upper bound was $0.5n + \mathcal{O}(\log n)$ [ACFN15]. On the other hand, for two players, Yao [Yao79] proved that $\mathrm{D}_2(\mathsf{Eval}_{G^n}) = \Omega(n)$ (for nontrivial G). But, it is an open problem whether $\mathrm{D}_3(\mathsf{Eval}_{G^n}) = \Omega(n)$ for three players.

Chandra, Furst and Lipton [CFL83] found that the deterministic communication complexity of many problems in the NOF model can be recast as Ramsey theory problems. In particular, the deterministic communication complexity of Eval_{G^n} problem can be characterized in terms of corner-free sets in $G^n \times G^n$. More precisely, as a generalization a result in [CFL83] (Theorem 4.2 in [CFL83]), [BGG06] showed that:

Lemma 3.2.9 ([BGG06])**.** $\log(c_{\angle}(G^n)) \leq \mathrm{D}_3(\mathsf{Eval}_{G^n}) \leq 2 + \log(c_{\angle}(G^n))$.

From Lemma 3.2.9 and Proposition 3.2.8, it follows that $\Theta(H_{\mathrm{cor},G}) < |G|^2$ would imply that $\mathrm{D}_3(\mathsf{Eval}_{G^n}) = \Omega(n)$, and also that lower bounds on $r_{\angle}(G^n)$ give upper bounds on $\mathrm{D}_3(\mathsf{Eval}_{G^n})$. For $G = \mathbb{F}_2$, the best-known upper bound on $\mathrm{D}_3(\mathsf{Eval}_{\mathbb{F}_2^n})$ is $0.5n + \mathcal{O}(\log n)$ [ACFN15] which we improve later on.

Putting all the claims till now together, the open problem that motivates our work in this chapter, and that is central in NOF communication complexity and combinatorics, asks:

Problem 3.2.10. Are the following three equivalent statements true?

- $\mathrm{D}_3(\mathsf{Eval}_{G^n}) = \Omega(n)$
- $r_{\angle}(G^n) \leq \mathcal{O}(c^n)$ for some $c < |G|^2$
- $\Theta(H_{\mathrm{cor},G}) < |G|^2$.

3.2.2 The capset problem

A *three-term arithmetic progression* in G is a three-element set of the form $\{x, x+\lambda, x+2\lambda\}$ for some $x, \lambda \in G$ and $\lambda \neq 0$. Let $r_3(G)$ be the size of the largest subset $S \subseteq G$ such that no three elements in S form a three-term arithmetic progression.

A three-term-arithmetic-progression-free subset of $\mathbb{F}_3^n$ is also called a *cap set*. The notorious cap set problem is to determine how $r_3(\mathbb{F}_3^n)$ grows when n goes to infinity. A priori we have that $2^n \leq r_3(\mathbb{F}_3^n) \leq 3^n$. Using Fourier methods and the density increment argument of Roth, the upper bound $r_3(\mathbb{F}_3^n) \leq \mathcal{O}(3^n/n)$ was obtained by Meshulam [Mes95], and improved only as late as 2012 to $\mathcal{O}(3^n/n^{1+\epsilon})$ for some positive constant ϵ by Michael Bateman and Nets Hawk Katz in [BK12]. Until recently it was not known whether $r_3(\mathbb{F}_3^n)$ grows like $3^{n-o(n)}$ or like $c^{n-o(n)}$ for some $c < 3$. Gijswijt and Ellenberg solved this question in 2017, showing that $r_3(\mathbb{F}_3^n) \leq 2.756^{n+o(n)}$ [EG17]. The best lower bound is $2.2174^n \leq r_3(\mathbb{F}_3^n)$ by Edel [Ede04]. In particular, using Lemma 3.2.12, this implies the lower bound $3^n \cdot 2.2174^n = 6.6522^n \leq r_{\angle}(\mathbb{F}_3^n)$ for the corner problem. We will improve this lower bound in Theorem 3.3.5.

We may phrase the cap set problem as a Shannon capacity of hypergraph $H_{\mathrm{cap}} = (V, E)$ with $V = \{0, 1, 2\}$ and $E = \{(a, a+\lambda, a+2\lambda)\}$ for all $a, \lambda \in \mathbb{F}_3$ and $\lambda \neq 0$. It is easy to verify that the H_{cap} is undirected 3-uniform hypergraph with one edge $E = \{(0, 1, 2)\}$. By the construction, the independence number $\alpha(H_{\mathrm{cap}}^{\boxtimes n})$ equals $r_3(\mathbb{F}_3^n)$, and thus the Shannon capacity of H_{cap} determines the rate of growth of $r_3(\mathbb{F}_3^n)$.

Remark 3.2.11. By choosing $S = \{(0,0), (0,1), (0,2)\} \subseteq \mathbb{F}_3 \times \mathbb{F}_3$. Then, the cap set problem is a special case of the generalized multidimensional Szemerédi problem.

Moreover, following [Zha19, Corollary 3.24], there is a simple relation between corner-free sets and three-term-arithmetic-progression-free sets:

Lemma 3.2.12. $p^n\, r_3(\mathbb{F}_p^n) \leq r_\angle(\mathbb{F}_p^n)$

Proof. Let $S \subseteq \mathbb{F}_p^n$ be a subset that is free of three-term arithmetic progressions. Define the subset $T = \{(x, y) : x - y \in S\}$. Then T is a corner-free set of size $p^n|S|$. Indeed, if $(x, y), (x + \lambda, y), (x, y + \lambda)$ are elements of T, then $x - y, x + \lambda - y, x - y - \lambda$ are in S and these elements form a three-term arithmetic progression. □

3.2.3 The Lshape and square problems

There are some other combinatorial problems that can be obtained by taking some special non-degenerate sets $S \subseteq \mathbb{F}_p \times \mathbb{F}_p$. For instance, the *Lshape* is a set of points $\{(x, y), (x + \lambda, y), (x, y + \lambda), (x, y + 2\lambda)\}$ for all $x, y, \lambda \in \mathbb{F}_p^n$ and $\lambda \neq 0$. Denote $r_{\mathrm{L}}(\mathbb{F}_p^n)$ to be the size of largest subset of $\mathbb{F}_p^n \times \mathbb{F}_p^n$ that does not contain any nontrivial configuration of the form $\{(x, y), (x + \lambda, y), (x, y + \lambda), (x, y + 2\lambda)\}$. This problem of determining $r_{\mathrm{L}}(\mathbb{F}_p^n)$ is called the *Lshape* problem. The Lshape problem has been studied in [Pel22], and the author has shown the following reasonable upper bounds for $r_{\mathrm{L}}(\mathbb{F}_p^n)$

$$r_{\mathrm{L}}(\mathbb{F}_p^n) \leq \frac{p^{2n}}{\log_m n}\,,$$

for some large constant m. Now, consider a $S = \{(0, 0), (1, 0), (0, 1), (0, 2)\} \subseteq \mathbb{F}_p \times \mathbb{F}_p$, it is easy to verify that S is non-degenerate. Then the Lshape problem can be seen as a special case of generalized multidimensional Szemerédi problem. In other words, we have $r_{\mathrm{L}}(\mathbb{F}_p^n) = r_S(\mathbb{F}_p^n)$. We denote by $H_{\mathrm{Lshape},\mathbb{F}_p}$ the directed 4-uniform hypergraph corresponding to the Lshape problem over $\mathbb{F}_p^n \times \mathbb{F}_p^n$.

Another shape we can consider is the *square* shape which is obtained by taking a non-degenerate set $S = \{(0, 0), (1, 0), (0, 1), (1, 1)\} \subseteq \mathbb{F}_p \times \mathbb{F}_p$. Let $r_\square(\mathbb{F}_p^n)$ be the size of largest subset of $\mathbb{F}_p^n \times \mathbb{F}_p^n$ that does not contain nontrivial configuration of the form $\{(x, y), (x + \lambda, y), (x, y + \lambda), (x + \lambda, y + \lambda)\}$ for some $x, y, \lambda \in \mathbb{F}_p^n$ and $\lambda \neq 0^n$. It is easy to verify that $r_\square(\mathbb{F}_p^n) = r_S(\mathbb{F}_p^n)$. The problem of determining $r_\square(\mathbb{F}_p^n)$ is often referred to as the *square* problem. Denote by $H_{\mathrm{square},\mathbb{F}_p}$ the directed 4-uniform hypergraph corresponding to the square problem.

The square problem can also be considered as a generalization of Graham's question [Erd73] over $\mathbb{F}_p^n$. Namely, in an integer grid $[N] \times [N]$, is it true that for any $\delta > 0$, there exists $N_0 = N_0(\delta)$ such that for any set $A \subseteq [N] \times [N]$, of size at least δN where $N > N_0$, one can find a quadruple of the form $\{(x, y), (x+\lambda, y), (x, y+\lambda), (x+\lambda, y+\lambda)\}$ in A for some $\lambda \neq 0$. An affirmative answer for this question can be found in [BB07, Pre15]. For more information, we refer the reader to [Shk15, SS21].

3.3 Lower bounds on Shannon capacity from combinatorial degenerations

We discuss three methods to obtain lower bounds on the Shannon capacity of directed k-uniform hypergraphs: the combinatorial degeneration method, the acyclic set method

and the probabilistic method. We apply these methods to the problems of constructing large corner-free, Lshape-free, square-free sets. As a consequence of the construction of corner-free set, we provide a better NOF communication protocols for the Eval function.

3.3.1 Combinatorial degeneration method

We first introduce the combinatorial degeneration method for lower bounding the Shannon capacity. Combinatorial degeneration is an existing concept from algebraic complexity theory introduced by Strassen in [Str87, Section 6, in particular Theorem 6.1][3]. In that original setting it was used as part of the construction of fast matrix multiplication algorithms [BCS97, Definition 15.29 and Lemma 15.31], and, in a broader setting, combinatorial degeneration was used to construct large induced matchings in [ASU13, Lemma 3.9], [AW18, Lemma 5.1] and [CVZ18, Theorem 4.11]. However, we will be using it in a novel manner in order to construct independent sets instead of induced matchings. We will subsequently apply the combinatorial degeneration method to get new bounds for the corner, square, Lshape problems. We expect the method to be useful in the study of other problems besides these problems as well. First we must define combinatorial degeneration.

Definition 3.3.1 (Combinatorial degeneration)**.** Let $I_1, \ldots, I_k$ be finite sets. 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 every $x = (x_1, \ldots, x_k) \in I_1 \times \cdots \times I_k$, if $x \in \Psi \setminus \Phi$, then $\sum_{i=1}^k u_i(x_i) > 0$, and if $x \in \Phi$, then $\sum_{i=1}^k u_i(x_i) = 0$.

Example 3.3.2. As a quick example of a combinatorial degeneration, let

$$\Phi = \{(0,0,0), (1,1,0), (1,0,1)\},$$
$$\Psi = \{(0,0,0), (1,1,0), (1,0,1), (0,1,1)\}.$$

Then we have a combinatorial degeneration $\Psi \trianglerighteq \Phi$ by picking the maps $u_1(0) = u_2(0) = u_3(0) = 0$, and $u_1(1) = -1$, $u_2(1) = u_3(1) = 1$.

We apply combinatorial degeneration in the following fashion to get Shannon capacity lower bounds:

Theorem 3.3.3 (Combinatorial degeneration method)**.** *Let $H = (V, E)$ be a directed k-uniform hypergraph. Let $S \subseteq V$. Let $\Psi = E \cup \{(v, \ldots, v) : v \in V\}$ and let $\Phi = \{(v, \ldots, v) : v \in S\}$ and suppose that $\Psi \trianglerighteq \Phi$. Then $\Theta(H) \geq |S|$.*

Proof. Let u_i be the maps given by the combinatorial degeneration $\Psi \trianglerighteq \Phi$. Let n be any multiple of $|S|$. Let $(x^{(1)}, \ldots, x^{(k)}) \in \Psi^{\otimes n}$. Suppose for every $i \in [k]$ that the n elements in the tuple $x^{(i)} = (x_1^{(i)}, \ldots, x_n^{(i)})$ are uniformly distributed over S, so that every element of S appears $n/|S|$ times in $x^{(i)}$. Then, using that $\sum_{i=1}^k u_i(s) = 0$ for every $s \in S$ and the

[3]The precise connection to [Str87] is as follows. Strassen defines the notion of *M-degeneration* on tensors. In our terminology, a tensor is an M-degeneration of another tensor, if the support of the first is a combinatorial degeneration of the support of the second. The terminology "combinatorial degeneration", which does not refer to tensors, but rather directly to their supports (hence the adjective "combinatorial"), was introduced in [BCS97, Definition 15.29].

uniformity of $x^{(i)}$, we have

$$\sum_{i=1}^{k}\sum_{j=1}^{n} u_i(x_j^{(i)}) = \frac{n}{|S|}\sum_{s\in S}\sum_{i=1}^{k} u_i(s) = 0. \tag{3.3}$$

For every $j \in [n]$, since $(x_j^{(1)}, \ldots, x_j^{(k)}) \in \Psi$, we have $\sum_{i=1}^{k} u_i(x_j^{(i)}) \geq 0$. Suppose that there is an index $j \in [n]$ such that $(x_j^{(1)}, \ldots, x_j^{(k)}) \notin \Phi$. Then

$$\sum_{i=1}^{k} u_i(x_j^{(i)}) > 0.$$

As a consequence, $\sum_{j=1}^{n}\sum_{i=1}^{k} u_i(x_j^{(i)}) > 0$, which contradicts (3.3). Thus the uniform strings in S^n form an independent set in $H^{\boxtimes n}$. There are

$$\binom{|S|\frac{n}{|S|}}{\frac{n}{|S|}, \ldots, \frac{n}{|S|}} \geq \frac{|S|^n}{(n+1)^{|S|}}$$

such strings. The inequality follows from the fact that the largest multinomial coefficient is the central one, i.e., $\binom{n}{n_1,\ldots,n_{|S|}} \leq \binom{n}{\frac{n}{|S|},\ldots,\frac{n}{|S|}}$ and the number of possible partitions of n into $|S|$ parts is at most $(n+1)^{|S|}$. □

Remark 3.3.4. The above proof of Theorem 3.3.3 gives in fact the precise lower bound

$$\alpha(H^{\boxtimes n}) \geq \frac{|S|^n}{(n+1)^{|S|}}. \tag{3.4}$$

This lower bound is optimal up to a poly(n) factor. The following more careful analysis improves this poly(n) factor, but may safely be skipped when the reader is satisfied by the lower bound of (3.4).

We may without loss of generality assume that $S = V$. For $p \in \mathbb{Z}$, let $[V^n]_p^{(i)} \subseteq V^n$ be the subset of all elements $(x_1, \ldots, x_n) \in V^n$ such that $\sum_{j=1}^{n} u_i(x_j) = p$. For $p_1, \ldots, p_k \in \mathbb{Z}$, we let $[\Psi^{\otimes n}]_{p_1,\ldots,p_k} \subseteq \Psi^{\otimes n}$ denote the subset of all elements $(x^{(1)}, \ldots, x^{(k)}) \in \Psi^{\otimes n}$ such that for every $i \in [k]$ we have $\sum_{j=1}^{n} u_i(x_j^{(i)}) = p_i$. Thus $[\Psi^{\otimes n}]_{p_1,\ldots,p_k} = \Psi^{\otimes n} \cap ([V^n]_{p_1}^{(1)} \times \cdots \times [V^n]_{p_1}^{(k)})$. Then

$$\Psi^{\otimes n} = \bigsqcup_{p_1,\ldots,p_k} [\Psi^{\otimes n}]_{p_1,\ldots,p_k}$$

and from the definition of a combinatorial degeneration we get

$$\Phi^{\otimes n} = \bigsqcup_{\substack{p_1,\ldots,p_k:\\ \sum_{i=1}^{k} p_i = 0}} [\Psi^{\otimes n}]_{p_1,\ldots,p_k}. \tag{3.5}$$

Since $\Phi^{\otimes n}$ only contains elements of the form $(x, \ldots, x)$, we see that if $[\Psi^{\otimes n}]_{p_1,\ldots,p_k} \neq \emptyset$ and $\sum_{i=1}^{k} p_i = 0$, then the elements of $[\Psi^{\otimes n}]_{p_1,\ldots,p_k}$ are all the elements $(x, \ldots, x)$ going over all $x \in \cap_{i=1}^{k} [V^n]_{p_i}^{(i)}$. Thus $\alpha(H^{\otimes n}) \geq |[\Psi^{\otimes n}]_{p_1,\ldots,p_k}|$ for any choice of $p_1, \ldots, p_k$ such that $\sum_{i=1}^{k} p_i = 0$.

One good choice of $p_1, \dots, p_k$ is obtained as follows, and lets us recover the lower bound in (3.4). For notational simplicity we are still assuming $S = V$. Let $(x_1, \dots, x_n) \in V^n$ be any element that is uniform on S. For every $i \in [k]$ let $p_i = \sum_{j=1}^n u_i(x_j)$. Note that for every $i \in [k]$ the value of p_i remains the same if we had picked another uniform element $(x_1, \dots, x_n) \in V^n$. We claim that $\sum_{i=1}^k p_i = 0$. To prove this, let $(x^{(1)}, \dots, x^{(k)}) \in \Psi^n$ be any element for which every $x^{(i)}$ is uniform on S. Then in the same way as in (3.3) we have $p_1 + \cdots + p_k = \sum_i \sum_j u_i(x_j^{(i)}) = 0$, using that for every $s \in S$ we have $\sum_i u_i(s) = 0$. Finally, note that $[\Psi^{\otimes n}]_{p_1,\dots,p_k}$ contains all elements $(x^{(1)}, \dots, x^{(k)}) \in \Psi^{\otimes n}$ for which every $x^{(i)}$ is uniform. Therefore, with this choice we recover a bound that is at least as good as (3.4).

Another choice of $p_1, \dots, p_k$ (that leads to an incomparable lower bound) is obtained as follows. Note that if $[\Psi^{\otimes n}]_{p_1,\dots,p_k} \neq \emptyset$, then $n \min_{x \in V} u_i(x) \leq p_i \leq n \max_{x \in V} u_i(x)$. Thus the number of nonzero summands in (3.5) is at most $c_{|S|} n^{k-1}$ for a constant $c_{|S|}$ that depends only on $|S|$. Therefore, there is a choice of $p_1, \dots, p_k$ with $\sum_{i=1}^k p_i = 0$ such that

$$\alpha(H^{\boxtimes n}) \geq \left|[\Psi^{\otimes n}]_{p_1,\dots,p_k}\right| \geq \frac{|\Phi^{\otimes n}|}{c_{|S|} n^{k-1}} = \frac{|S|^n}{c_{|S|} n^{k-1}}$$

which improves on (3.4) in some parameter regimes.

In order to construct combinatorial degenerations we employ integer linear programming. For any directed k-uniform hypergraph $H = (V, E)$, we naturally define $\beta(H)$ to be the size of the largest subset $S \subseteq V$ such that $\{(v, \dots, v) : v \in S\}$ is a combinatorial degeneration of $E \cup \{(v, \dots, v) : v \in V\}$. Clearly, $\Theta(H) \geq \beta(H)$ by Theorem 3.3.3.

To state the integer program, we let t be a variable that takes values in $\{0,1\}^{|V|}$ and let $u_1, \dots, u_k$ be variables that take values in $\mathbb{Z}^{|V|}$. We choose $M \in \mathbb{N}$ large enough. The parameter $\beta(H)$ can be then computed by the following integer linear program:

$$\boxed{\begin{array}{rrll} \max & \sum_{i \in V} t(i) & & \\ \text{subject to} & u_1(i_1) + \cdots + u_k(i_k) & \geq 1 & \forall (i_1, \dots, i_k) \in E, \\ & 1 - t(i) \leq u_1(i) + \cdots + u_k(i) & \leq M(1 - t(i)) & \forall i \in V \end{array}} \tag{3.6}$$

Indeed, if $(t, u_1, \dots, u_k)$ is a feasible solution of the program (3.6), then $\{(v, \dots, v) : v \in S\}$ is a combinatorial degeneration of $E \cup \{(v, \dots, v) : v \in V\}$ by choosing k integer maps $u_1, \dots, u_k$, where $S = \{i \in V : t(i) = 1\}$. Therefore, one has $\beta(H) \geq A$ (A is a maximum value of program (3.6)). On the other hand, for any $S \subseteq V$ such that if there is a combinatorial degeneration from $E \cup \{(v, \dots, v) : v \in V\}$ to $\{(v, \dots, v) : v \in S\}$ with k integer maps $u_1, \dots, u_k$, by defining $t \in \{0,1\}^{|V|}$ so that $t(i) = 1$ iff $i \in S$, we have $(t, u_1, \dots, u_k)$ is a feasible solution of the program (3.6). Thus, $\beta(H) \leq A$.

As a first application of the combinatorial degeneration method, we prove the following new bound for corners over $\mathbb{F}_3^n$ and $\mathbb{F}_2^n$.

Theorem 3.3.5. $\Theta(H_{\mathrm{cor},\mathbb{F}_3}) \geq 7$.

In other words, $7^n/\mathrm{poly}(n) \leq r_\angle(\mathbb{F}_3^n)$. This improves on the lower bound $6.6522^n \leq r_\angle(\mathbb{F}_3^n)$ that can be obtained from Edel's construction of cap sets [Ede04] and Lemma 3.2.12.

As a consequence of the new lower bound, we find the bounds $c_\angle(\mathbb{F}_3^n) \leq \mathcal{O}(\mathrm{poly}(n)(\frac{9}{7})^n)$ and $\mathrm{D}_3(\mathsf{Eval}_{\mathbb{F}_3^n}) \leq n\log(9/7) + \mathcal{O}(\log n) \leq 0.37n + \mathcal{O}(\log n)$. Previously, only the weaker bound $\mathrm{D}_3(\mathsf{Eval}_{\mathbb{F}_3^n}) \leq n + \mathcal{O}(\log n)$ was known [LPS18].[4]

Proof. Let Ψ be the support of the adjacency tensor of $H_{\mathrm{cor},\mathbb{F}_3}$. We label each pair (a,b) for $a,b \in \{0,1,2\}$ by the integer number $3a+b$. The hypergraph $H_{\mathrm{cor},\mathbb{F}_3}$ has vertex set $V = \{0,1,3,4,5,6,7,8\}$ and the set Ψ is given by

$$\begin{aligned}\Psi = \{&(0,0,0),(1,1,1),(2,2,2),(3,3,3),(4,4,4),(5,5,5),(6,6,6),(7,7,7),(8,8,8),\\ &(0,3,1),(0,6,2),(1,4,2),(1,7,0),(2,5,0),(2,8,1),(3,6,4),(3,0,5),(4,7,5),\\ &(4,1,3),(5,8,3),(5,2,4),(6,0,7),(6,3,8),(7,1,8),(7,4,6),(8,2,6),(8,5,7)\}\,.\end{aligned}$$

Let $S \subseteq V(H_{\mathrm{cor},\mathbb{F}_3})$ be the subset consisting of the following seven vertices:

$$S \coloneqq \{0,1,2,3,4,7,8\}\,.$$

One directly verifies that the maps $u_i : V \to \mathbb{Z}$ provided in the following table give a combinatorial degeneration from Ψ to $\Phi_S \coloneqq \{(v,v,v) : v \in S\}$.

vertex	u_1	u_2	u_3
0	-5	1	4
1	-5	4	1
2	-5	1	4
3	-5	5	0
4	-3	2	1
5	-1	5	5
6	-1	5	5
7	-3	2	1
8	-5	5	0

We conclude that $\Theta(H_{\mathrm{cor},\mathbb{F}_3}) \geq 7$. □

In the previous proof we only considered the first power of the relevant hypergraph. For the next result we will be able to get good bounds by considering higher powers.

Theorem 3.3.6. *$\beta(H_{\mathrm{cor},\mathbb{F}_2}^{\boxtimes 2}) \geq 11$ and $\beta(H_{\mathrm{cor},\mathbb{F}_2}^{\boxtimes 3}) \geq 39$, as a consequence $\Theta(H_{\mathrm{cor},\mathbb{F}_2}) \geq 3.39$.*

In other words, $3.39^n/\mathrm{poly}(n) \leq r_\angle(\mathbb{F}_2^n)$. As a consequence, we have the upper bound $c_\angle(\mathbb{F}_2^n) \leq \mathcal{O}(\mathrm{poly}(n)1.18^n)$ for the corner problem and the upper bound $\mathrm{D}_3(\mathsf{Eval}_{\mathbb{F}_2^n}) \leq 0.24n + \mathcal{O}(\log n)$ for the Eval problem.

Proof. Let $H = H_{\mathrm{cor},\mathbb{F}_2} \boxtimes H_{\mathrm{cor},\mathbb{F}_2}$. We will show $\beta(H) \geq 11$, which implies $\Theta(H_{\mathrm{cor},\mathbb{F}_2}) \geq \sqrt{11}$. Let Ψ be the support of the adjacency tensor of H. Then Ψ is this rather large set

[4] We note that the NOF protocol in [ACFN15] for $\mathsf{Eval}_{\mathbb{F}_2^n}$ does not in any direct way generalize to $\mathsf{Eval}_{\mathbb{F}_3^n}$ as far as we know.

of 64 triples:

$$\begin{aligned}\Psi = \{&((0,0),(0,0),(0,0)),\ ((0,1),(0,1),(0,1)),\ ((0,2),(0,2),(0,2)),\ ((0,3),(0,3),(0,3)),\\ &((1,0),(1,0),(1,0)),\ ((1,1),(1,1),(1,1)),\ ((1,2),(1,2),(1,2)),\ ((1,3),(1,3),(1,3)),\\ &((2,0),(2,0),(2,0)),\ ((2,1),(2,1),(2,1)),\ ((2,2),(2,2),(2,2)),\ ((2,3),(2,3),(2,3)),\\ &((3,0),(3,0),(3,0)),\ ((3,1),(3,1),(3,1)),\ ((3,2),(3,2),(3,2)),\ ((3,3),(3,3),(3,3)),\\ &((0,0),(0,2),(0,1)),\ ((0,0),(2,0),(1,0)),\ ((0,0),(2,2),(1,1)),\ ((0,1),(0,3),(0,0)),\\ &((0,1),(2,1),(1,1)),\ ((0,1),(2,3),(1,0)),\ ((0,2),(0,0),(0,3)),\ ((0,2),(2,0),(1,3)),\\ &((0,2),(2,2),(1,2)),\ ((0,3),(0,1),(0,2)),\ ((0,3),(2,1),(1,2)),\ ((0,3),(2,3),(1,3)),\\ &((1,0),(1,2),(1,1)),\ ((1,0),(3,0),(0,0)),\ ((1,0),(3,2),(0,1)),\ ((1,1),(1,3),(1,0)),\\ &((1,1),(3,1),(0,1)),\ ((1,1),(3,3),(0,0)),\ ((1,2),(1,0),(1,3)),\ ((1,2),(3,0),(0,3)),\\ &((1,2),(3,2),(0,2)),\ ((1,3),(1,1),(1,2)),\ ((1,3),(3,1),(0,2)),\ ((1,3),(3,3),(0,3)),\\ &((2,0),(0,0),(3,0)),\ ((2,0),(0,2),(3,1)),\ ((2,0),(2,2),(2,1)),\ ((2,1),(0,1),(3,1)),\\ &((2,1),(0,3),(3,0)),\ ((2,1),(2,3),(2,0)),\ ((2,2),(0,0),(3,3)),\ ((2,2),(0,2),(3,2)),\\ &((2,2),(2,0),(2,3)),\ ((2,3),(0,1),(3,2)),\ ((2,3),(0,3),(3,3)),\ ((2,3),(2,1),(2,2)),\\ &((3,0),(1,0),(2,0)),\ ((3,0),(1,2),(2,1)),\ ((3,0),(3,2),(3,1)),\ ((3,1),(1,1),(2,1)),\\ &((3,1),(1,3),(2,0)),\ ((3,1),(3,3),(3,0)),\ ((3,2),(1,0),(2,3)),\ ((3,2),(1,2),(2,2)),\\ &((3,2),(3,0),(3,3)),\ ((3,3),(1,1),(2,2)),\ ((3,3),(1,3),(2,3)),\ ((3,3),(3,1),(3,2))\}.\end{aligned}$$

Let $S \subseteq V(H)$ be the subset consisting of the following eleven vertices:

$$S \coloneqq \{(0,0),(0,1),(1,0),(1,2),(1,3,)(2,0),(2,1),(2,2),(3,1),(3,2),(3,3)\}\,.$$

One directly verifies that the maps $u_i : \{0,1,2,3\}^2 \to \mathbb{Z}$ provided in the following table give a combinatorial degeneration from Ψ to $\Phi_S \coloneqq \{(v,v,v) : v \in S\}$.

vertex	u_1	u_2	u_3
$(0,0)$	-10	0	10
$(0,1)$	-10	0	10
$(0,2)$	10	1	-1
$(0,3)$	10	1	1
$(1,0)$	-10	0	10
$(1,1)$	-8	10	10
$(1,2)$	7	3	-10
$(1,3)$	1	5	-6
$(2,0)$	-5	1	4
$(2,1)$	-6	1	5
$(2,2)$	9	1	-10
$(2,3)$	10	3	-7
$(3,0)$	-3	1	10
$(3,1)$	-8	1	7
$(3,2)$	8	1	-9
$(3,3)$	9	-1	-8

We make it easier to verify this by listing every element $e = (v_1, v_2, v_3) \in \Psi$ again, together with the evaluation $(u_1(v_1), u_2(v_2), u_3(v_3))$, the sum of the evaluations $\sum_i u_i(v_i)$, and whether e is in Φ or in $\Psi \setminus \Phi$:

(v_1, v_2, v_3)	(u_1, u_2, u_3)	$\sum_i u_i$	
((0,0), (0,0), (0,0))	$(-10, 0, 10)$	0	Φ
((0,1), (0,1), (0,1))	$(-10, 0, 10)$	0	Φ
((0,2), (0,2), (0,2))	$(10, 1, -1)$	10	$\Psi \setminus \Phi$
((0,3), (0,3), (0,3))	$(10, 1, 1)$	12	$\Psi \setminus \Phi$
((1,0), (1,0), (1,0))	$(-10, 0, 10)$	0	Φ
((1,1), (1,1), (1,1))	$(-8, 10, 10)$	12	$\Psi \setminus \Phi$
((1,2), (1,2), (1,2))	$(7, 3, -10)$	0	Φ
((1,3), (1,3), (1,3))	$(1, 5, -6)$	0	Φ
((2,0), (2,0), (2,0))	$(-5, 1, 4)$	0	Φ
((2,1), (2,1), (2,1))	$(-6, 1, 5)$	0	Φ
((2,2), (2,2), (2,2))	$(9, 1, -10)$	0	Φ
((2,3), (2,3), (2,3))	$(10, 3, -7)$	6	$\Psi \setminus \Phi$
((3,0), (3,0), (3,0))	$(-3, 1, 10)$	8	$\Psi \setminus \Phi$
((3,1), (3,1), (3,1))	$(-8, 1, 7)$	0	Φ
((3,2), (3,2), (3,2))	$(8, 1, -9)$	0	Φ
((3,3), (3,3), (3,3))	$(9, -1, -8)$	0	Φ
((0,0), (0,2), (0,1))	$(-10, 1, 10)$	1	$\Psi \setminus \Phi$
((0,0), (2,0), (1,0))	$(-10, 1, 10)$	1	$\Psi \setminus \Phi$
((0,0), (2,2), (1,1))	$(-10, 1, 10)$	1	$\Psi \setminus \Phi$
((0,1), (0,3), (0,0))	$(-10, 1, 10)$	1	$\Psi \setminus \Phi$
((0,1), (2,1), (1,1))	$(-10, 1, 10)$	1	$\Psi \setminus \Phi$
((0,1), (2,3), (1,0))	$(-10, 3, 10)$	3	$\Psi \setminus \Phi$
((0,2), (0,0), (0,3))	$(10, 0, 1)$	11	$\Psi \setminus \Phi$
((0,2), (2,0), (1,3))	$(10, 1, -6)$	5	$\Psi \setminus \Phi$
((0,2), (2,2), (1,2))	$(10, 1, -10)$	1	$\Psi \setminus \Phi$
((0,3), (0,1), (0,2))	$(10, 0, -1)$	9	$\Psi \setminus \Phi$
((0,3), (2,1), (1,2))	$(10, 1, -10)$	1	$\Psi \setminus \Phi$
((0,3), (2,3), (1,3))	$(10, 3, -6)$	7	$\Psi \setminus \Phi$
((1,0), (1,2), (1,1))	$(-10, 3, 10)$	3	$\Psi \setminus \Phi$
((1,0), (3,0), (0,0))	$(-10, 1, 10)$	1	$\Psi \setminus \Phi$
((1,0), (3,2), (0,1))	$(-10, 1, 10)$	1	$\Psi \setminus \Phi$
((1,1), (1,3), (1,0))	$(-8, 5, 10)$	7	$\Psi \setminus \Phi$

(v_1, v_2, v_3)	(u_1, u_2, u_3)	$\sum_i u_i$	
((1,1), (3,1), (0,1))	$(-8, 1, 10)$	3	$\Psi \setminus \Phi$
((1,1), (3,3), (0,0))	$(-8, -1, 10)$	1	$\Psi \setminus \Phi$
((1,2), (1,0), (1,3))	$(7, 0, -6)$	1	$\Psi \setminus \Phi$
((1,2), (3,0), (0,3))	$(7, 1, 1)$	9	$\Psi \setminus \Phi$
((1,2), (3,2), (0,2))	$(7, 1, -1)$	7	$\Psi \setminus \Phi$
((1,3), (1,1), (1,2))	$(1, 10, -10)$	1	$\Psi \setminus \Phi$
((1,3), (3,1), (0,2))	$(1, 1, -1)$	1	$\Psi \setminus \Phi$
((1,3), (3,3), (0,3))	$(1, -1, 1)$	1	$\Psi \setminus \Phi$
((2,0), (0,0), (3,0))	$(-5, 0, 10)$	5	$\Psi \setminus \Phi$
((2,0), (0,2), (3,1))	$(-5, 1, 7)$	3	$\Psi \setminus \Phi$
((2,0), (2,2), (2,1))	$(-5, 1, 5)$	1	$\Psi \setminus \Phi$
((2,1), (0,1), (3,1))	$(-6, 0, 7)$	1	$\Psi \setminus \Phi$
((2,1), (0,3), (3,0))	$(-6, 1, 10)$	5	$\Psi \setminus \Phi$
((2,1), (2,3), (2,0))	$(-6, 3, 4)$	1	$\Psi \setminus \Phi$
((2,2), (0,0), (3,3))	$(9, 0, -8)$	1	$\Psi \setminus \Phi$
((2,2), (0,2), (3,2))	$(9, 1, -9)$	1	$\Psi \setminus \Phi$
((2,2), (2,0), (2,3))	$(9, 1, -7)$	3	$\Psi \setminus \Phi$
((2,3), (0,1), (3,2))	$(10, 0, -9)$	1	$\Psi \setminus \Phi$
((2,3), (0,3), (3,3))	$(10, 1, -8)$	3	$\Psi \setminus \Phi$
((2,3), (2,1), (2,2))	$(10, 1, -10)$	1	$\Psi \setminus \Phi$
((3,0), (1,0), (2,0))	$(-3, 0, 4)$	1	$\Psi \setminus \Phi$
((3,0), (1,2), (2,1))	$(-3, 3, 5)$	5	$\Psi \setminus \Phi$
((3,0), (3,2), (3,1))	$(-3, 1, 7)$	5	$\Psi \setminus \Phi$
((3,1), (1,1), (2,1))	$(-8, 10, 5)$	7	$\Psi \setminus \Phi$
((3,1), (1,3), (2,0))	$(-8, 5, 4)$	1	$\Psi \setminus \Phi$
((3,1), (3,3), (3,0))	$(-8, -1, 10)$	1	$\Psi \setminus \Phi$
((3,2), (1,0), (2,3))	$(8, 0, -7)$	1	$\Psi \setminus \Phi$
((3,2), (1,2), (2,2))	$(8, 3, -10)$	1	$\Psi \setminus \Phi$
((3,2), (3,0), (3,3))	$(8, 1, -8)$	1	$\Psi \setminus \Phi$
((3,3), (1,1), (2,2))	$(9, 10, -10)$	9	$\Psi \setminus \Phi$
((3,3), (1,3), (2,3))	$(9, 5, -7)$	7	$\Psi \setminus \Phi$
((3,3), (3,1), (3,2))	$(9, 1, -9)$	1	$\Psi \setminus \Phi$

Indeed, we see that $\sum_i u_i(v_i)$ is always nonnegative, and equals 0 if and only if $(v_1, v_2, v_3) \in \Phi$. Therefore we obtain $\beta(H) \geq 11$.

For the construction to prove $\beta(H^{\boxtimes 3}_{\mathrm{cor},\mathbb{F}_2^n}) \geq 39$. Let $S \subseteq V(H)$ be the subset consisting of the following thirty-nine vertices:

$$\begin{aligned} S \coloneqq \{&(0,0,0), (0,0,1), (0,0,2), (0,1,0), (0,1,2), (0,1,3), (0,2,1), (0,2,3), (0,3,0), (0,3,1), \\ &(0,3,3), (1,0,2), (1,1,0), (1,1,1), (1,1,3), (1,2,1), (1,2,2), (1,2,3), (1,3,1), (1,3,2), \\ &(1,3,3), (2,0,0), (2,0,1), (2,0,3), (2,1,2), (2,1,3), (2,2,0), (2,2,2), (2,3,0), (2,3,2), \\ &(2,3,3), (3,0,0), (3,0,1), (3,0,2), (3,1,2), (3,2,0), (3,2,1), (3,2,3), (3,3,2)\}. \end{aligned}$$

One directly verifies that the maps $u_i : \{0,1,2,3\}^3 \to \mathbb{Z}$ provided in the following table give a combinatorial degeneration from $\Psi = \{(v,v,v) : v \in V(H^{\boxtimes 3}_{\mathrm{cor},\mathbb{F}_2^n})\} \cup E(H^{\boxtimes 3}_{\mathrm{cor},\mathbb{F}_2^n})$ to $\Phi_S \coloneqq \{(v,v,v) : v \in S\}$.

vertex	u_1	u_2	u_3
$(0,0,0)$	1	-4	3
$(0,0,1)$	-15	-4	19
$(0,0,2)$	-3	-17	20
$(0,0,3)$	20	20	20
$(0,1,0)$	-5	7	-2
$(0,1,1)$	20	20	20
$(0,1,2)$	-5	20	-15
$(0,1,3)$	1	0	-1
$(0,2,0)$	20	5	20
$(0,2,1)$	5	-4	-1
$(0,2,2)$	20	20	7
$(0,2,3)$	1	18	-19
$(0,3,0)$	-4	3	1
$(0,3,1)$	20	-20	0
$(0,3,2)$	20	20	20
$(0,3,3)$	20	-20	0
$(1,0,0)$	20	20	20
$(1,0,1)$	20	20	20
$(1,0,2)$	-2	-14	16
$(1,0,3)$	9	20	20
$(1,1,0)$	-16	20	-4
$(1,1,1)$	0	1	-1
$(1,1,2)$	20	20	20
$(1,1,3)$	-4	5	-1
$(1,2,0)$	20	20	20
$(1,2,1)$	-5	1	4
$(1,2,2)$	18	2	-20
$(1,2,3)$	20	0	-20
$(1,3,0)$	20	20	20
$(1,3,1)$	-4	-11	15
$(1,3,2)$	1	2	-3
$(1,3,3)$	20	-15	-5

vertex	u_1	u_2	u_3
$(2,0,0)$	5	-5	0
$(2,0,1)$	-4	4	0
$(2,0,2)$	20	20	20
$(2,0,3)$	-15	5	10
$(2,1,0)$	20	20	20
$(2,1,1)$	20	20	20
$(2,1,2)$	-9	7	2
$(2,1,3)$	1	1	-2
$(2,2,0)$	2	5	-7
$(2,2,1)$	20	20	20
$(2,2,2)$	9	1	-10
$(2,2,3)$	20	20	20
$(2,3,0)$	-4	0	4
$(2,3,1)$	20	20	20
$(2,3,2)$	4	-10	6
$(2,3,3)$	12	-9	-3
$(3,0,0)$	15	-16	1
$(3,0,1)$	-19	6	13
$(3,0,2)$	-3	-17	20
$(3,0,3)$	3	20	20
$(3,1,0)$	-1	19	20
$(3,1,1)$	20	20	13
$(3,1,2)$	-17	-3	20
$(3,1,3)$	20	20	20
$(3,2,0)$	-2	4	-2
$(3,2,1)$	8	7	-15
$(3,2,2)$	9	20	20
$(3,2,3)$	10	0	-10
$(3,3,0)$	20	20	20
$(3,3,1)$	20	20	20
$(3,3,2)$	-7	-2	9
$(3,3,3)$	20	20	20

□

We can see that any lower bound for corner problem is also a lower bound for the Lshape and square problems. In Table 3.1, we compute the values of $\beta(H)$ and independence number $\alpha(H)$ for hypergraphs $H = \{\text{corner}, \text{square}, \text{Lshape}\}$ over $G = \mathbb{F}_p$ on some small values of p. As far as we know, these values in the table are the best-known lower bounds for these problems.

	corner		square		Lshape	
p	$\beta(H_{\text{cor},\mathbb{F}_p})$	$\alpha(H_{\text{cor},\mathbb{F}_p})$	$\beta(H_{\text{square},\mathbb{F}_p})$	$\alpha(H_{\text{square},\mathbb{F}_p})$	$\beta(H_{\text{Lshape},\mathbb{F}_p})$	$\alpha(H_{\text{Lshape},\mathbb{F}_p})$
2	3	2	3	3		
3	7	6	7	6	7	6
4	12	8	12	10	12	8
5	19	12	20	16	21	15
7	37	22	38	28	42	28
8	48	26	48	37	48	32
9	61	36	62	45	63	41

Table 3.1: The value of $\beta(H)$ and $\alpha(H)$ for $H = \{\text{corner}, \text{square}, \text{Lshape}\}$ over $\mathbb{F}_p$ for $p < 10$.

Moreover, over $\mathbb{F}_3$ we can compute the $\beta(H^{\boxtimes 2})$ for hypergraph $H = \{\text{square}, \text{Lshape}\}$

to obtain a better lower bound for the square and Lshape problems. Namely,

Lemma 3.3.7. *We have,* $\beta(H^{\boxtimes 2}_{\text{square},\mathbb{F}_3}) = 58$ *and* $\beta(H^{\boxtimes 2}_{\text{Lshape},\mathbb{F}_3}) = 59$*. As a consequence,* $r_{\square}(\mathbb{F}_3^n) \geq 7.61^n/\text{poly}(n)$*,* $r_{\text{L}}(\mathbb{F}_3^n) \geq 7.68^n/\text{poly}(n)$*.*

For any directed k-uniform hypergraph H, it is clear that $\alpha(H) \leq \beta(H)$. The combinatorial degeneration method has shown improvements on the lower bounds for corner, square, Lshape hypergraphs compared to using the independence numbers in the Table 3.1. On undirected hypergraphs (e.g. cap set hypergraph), we show that there is no difference between the combinatorial degeneration method and the independence number method.

Theorem 3.3.8. *For any undirected k-uniform hypergraph* H*, then* $\beta(H) = \alpha(H)$*.*

Proof. It is clear that $\alpha(H) \leq \beta(H)$ for any undirected k-uniform hypergraph H. Let S be a set such that there is combinatorial degeneration from $V(H) \cup E(H)$ to S. Let $u_1, \ldots, u_k$ be maps corresponding to the combinatorial degeneration. For any k vertices $v_1, \ldots, v_k$ in S, we will prove $(v_1, \ldots, v_k) \notin E(H)$. Indeed, assume that there is an edge $(v_1, \ldots, v_k)$, since H is an undirected hypergraph, so for any $\pi \in \mathfrak{S}_k$ we have $(v_{\pi(1)}, \ldots, v_{\pi(k)})$ also in $E(H)$. On the other hand, since $(v_1, \ldots, v_k) \in E(H)$, this implies $\sum_{i=1}^k u_i(v_{\pi(i)}) > 0$ for any $\pi \in \mathfrak{S}_k$, thus $\sum_{\pi \in \mathfrak{S}_k} \sum_{i=1}^k u_i(v_{\pi(i)}) > 0$, which contradicts with $\sum_{i=1}^k u_i(v) = 0$ for all $v \in S$. Therefore, the set S is an independent set of H and $\alpha(H) \geq \beta(H)$, this proves the claim.

□

We have yet to develop structural understanding of how the above combinatorial degenerations that exhibit the new capacity lower bounds arise, and leave the investigation of further generalizations and improvements to future work. As a partial remedy to our limited understanding, we introduce in the next section the *acyclic method* as a tool to construct combinatorial degenerations. While the acyclic method does not recover the bounds of Theorem 3.3.6 and Theorem 3.3.5, it has the merits of being transparent and simple to apply.

3.3.2 Acyclic set method

The acyclic set method that we are about to introduce is modeled on the fact that the Shannon capacity of a directed graph G is at least the size of any induced acyclic subgraph of G [BM85]. We introduce the concept of an *acyclic set* in a directed k-uniform hypergraph as an extension of the notion of an induced acyclic subgraph.

Definition 3.3.9. Let H be a directed k-uniform hypergraph. We associate to H the directed graph G_H with vertices $V(G) = V(H)$ and edges $E(G) = \{(a_1, a_2) : (a_1, a_2, \ldots, a_k) \in E \text{ for some } a_3, \ldots, a_k\}$. For any subset $A \subseteq V$ let $H[A]$ denote the subhypergraph of H induced by A, that is, $H[A]$ is the directed k-uniform hypergraph with vertices S and edges $E \cap A^{\times k}$. We call a subset $A \subseteq V$ an *acyclic set* of H if the directed graph $G_{H[A]}$ is a directed acyclic graph.

Note that, if A is an independent set of H, then $E(H[A]) = \emptyset$ and thus $E(G_{H[A]}) = \emptyset$, and in particular A is an acyclic set of H. On the other hand, acyclic sets are not necessarily independent sets. However, the existence of an acyclic set does imply strong lower bounds on the Shannon capacity (via combinatorial degeneration, as we will see):

Theorem 3.3.10. *Let H be a directed k-uniform hypergraph. For any acyclic set A of H, we have $\Theta(H) \geq |A|$.*

Theorem 3.3.10 follows directly from the combinatorial degeneration method (Theorem 3.3.3) and the following lemma:

Lemma 3.3.11. *Let $H = (V, E)$ be a directed k-uniform hypergraph. Let A be an acyclic set of H. Then there is a combinatorial degeneration from $E \cup \{(v, \ldots, v) : v \in V\}$ to $\Phi = \{(v, \ldots, v) : v \in A\}$.*

Proof. We may assume that $A = V = [n]$. The proof for the case that $A \subsetneq V$ is a simple adaptation. Recall that we construct the directed graph G associated to H with the same vertex set as H and the edges as follows: for every edge $e = (a_1, a_2, \ldots, a_k)$ in H we add the edge (a_1, a_2) to G. Since V is an acyclic set we have that G is a directed acyclic graph. Therefore, we have a topological ordering on the vertices of G. A topological ordering is a total ordering $>$ on the vertices such that if (u, v) forms an edge then $u < v$. Assume that this ordering is $1 > 2 > \cdots > n$. For each vertex $i \in [n]$, we define $u_1(i) = -i$, $u_2(i) = i$, $u_3(i) = \cdots = u_k(i) = 0$. For every $i \in [n]$ we clearly have $u_1(i) + u_2(i) + \cdots + u_k(i) = 0$. For each edge $e = (a_1, a_2, \ldots, a_k)$ in H we have $u_1(a_1) + u_2(a_2) + \cdots + u_k(a_k) > 0$ because of the topological ordering and since we have the edge (a_1, a_2) in G. Therefore we have a combinatorial degeneration from $E \cup \{(v, \ldots, v) : v \in V\}$ to $\{(v, \ldots, v) : v \in V\}$. For the case $A \subsetneq V$ the proof is similar except that we define $u_1(i), u_2(i), \ldots, u_k(i)$ to be some large integer number for each $i \in V \setminus A$. □

As can be seen from the proof of Lemma 3.3.11, the combinatorial degenerations that result from acyclic sets have a special form, and in particular the acyclic set method does not recover the full power of the combinatorial degeneration method. However the acyclic set method is much easier to apply than the combinatorial degeneration method. For example, we can use the acyclic set method to quickly see that $\Theta(H_{\mathrm{cor},\mathbb{F}_2}) \geq 3$. Namely, it is verified directly that the set $S = \{0, 1, 2\}$ of size three is an acyclic set in $H_{\mathrm{cor},\mathbb{F}_2}$, which implies the claim by Theorem 3.3.10.

Finally, we note that for directed graphs ($k = 2$) the combinatorial degeneration method can be used to characterize whether the Shannon capacity is full or not.

Theorem 3.3.12. *Let $G = (V, E)$ be a directed graph. Then $\Theta(G) = |V|$ if and only if there is a combinatorial degeneration from $E \cup \{(v, v) : v \in V\}$ to $\{(v, v) : v \in V\}$.*

Proof. The *if* direction follows directly from Theorem 3.3.3. For the *only if* direction, it is shown in [BM85] that if $\Theta(G) = |V|$, then G is an acyclic graph. Then, applying Lemma 3.3.11 for $k = 2$ proves the claim. □

Remark 3.3.13. We note that for any directed graph $G = (V, E)$, by Theorem 3.3.12 and Lemma 3.3.10, the graph G is acyclic if and only if there is a combinatorial degeneration from $\{(v, v) : v \in V\} \cup E$ to $\{(v, v) : v \in V\}$. In other words, the notion of combinatorial degeneration in directed k-uniform hypergraphs can be seen as a generalization of the notion of acyclicity in directed graphs.

3.3.3 Probabilistic method

We finish this section with a simple and general method for obtaining lower bounds on the Shannon capacity. For any element $g \in G$, the set $\{(g, g+\lambda) : \lambda \in G\}$ is an independent set of $H_{\mathrm{cor},G}$, and therefore we have $\Theta(H_{\mathrm{cor},G}), \Theta(H_{\mathrm{square},G}), \Theta(H_{\mathrm{Lshape},G})$ are at least $|G|$, which we think of as the trivial lower bound. By using a simple probabilistic argument (which does not use much of the structure of $H_{\mathrm{cor},G}$), we show the following nontrivial lower bound for $\Theta(H_{\mathrm{cor},G}), \Theta(H_{\mathrm{square},G}), \Theta(H_{\mathrm{Lshape},G})$.

Proposition 3.3.14. *For any finite Abelian group G, we have all quantities $\Theta(H_{\mathrm{cor},G})$, $\Theta(H_{\mathrm{square},G})$, $\Theta(H_{\mathrm{Lshape},G})$ are at least $|G|^{3/2}$.*

Proof. Let $|G| = m$ and $n \in \mathbb{N}$. We will show that $\Theta(H_{\mathrm{cor},G}) \geq |G|^{3/2}$, as a consequence we also have $\Theta(H_{\mathrm{square},G}), \Theta(H_{\mathrm{Lshape},G})$ are at least $|G|^{3/2}$.

Recall that the hypergraph $H_{\mathrm{cor},G}^{\boxtimes n}$ has vertices given by the elements of $G^n \times G^n$ and edges given by the corners in $G^n \times G^n$. Let $p = 1/\sqrt{3(m^n - 1)}$ and choose the subset A of $V(H_{\mathrm{cor},G}^{\boxtimes n})$ randomly by choosing any element $(g_1, g_2) \in G^n \times G^n$ to be in the set A with probability p. Let H_A be the directed subhypergraph of $H_{\mathrm{cor},G}^{\boxtimes n}$ induced by A. We have $\mathbb{E}[|V(H_A)|] = m^{2n}p$. Let e be any edge of $H_{\mathrm{cor},G}^{\boxtimes n}$. Then e is of the form

$$e = \big((g_1, g_2), (g_1 + \lambda, g_2), (g_1, g_2 + \lambda)\big)$$

for some $g_1, g_2, \lambda \in G^n$ and $\lambda \neq 0$. Since (g_1, g_2), $(g_1 + \lambda, g_2)$ and $(g_1, g_2 + \lambda)$ are different, and for each the probability of being in A is p, we have that $\Pr[e \in E(H_A)] = p^3$. Therefore, since $|E(H_{\mathrm{cor},G}^{\boxtimes n})| = m^{2n}(m^n - 1)$, we have $\mathbb{E}[|E(H_A)|] = m^{2n}(m^n - 1)p^3$. On the other hand, for any hypergraph H we have $\alpha(H) \geq |V(H)| - |E(H)|$. Therefore

$$\alpha(H_{\mathrm{cor},G}^{\boxtimes n}) \geq \mathbb{E}[|V(H_A)|] - \mathbb{E}[|E(H_A)|] = \frac{2m^{2n}}{3\sqrt{3(m^n - 1)}}\,.$$

Thus find the lower bound $\Theta(H_{\mathrm{cor},G}) = \lim_{n\to\infty} \alpha(H_{\mathrm{cor},G}^{\boxtimes n})^{1/n} \geq m^{3/2}$. □

The idea in the proof of Proposition 3.3.14 to apply the probabilistic method to lower bound the number of remaining elements after a "pruning" procedure (in this case, pruning vertices that induce edges) goes back to [CW87]. A similar probabilistic method construction is the driving component in the recent new upper bound on the matrix multiplication exponent ω [AW21].

In terms of the corner problem, the lower bound on the Shannon capacity in Proposition 3.3.14 for $G = \mathbb{F}_2$ corresponds to the upper bound $c_{\angle}(\mathbb{F}_2^n) \leq \mathcal{O}(n2^{n/2})$ (via Proposition 3.2.8). This upper bound is similar to the bound provided in [LPS18, Corollary 26 in the ITCS version].

Remark 3.3.15. The proof of Proposition 3.3.14 directly extends from 2-dimensional corners to k-dimensional corners, which are sets of the form

$$\{(x_1, x_2, \ldots, x_k), (x_1 + \lambda, x_2, \ldots, x_k), \ldots, (x_1, x_2, \ldots, x_k + \lambda)\}.$$

From a similar probabilistic method argument as in the proof of Proposition 3.3.14, choosing each $(x_1, \ldots, x_k) \in (G^n)^{\times k}$ independently at random with probability $p = \frac{1}{[(k+1)(|G|^n-1)]^{1/k}}$, we get

$$r_{k,\angle}(G^n) \geq \frac{k|G|^{kn}}{|G|^{n/k}(k+1)^{\frac{k+1}{k}}},$$

where $r_{k,\angle}(G^n)$ is the size of the largest k-dimensional corner free set in $G^{n\times k}$. As a consequence one has $\Theta(H_{k,\mathrm{cor},G}) \geq |G|^{k-1/k}$, where $H_{k,\mathrm{cor},G}$ is directed $(k+1)$-uniform hypergraph that construct for the k-dimensional corner.

Just like the Eval problem on 3 players is closely related to 2-dimensional corners in $(G^n)^{\times 2}$, the Eval function on $k+1$ players is closely related to k-dimensional corners in $(G^n)^{\times k}$. By a similar argument as the proof of Lemma 3.2.9 we have that the $k+1$ player NOF complexity is upper bounded by $\mathrm{D}_{k+1}(\mathsf{Eval}_{G^n}) \leq k + c_{k,\angle}(G^n)$, where $c_{k,\angle}(G^n)$ is minimum number of colors that we can use to color $(G^n)^{\times k}$ such that no k-dimensional corner is monochromatic. We also have similar to Proposition 3.2.8 the relation between $r_{k,\angle}(G^n)$ and $c_{k,\angle}(G^n)$ given by

$$\frac{|G|^{kn}}{r_{k,\angle}(G^n)} \leq c_{k,\angle}(G^n) \leq \frac{nk|G|^n \log(|G|)}{r_{k,\angle}(G^n)},$$

which is proved in [LPS18]. Furthermore from the lower bound of $r_{k,\angle}(G^n)$, we have

$$\mathrm{D}_{k+1}(\mathsf{Eval}_{G^n}) \leq \frac{n}{k}\log|G| + \log n + \log\log|G| + (1+\frac{1}{k})\log(1+k) + k\,.$$

If we take $k = \log n$ (for instance), then $\mathrm{D}_{k+1}(\mathsf{Eval}_{G^n}) \leq \frac{n}{\log n}\log|G| + \mathcal{O}(\log n)$, that is, we obtain a sublinear upper bound for $\mathrm{D}_{\log n}(\mathsf{Eval}_{G^n})$ in n.

We note that when $G = \mathbb{F}_2$, a better lower bound for $r_{k,\angle}(\mathbb{F}_2^n)$ with $k = \log n$ is given by [ACFN15]. Namely, the authors showed that $r_{k,\angle}(\mathbb{F}_2^n) \geq \frac{2^{nk}}{n^{c\log n}}$ for some constant $c > 0$.

3.4 Limitations of current upper bound methods for Shannon capacity

Our result in this section is a strong limitation of current tensor rank methods to effectively upper bound the Shannon capacity of hypergraphs. This limitation is caused by induced matchings and applies to various combinatorial problems including the corner problem. The main point is to describe the induced matching barrier and apply it to the corner problem.

3.4.1 Induced matchings and tightness

Now we discuss the notion of induced matchings and tight sets. Then we will discuss Strassen's theorem that gives a construction of large induced matchings under a tightness condition.

Let $H = (V, E)$ be a directed k-uniform hypergraph with adjacency tensor A. Let Φ_H be the support of A. A subset $D \subseteq \Phi_H$ is called a *matching* if any two distinct elements

$a, b \in D$ differ in all k coordinates, that is, $a_i \neq b_i$ for all $i \in [k]$. We call a matching $D \subseteq \Phi_H$ an *induced matching* if $D = \Phi_H \cap (D_1 \times \cdots \times D_k)$, where $D_i = \{a_i : a \in D\}$ is the projection of D onto the i-th coordinate. We denote by $\mathrm{Q}_{\mathrm{IM}}(\Phi_H)$ the maximum size of an induced matching $D \subseteq \Phi_H$. The quantity $\mathrm{Q}_{\mathrm{IM}}(\Phi_H)$ is called *induced matching number* of H.

For two directed k-uniform hypergraphs $G = (V_G, E_G)$ and $H = (V_H, E_H)$, let Φ_G and Φ_H be the support of the adjacency tensors of G and H, respectively. We define the product $\Phi_G \times \Phi_H \subseteq (V_G \times V_H) \times \cdots \times (V_G \times V_H)$ by $\Phi_G \times \Phi_H \coloneqq \{((a_1, b_1), \ldots, (a_k, b_k)) : a \in \Phi_G, b \in \Phi_H\}$. The *asymptotic induced matching number* of H is defined as $\tilde{\mathrm{Q}}_{\mathrm{IM}}(\Phi_H) \coloneqq \lim_{n\to\infty} \mathrm{Q}_{\mathrm{IM}}(\Phi_H^{\times n})^{1/n} = \sup_n \mathrm{Q}_{\mathrm{IM}}(\Phi_H^{\times n})^{1/n}$.

The induced matching number should be thought of as the combinatorial version of the subrank, which was introduced in section 2.2, as follows. Let Φ_H be the support of the adjacency tensor A_H of a directed k-uniform hypergraph H. Then the induced matching number $\mathrm{Q}_{\mathrm{IM}}(\Phi_H)$ is the largest number n such that $\langle n \rangle$ can be obtained from A_H using a restriction that consists of matrices that have at most one nonzero entry in each row and in each column. Therefore, $\mathrm{Q}_{\mathrm{IM}}(\Phi_H) \leq \mathrm{Q}(A_H)$.

Lemma 3.4.1. *Let H be a directed k-uniform hypergraph and A_H be its adjacency tensor with support $\Phi_H = \mathrm{supp}(A_H)$. Then*

$$\Theta(H) \leq \tilde{\mathrm{Q}}_{\mathrm{IM}}(\Phi_H) \leq \tilde{\mathrm{Q}}(A_H).$$

Proof. We begin with the first inequality. Let S be an independent set of $H^{\boxtimes n}$. We have $\Phi_H^{\times n} = \mathrm{supp}(A_H^{\otimes n})$. Thus $\Phi_H^{\times n} \cap (S \times S \cdots \times S) = \{(a, \ldots, a) : a \in S\}$. This means that $|S| \leq \mathrm{Q}_{\mathrm{IM}}(\Phi_H^{\times n})$. We conclude $\Theta(H) \leq \tilde{\mathrm{Q}}_{\mathrm{IM}}(\Phi_H)$. The second inequality follows from the already established inequality $\mathrm{Q}_{\mathrm{IM}}(\Phi_H) \leq \mathrm{Q}(A_H)$. □

Next, we discuss *tight sets*, a notion introduced by Strassen [Str91].

Definition 3.4.2 ([Str91], see also [CVZ18])**.** Let $I_1, \ldots, I_k$ be finite sets. We call any subset $\Phi \subseteq I_1 \times \cdots \times I_k$ *tight* if there are injective maps $u_i : I_i \to \mathbb{Z}$ for every $i \in [k]$ such that:

$$u_1(a_1) + \cdots + u_k(a_k) = 0 \text{ for every } (a_1, \ldots, a_k) \in \Phi.$$

When Φ_H is tight, the asymptotic induced matching number is essentially known, and can be described as a simple optimization. To explain the precise formula we recall some definitions.

For any finite set X, let $\mathcal{P}(X)$ be the set of all distributions on X. For any probability distribution $P \in \mathcal{P}(X)$ the *Shannon entropy* of P is defined as $H(P) \coloneqq -\sum_{x\in X} P(x) \log_2 P(x)$ with $0 \log_2 0 = 0$. Given finite sets $I_1, \ldots, I_k$ and a probability distribution $P \in \mathcal{P}(I_1 \times \cdots \times I_k)$ on the product set $I_1 \times \cdots \times I_k$ we denote the *marginal distribution* of P on I_i by P_i, that is, $P_i(a) = \sum_{x:x_i=a} P(x)$ for any $a \in I_i$.

Theorem 3.4.3 ([Str91])**.** *Let H be a directed 3-uniform hypergraph. If Φ_H is tight, then*

$$\tilde{\mathrm{Q}}_{\mathrm{IM}}(\Phi_H) = \max_{P\in\mathcal{P}(\Phi_H)} \min_{i\in[3]} 2^{H(P_i)}.$$

In particular, Theorem 3.4.3 implies that, for any directed 3-uniform hypergraph $H = (V, E)$ if there is a distribution P on Φ_H such that every marginal distribution P_i is uniform on V, then Φ_H has asymptotically maximal induced matchings.

Note that Theorem 3.4.3 only applies to directed k-uniform hypergraphs for $k = 3$. For the higher-order case $k > 3$, an extension of the lower bound of Theorem 3.4.3 was proven in [CVZ19a, Theorem 1.2.4], by expanding the work of Coppersmith and Winograd [CW87] and Strassen [Str91].

Theorem 3.4.4 (Higher-order CW method [CVZ19a]). *Let $\Phi \subseteq I_1 \times \cdots \times I_k$ be a nonempty tight set. Let $u_1, \ldots, u_k$ be injective maps such that*

$$u_1(a_1) + \cdots + u_k(a_k) = 0 \text{ for every } (a_1, \ldots, a_k) \in \Phi.$$

For any $R \subseteq \Phi \times \Phi$, let $r(R)$ be the rank over $\mathbb{Q}$ of the $|R| \times k$ matrix with rows

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

where $u(x) \coloneqq (u_1(x_1), \ldots, u_k(x_k)) \in \mathbb{Z}^k$. Then

$$\log_2 \tilde{\mathrm{Q}}_{\mathrm{IM}}(\Phi) \geq \max_{P \in \mathcal{P}(\Phi)} \left(H(P) - (k-2) \max_{R \in \mathcal{R}(\Phi)} \frac{\max_{Q \in \mathcal{D}_{R,(P_1,\ldots,P_k)}} H(Q) - H(P)}{r(R)} \right) \tag{3.7}$$

where the parameters R and Q are taken over the following domains:

- *$\mathcal{R}(\Phi)$ is the set of all subsets $R \in \Phi \times \Phi$ such that $R \nsubseteq \{(x, x) : x \in \Phi\}$ and $R \subseteq \{(x, y) \in \Phi \times \Phi : x_i = y_i\}$ for some $i \in [k]$*
- *$\mathcal{D}_{R,(P_1,\ldots,P_k)}$ is the set of probability distributions on $R \in \Phi \times \Phi$ with marginal distributions equal to $P_1, \ldots, P_k$ respectively.*

3.4.2 Tight tensors: corner, square, Lshape

We will first apply Theorem 3.4.3 to the corner problem. Before that, we see how the tightness property is satisfied by the corner problem.

Theorem 3.4.5. *For any finite Abelian group $(G, +)$, let $\Phi_{H_{\mathrm{cor},G}}$ be the support of the adjacency tensor of $H_{\mathrm{cor},G}$. Then the set $\Phi_{H_{\mathrm{cor},G}}$ is tight.*

Proof. Let $m = |G|$ and ϕ be a bijection between G and $\{0, 1, \ldots, m-1\}$. We define

$$\begin{aligned} u_1((g_1, g_2)) &= \phi(g_1) + m\phi(g_2) \\ u_2((g_1, g_2)) &= m^2\phi(g_1 + g_2) - m\phi(g_2) \\ u_3((g_1, g_2)) &= -m^2\phi(g_1 + g_2) - \phi(g_1) \ . \end{aligned}$$

It is easy to check that the maps u_1, u_2, u_3 are injective and that for every triple of pairs $(g_1, g_2), (g_1 + \lambda, g_2), (g_1, g_2 + \lambda)$, it holds that

$$u_1((g_1, g_2)) + u_2((g_1 + \lambda, g_2)) + u_3((g_1, g_2 + \lambda)) = 0 \, .$$

This proves the claim. □

As a result of Theorem 3.4.5 and Theorem 3.4.3, we have that the asymptotic induced matching number of the corner hypergraph is maximal:

Corollary 3.4.6. *For any group G, $\tilde{\mathrm{Q}}_{\mathrm{IM}}(H_{\mathrm{cor},G}) = |G|^2$.*

Proof. We know that $\Phi_{H_{\mathrm{cor},G}}$ is tight by Theorem 3.4.5, and so we may apply Theorem 3.4.3. We take $P \in \mathcal{P}(\Phi_{H_{\mathrm{cor},G}})$ to be the uniform probability distribution. It then suffices to observe that every marginal distribution P_i is also uniform to obtain the claim. □

In particular, Corollary 3.4.6 implies that no better upper bound on $\Theta(H_{\mathrm{cor},G})$ can be obtained via tools that also upper bound $\tilde{\mathrm{Q}}_{\mathrm{IM}}(H_{\mathrm{cor},G})$. Such tools include the slice rank, the analytic rank, the geometric rank and the G-stable rank. We computed the maximum independent set and maximum induced matching for $H^{\boxtimes n}_{\mathrm{cor},\mathbb{F}_2}$ for small powers $n = 1, 2, 3$ (see Table 3.2) and we found that the maximum independent set is strictly smaller than the maximum induced matching for $n = 2$ and $n = 3$. This motivates the search for methods that go beyond the maximum induced matching barrier. For comparison, we also give the analogous numbers for the cap set hypergraph H_{cap} (which is an undirected hypergraph), where, interestingly, the maximum independent set and the maximum induced matching are equal.

$H^{\boxtimes n}_{\mathrm{cap}}$		
n	independence	induced matching
1	2	2
2	4	4
3	9	9

$H^{\boxtimes n}_{\mathrm{cor},\mathbb{F}_2}$		
n	independence	induced matching
1	2	2
2	8	9
3	24	32

Table 3.2: Independence number and induced matching number for small powers of the cap set hypergraph and corner hypergraph.

Similar to the construction for showing the adjacency tensor of corner hypergraph is tight. We prove in the following lemma that the support adjacency tensors of square and Lshape hypergraphs are tight.

Lemma 3.4.7. *For any finite Abelian group $(G, +)$, let $\Phi_{\mathrm{Lshape},G}$ and $\Phi_{\mathrm{square},G}$ be the support of the adjacency tensor of $H_{\mathrm{Lshape},G}$ and $H_{\mathrm{square},G}$, respectively. Then both $\Phi_{\mathrm{Lshape},G}$ and $\Phi_{\mathrm{square},G}$ are tight.*

Proof. Let $m = |G|$ and ϕ be bijection between G and $\{0, 1, \ldots, m-1\}$. We define

$$\begin{aligned}
u_1((g_1, g_2)) &= \phi(g_1) + m\phi(g_2) \\
u_2((g_1, g_2)) &= m^2\phi(g_1) - m\phi(g_2) \\
u_3((g_1, g_2)) &= m^2\phi(g_2) - m\phi(g_1) \\
u_4((g_1, g_2)) &= -m^2\phi(g_1 + g_2) + (m-1)\phi(g_1)\,.
\end{aligned}$$

It is easy to check that the maps u_1, u_2, u_3, u_4 are injective and that for every $(g_1, g_2), (g_1 + \lambda, g_2), (g_1, g_2 + \lambda), (g_1, g_2 + 2\lambda)$, it holds that

$$u_1((g_1, g_2)) + u_2((g_1 + \lambda, g_2)) + u_3((g_1, g_2 + \lambda)) + u_4((g_1, g_2 + 2\lambda)) = 0\,.$$

This shows that $\Phi_{\text{Lshape},G}$ is tight. A similar argument with the maps defined below will prove the $\Phi_{\text{square},G}$ is tight.

$$\begin{aligned}
u_1((g_1,g_2)) &= \phi(g_1) + m\phi(g_2)\\
u_2((g_1,g_2)) &= m^2\phi(g_1) - (m-1)\phi(g_2)\\
u_3((g_1,g_2)) &= m^2\phi(g_2) - 2\phi(g_1)\\
u_4((g_1,g_2)) &= -m^2\phi(g_1+g_2) + \phi(g_1) - \phi(g_2)\,.
\end{aligned}$$

□

We have shown that the support of the adjacency tensor of the square and Lshape hypergraphs are tight in Lemma 3.4.7. Similarly with the corner problem, we can ask whether $\tilde{\mathrm{Q}}_{\mathrm{IM}}(\Phi_{\text{Lshape},G})$ or $\tilde{\mathrm{Q}}_{\mathrm{IM}}(\Phi_{\text{square},G})$ is full or not? One potential approach is to use the Theorem 3.4.4 for $k = 4$, but computing the quantity in the right hand side of the equation (3.7) is nontrivial for both $\Phi_{\text{square},G}$ and $\Phi_{\text{Lshape},G}$. Thus, we leave this question for future work.

In next subsection, we propose a simple and generic method based on fractional coverings that in principle, does not suffer from the induced matchings barrier. For the corner problem, however, it gives the trivial bound. We use it to give a simple example of a graph for which the asymptotic induced matching is arbitrarily larger than the Shannon capacity.

3.4.3 Fractional cover method

We discuss an upper bound method for the Shannon capacity based on the fractional cover method. The fractional cover method was introduced in [FK00] for finding upper bounds for the Sperner capacity of directed graphs and further improved in [KPS05]. The method in [KPS05] is easily extended to upper bound the Shannon capacity of hypergraphs. We discuss this extension here.

We say that a real-valued function, called γ, on directed k-uniform hypergraphs is *sub-multiplicative* if $\gamma(G \boxtimes H) \leq \gamma(G)\gamma(H)$ for all G and H.

Definition 3.4.8 (Upper function)**.** Let γ be a map from directed k-uniform hypergraphs to real nonnegative numbers. We say that γ is an *upper function* if γ is at least the independence number α, and γ is sub-multiplicative.

Lemma 3.4.9. *Let $H_1, \ldots, H_n$ be directed k-uniform hypergraphs. Let $H = H_1 \boxtimes H_2 \boxtimes \cdots \boxtimes H_n$. Let γ be an upper function on directed k-uniform hypergraphs. Then*

$$\alpha(H) \leq \prod_{i=1}^{n} \gamma(H_i).$$

Proof. This follows immediately from the definition of an upper function (Definition 3.4.8). □

Definition 3.4.10. (Fractional cover). Let H be a directed k-uniform hypergraph with vertex set $V(H)$. A function $g : 2^{V(H)} \to \mathbb{R}_{\geq 0}$ is called a *fractional cover* of $V(H)$ if

$$\sum_{U \in \mathcal{F} : v \in U} g(U) \geq 1 \text{ for all } v \in V(H),$$

where $\mathcal{F}$ is the family of all subsets of $V(H)$.

The methods in [KPS05] are easily extended to get the following theorem.

Theorem 3.4.11. *For any directed k-uniform hypergraph H, and any upper function γ, we have*

$$\Theta(H) \leq \min_g \sum_{U \subseteq V(H)} g(U)\gamma(H[U]),$$

where the minimization is over all fractional covers g of $V(H)$, and $H[U]$ is the directed k-uniform hypergraph induced by U.

Proof. Let h be a nonnegative integer function from $2^{V(H)}$ to $\mathbb{Z}_{\geq 0}$. For $s \in \mathbb{Z}_{\geq 0}$, h is called an s-cover of $V(H)$ if $\sum_{U: v \in U} h(U) \geq s$ hold for all $v \in V(H)$. Then we have

$$\min_g \sum_{U \subseteq V(H)} g(U)\gamma(H[U]) = \inf_s \frac{1}{s} \min_h \sum_{U \subseteq V(H)} h(U)\gamma(H[U]), \tag{3.8}$$

where the minimization on the right-hand side is taken over all s-covers h and the minimization on the left-hand side is taken over all fractional covers g. Indeed, there is a fractional cover g of $V(H)$ that takes rational values and achieves the minimum on the left-hand of (3.8). Therefore, there exists an integer number s such that $h(U) = sg(U)$ is an integral s-cover of $V(H)$. In the other direction, if h is an s-cover of $V(H)$ then the function $g(U) = h(U)/s$ is a fractional cover of $V(H)$.

Let h be an s-cover of $V(H)$ and denote $\mathcal{U} = \{U_1, \ldots, U_m\}$ the multiset of subsets of $V(H)$ with $U \subseteq V(H)$ appearing $h(U)$ times in $\mathcal{U}$. For any independent set I of H, we have

$$\sum_{i=1}^{m} \alpha(H[U_i]) \geq \sum_{i=1}^{m} \alpha(H[U_i \cap I]) = \sum_{i=1}^{m} |U_i \cap I| \geq s|I|.$$

Fix s and let h be a nonnegative s-cover attained by the minimum on the right-hand side of the equation (3.8). For $n \in \mathbb{N}$, let $\mathcal{U}^n$ be the multiset of all n-fold Cartesian products of sets from $\mathcal{U}$. For any $A = U_1 \times U_2 \times \cdots \times U_n$, define a function $h^{(n)}(A) = h(U_1) \cdot h(U_2) \cdots h(U_n)$ then $h^{(n)}$ is an s^n-cover of $V(H^{\boxtimes n})$ and the set $A = U_1 \times \cdots \times U_n$ appear in $\mathcal{U}^n$ with $h(A)$ times. Let $I^{(n)}$ be a maximum independent set in $H^{\boxtimes n}$, we have

$$\begin{aligned}
s^n|I^{(n)}| &\leq \sum_{\times_{i=1}^n U_i \in \mathcal{U}^n} \alpha\left(H^{\boxtimes n}[\times_{i=1}^n U_i]\right) \\
&= \sum_{\times_{i=1}^n U_i \in \mathcal{U}^n} \alpha\left(H[U_1] \boxtimes \cdots \boxtimes H[U_n]\right) \\
&\leq \sum_{\times_{i=1}^n U_i \in \mathcal{U}^n} \prod_{i=1}^{n} \gamma(H[U_i]) \\
&= \left[\sum_{U_i \in \mathcal{U}} \gamma(H[U_i])\right]^n.
\end{aligned}$$

Recall that $I^{(n)}$ is a maximum independent set of $H^{\boxtimes n}$, by the definition of $\mathcal{U}$, we have

$$\alpha(H^{\boxtimes n}) = |I^{(n)}| \leq \frac{1}{s^n}\left[\sum_{U\subseteq V(H)} h(U)\gamma(H[U])\right]^n.$$

This implies

$$\Theta(H) \leq \inf_s \frac{1}{s}\left[\sum_{U\subseteq V(H)} h(U)\gamma(H[U])\right] = \min_g \sum_{U\subseteq V(H)} g(U)\gamma(H[U]),$$

completing the proof. □

We give a quick example using the above method of an undirected graph G with Shannon capacity strictly smaller than the asymptotic maximum induced matching of G.

Example 3.4.12. Let G be the undirected graph with adjacency matrix

$$\begin{pmatrix} J & I \\ I & J \end{pmatrix}$$

where I is the $n \times n$ identity matrix and J is the $n \times n$ all-ones matrix with $n \geq 2$. Clearly $\Theta(G) \geq 2$, since $\{1, n+2\}$ is an independent set and $\mathrm{Q}_{\mathrm{IM}}(G) \geq n$, since $\{(i, n+i) : i \in [n]\}$ is an induced matching. Therefore, $\tilde{\mathrm{Q}}_{\mathrm{IM}}(G) \geq n$. It remains to upper bound $\Theta(G)$. For this we will use the fractional cover method.

For any graph H, define $\gamma(H)$ as the matrix rank of the adjacency matrix of H (over some arbitrary but fixed field). Then γ is an upper function, because matrix rank is multiplicative under the tensor product. Let $V_1 = \{1, \ldots, n\}$ and $V_2 = \{n+1, \ldots, 2n\}$. Setting $g(V_1) = 1$ and $g(V_2) = 1$ we have that g is a fractional cover of G. By Theorem 3.4.11, we have

$$\Theta(G) \leq g(V_1)\,\mathrm{rank}(G[V_1]) + g(V_2)\,\mathrm{rank}(G[V_2]) \leq 2,$$

because $\mathrm{rank}(G[V_1]) = \mathrm{rank}(G[V_2]) = 1$. Therefore, we have $\Theta(G) = 2$.

3.5 Conclusion

In this chapter, we introduced the combinatorial degeneration method for finding lower bounds for Shannon capacity of directed k-uniform hypergraph. We then applied this method to improve the lower bound for the corner, square, Lshape, which are special case of generalized multidimensional Szemerédi problem. Finally, we pointed out how induced matchings in hypergraphs pose a barrier for existing tensor tools (such as slice rank, subrank, analytic rank, geometric rank and G-stable rank) to efficiently obtain an upper bound on the size of independent sets in hypergraphs. This implies a barrier for these tools to effectively establish lower bounds on the communication complexity on the NOF model of the Eval function over any group G.

Chapter 4

Symmetric subrank of tensors

This chapter is based on joint work with Matthias Christandl, Omar Fawzi, and Jeroen Zuiddam [CFTZ21].

4.1 Introduction

As we have seen in Chapter 3, various important problems in combinatorics are special cases of the problem of determining the independence number of a hypergraph. The parameter of interest in those problems is the Shannon capacity of the corresponding hypergraphs. Several tensor methods have been introduced in this context to find good upper bounds for the independence number, most notably slice rank [TS16], analytic rank [Lov19], subrank [Str87, Zui18] and related parameters. We have shown, however, that all those methods suffer from a barrier that renders them powerless in the case where the independence number is low but the tensors fitting the hypergraph have large induced matchings in their support. This "induced matching barrier" calls for an effort to find methods for upper bounding the independence numbers that can go below the induced matching number. To go beyond the induced matching barrier, we introduduce a new notion of tensor rank called the *symmetric subrank*. The symmetric subrank upper bounds the independence number of a hypergraph but not the induced matching number; we propose this method as a route to circumvent the induced matching barrier.

Recall that the subrank $\mathrm{Q}(f)$ of a tensor $f \in \mathbb{F}^{d_1} \otimes \cdots \otimes \mathbb{F}^{d_k}$ measures the largest number r such that the diagonal tensor $\langle r \rangle = \sum_{i=1}^{r} e_i \otimes \cdots \otimes e_i \in \mathbb{F}^r \otimes \cdots \otimes \mathbb{F}^r$ (where the e_i form the standard basis of $\mathbb{F}^r$) can be obtained from f by acting with linear operations $A_i : \mathbb{F}^{d_i} \to \mathbb{F}^r$ on f, that is $\langle r \rangle = (A_1 \otimes \cdots \otimes A_k) f$, the *symmetric subrank* $\mathrm{Q_s}(f)$ of a tensor $f \in \mathbb{F}^d \otimes \cdots \otimes \mathbb{F}^d$ we define as the largest number r such that there is a linear map $A : \mathbb{F}^n \to \mathbb{F}^r$ so that $\langle r \rangle = (A \otimes \cdots \otimes A) f$. Generally, $\mathrm{Q_s}(f) \leq \mathrm{Q}(f) \leq \mathrm{SR}(f) \leq d$, where $\mathrm{SR}(f)$ denote the slice rank. The relation between the symmetric subrank and the subrank is analogous to the relation between the symmetric rank and the rank [CGLM08]. Note that unlike the symmetric rank which only makes sense for symmetric tensors, the symmetric subrank can be defined for any tensor. We will see that the symmetric subrank can be used to give an upper bound for the independence number but can still be strictly smaller than the subrank, of which we give a precise analysis.

4.1.1 Our results

In this section we discuss our main results on the symmetric subrank, the asymptotic symmetric subrank for symmetric tensors and the symmetric quantum functional.

Symmetric subrank

We first investigate several properties of the symmetric subrank. It is simple to see that for a hypergraph H with adjacency tensor A_H, $\mathrm{Q_s}(A_H)$ provides an upper bound on the maximum independent set of H. In general, the symmetric subrank $\mathrm{Q_s}(A_H)$ leads to a better bound compared to the subrank:

- There exists a directed graph H such that over $\mathbb{F}_2$, $\mathrm{Q_s}(A_H)$ can be smaller than the maximum induced matching (Example 4.2.3).
- There exists a directed graph H such that over $\mathbb{C}$, $\mathrm{Q_s}(A_H) < \mathrm{Q}(A_H)$ (Example 4.2.8).

However, in some settings, we can show that they are equal:

- For any undirected hypergraph H on d vertices, then over $\mathbb{C}$, $\mathrm{Q}(A_H) = d$ implies that $\mathrm{Q_s}(A_H) = d$ (Theorem 4.2.12).

Similar to Comon's question about tensor rank, which was recently answered negatively by Shitov [Shi18], we asked to find a symmetric tensor f over $\mathbb{C}$ satisfying $\mathrm{Q_s}(f) < \mathrm{Q}(f)$. Subsequently, we have been informed by Shitov [Shi22] that he can construct such an example f over $\mathbb{C}$.

Asymptotic symmetric subrank for symmetric tensors

We prove a strong asymptotic connection between the symmetric subrank and the subrank for symmetric tensors. Recall the definition of the asymptotic symmetric subrank $\tilde{\mathrm{Q}}_\mathrm{s}(f) = \lim_{n\to\infty} \mathrm{Q_s}(f^{\otimes n})^{1/n}$ and the asymptotic subrank $\tilde{\mathrm{Q}}(f) = \lim_{n\to\infty} \mathrm{Q}(f^{\otimes n})^{1/n}$. The important property of $\tilde{\mathrm{Q}}_\mathrm{s}$ is that it directly gives an upper bound on the Shannon capacity of hypergraphs, i.e., for any directed k-uniform hypergraph H, we have $\Theta(H) \leq \tilde{\mathrm{Q}}_\mathrm{s}(A_H)$, where A_H is the adjacency tensor of H (see Proposition 4.3.2). Over appropriate fields[1], for any symmetric tensor f, we can prove that [2] (see Theorem 4.3.4)

$$\tilde{\mathrm{Q}}_\mathrm{s}(f) = \tilde{\mathrm{Q}}(f).$$

In fact we prove a far stronger result that the "asymptotic restriction preorder" and the "asymptotic symmetric restriction preorder" coincide on symmetric tensors, which has a strong implication in the theory of Strassen's asymptotic spectra of tensors developed in [Str86, Str88, Str88, Str91, Tob91, Bür90]. (See also the recent works [CVZ18, Zui18, WZ21].) Specifically, our result implies that the asymptotic spectrum of symmetric tensors X_s can be obtained as a natural projection of the asymptotic spectrum of tensors. While this does not immediately tell us what X_s is (since it is not known whether X contains any

[1] algebraically closed of characteristic at least $k+1$, where k is the order of the tensor

[2] The result implies that for undirected hypergraphs (e.g. capset hypergraph) $\tilde{\mathrm{Q}}_\mathrm{s}$ cannot go below the asymptotic induced matching number.

elements besides the quantum functionals that were introduced in [CVZ18]), it does give us a lot of information about how the symmetric and non-symmetric theories interact.

Comon [CGLM08] conjectured that rank and symmetric rank coincide on symmetric tensors. Shitov [Shi18] gave a counterexample to Comon's conjecture. Our results can be interpreted as saying that Comon's conjecture is true *asymptotically* for rank, subrank and the restriction preorder. This is discussed further in Section 4.3.2.

Symmetric quantum functional

To find upper bounds on $\tilde{\mathrm{Q}}_{\mathrm{s}}$, we introduce the natural symmetric analogue of the quantum functionals of [CVZ18] using the diagonal action of the group GL_d on $(\mathbb{C}^d)^{\otimes k}$ instead of the action of the group $\mathrm{GL}_d^{\times k}$ on $(\mathbb{C}^d)^{\otimes k}$. The symmetric quantum functional F applied to a tensor $f \in (\mathbb{C}^d)^{\otimes k}$ is obtained by constructing the k-partite density operator $\rho(f) = \frac{ff^\dagger}{\|f\|^2}$ and computing the von Neumann entropy of the average of the k marginals. We refer to Section 4.5 for a precise definition. The symmetric quantum functional does give an upper bound on the asymptotic symmetric subrank, and thus also on the Shannon capacity of hypergraphs, but unfortunately it gives trivial bounds for corner hypergraph: For any tensor f, the asymptotic symmetric subrank is bounded by the symmetric quantum functional (see Theorem 4.5.7):

$$\limsup_{n\to\infty} \mathrm{Q}_{\mathrm{s}}(f^{\otimes n})^{1/n} \leq F(f) \ .$$

However, it cannot overcome the induced matching barrier as we show that for any tensor f:

$$\limsup_{n\to\infty} \mathrm{SR}(f^{\otimes n})^{1/n} \leq F(f) \ .$$

In particular, when $f = A_{H_{\mathrm{cor}},G}$ is the adjacency tensor of corner hypergraph, the symmetric quantum functional give a trivial bound. In fact, we can show that for symmetric tensors f, the asymptotic slice rank is equal to symmetric quantum functional (see Theorem 4.5.7):

$$\limsup_{n\to\infty} \mathrm{SR}(f^{\otimes n})^{1/n} = F(f) \ . \tag{4.1}$$

Equation (4.1) also gives an alternative symmetric description of the quantum functional with uniform weight $\theta = (1/k, \ldots, 1/k)$ from [CVZ18] on symmetric tensors. This description may be advantageous in the development of numerical algorithms.

Outline of the chapter. We start by introducing the symmetric subrank and discuss its basic properties in Section 4.2. Then, in Section 4.3, we analyze the asymptotic symmetric subrank in a general fashion. Next, we discuss the asymptotic spectrum of symmetric tensors in Section 4.4. Finally, in Section 4.5, we construct a symmetric quantum functional for tensors and study its properties.

4.2 Symmetric subrank

In this section we define the symmetric subrank. Then we discuss basic properties and separations, which we do in three main parts each of which focuses on a subclass of tensors: non-symmetric (i.e., general) tensors, symmetric matrices and symmetric tensors.

4.2.1 Symmetric subrank

The subrank of a k-tensor f, denoted $\mathrm{Q}(f)$, was defined in the Section 2.2 as the size of the largest diagonal tensor that can be obtained from f by acting with k linear maps $A^{(1)}, \ldots, A^{(k)}$ on the k dimensions of f. The *symmetric* subrank of a tensor is defined in the same way with the extra requirement that all linear maps $A^{(i)}$ must be the same.

Definition 4.2.1 (Symmetric restriction and symmetric subrank)**.** For any two (not necessarily symmetric) tensors $f \in (\mathbb{F}^d)^{\otimes k}$ and $g \in (\mathbb{F}^e)^{\otimes k}$, we say that f *symmetrically restricts to* g, and we write $g \leq_{\mathrm{s}} f$, if there exists a linear map $A : \mathbb{F}^d \to \mathbb{F}^e$ such that $g = A^{\otimes k} f$. Thus $g \leq_{\mathrm{s}} f$ if and only if there is an $e \times d$ matrix A such that for all $i_1, \ldots, i_k \in [e]$ we have that

$$g_{i_1,\ldots,i_k} = \sum_{j_1,\ldots,j_k \in [d]} A_{i_1,j_1} \cdots A_{i_k,j_k} f_{j_1,\ldots,j_k}.$$

We define the *symmetric subrank* of f as the largest number r such that the diagonal tensor $\langle r \rangle$ is a symmetric restriction of f, that is,

$$\mathrm{Q}_{\mathrm{s}}(f) = \max\{r \in \mathbb{N} : \langle r \rangle \leq_s f\}.$$

Our main motivation for introducing the symmetric subrank is to upper bound the independence number of k-uniform hypergraphs:

Proposition 4.2.2. *Let $H = (V, E)$ be a directed k-uniform hypergraph with n vertices with adjacency tensor A_H. Then, $\alpha(H) \leq \mathrm{Q}_{\mathrm{s}}(A_H)$, where the symmetric subrank can be understood over any field $\mathbb{F}$. In fact, for any field $\mathbb{F}$ and tensor $f \in (\mathbb{F}^n)^{\otimes k}$ with support included in the support of A_H and with diagonal entries $f_{i,\ldots,i} = 1$ for all $i \in [n]$, we have $\alpha(H) \leq \mathrm{Q}_{\mathrm{s}}(f)$.*

Proof. Let $S = \{i_1, \ldots, i_r\}$ be an independent set of H with size $r \leq n$. Then take a matrix B that has size $r \times n$ such that $B_{j,i_j} = 1$ for all $j \in [r]$, and other entries equal to 0. Then the tensor $t = (B^{\otimes k}) \cdot f$ can be written as $t_{j_1,\ldots,j_k} = f_{i_{j_1},\ldots,i_{j_k}}$. As S is an independent set, we have $(A_H)_{i_{j_1},\ldots,i_{j_k}} = 1$ if and only if $j_1 = j_2 = \cdots = j_k$. This means that $t = \langle r \rangle$. Moreover, if any hypergraph is obtained from H by deleting some edges, then its independent number is at least $\alpha(H)$, which proves the desired result. □

Let H be a directed k-uniform hypergraph with adjacency tensor A_H. Let Φ_H be the support of A_H. We have seen from Section 2.2 and Proposition 4.2.2 that the symmetric subrank $\mathrm{Q}_{\mathrm{s}}(A_H)$ and the subrank $\mathrm{Q}(A_H)$ can be used to upper bound the independence number $\alpha(H)$. However, the Lemma 3.4.1 stated that $\mathrm{Q}(A_H)$ (over any field) cannot give good bounds when $\mathrm{Q}_{\mathrm{IM}}(\Phi_H)$ (recall that $\mathrm{Q}_{\mathrm{IM}}(\Phi_H)$ is the induced matching number of H, which is introduced in Section 3.4.1) is much larger than $\alpha(H)$. We may thus think of

$\mathrm{Q}_{\mathrm{IM}}(\Phi_H)$ as a barrier for $\mathrm{Q}(A_H)$ to give good upper bounds on $\alpha(H)$. Many other tensors methods (slice rank, partition rank, analytic rank, geometric rank, G-stable rank) are also lower bounded by this barrier $\mathrm{Q}_{\mathrm{IM}}(\Phi_H)$. We will see that indeed the symmetric subrank $\mathrm{Q}_{\mathrm{s}}(A_H)$ can be strictly smaller than $\mathrm{Q}_{\mathrm{IM}}(\Phi_H)$ in the following example.

Example 4.2.3. Let C_5 be the directed cycle graph on vertices $\{1, \ldots, 5\}$ with edge set $\{(1,2),(2,3),(3,4),(4,5),(5,1)\}$. Let

$$f = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 \end{pmatrix}$$

be the adjacency matrix of C_5 over $\mathbb{F}_2$. Then $\mathrm{Q}_{\mathrm{IM}}(\Phi_{C_5})$ is the size of the largest submatrix of f that is an identity matrix up to permutation. We see that $\mathrm{Q}_{\mathrm{IM}}(\Phi_{C_5}) = 3$. On the other hand, we compute directly that $\mathrm{Q}_{\mathrm{s}}(f) = 2$ over $\mathbb{F}_2$.

We note that for symmetric tensors, a natural dual to the symmetric subrank called symmetric rank is well-studied [CGLM08]. The symmetric rank $\mathrm{R}_{\mathrm{s}}(f)$ of a symmetric tensor f is defined as the smallest number r such that $f \leq_{\mathrm{s}} \langle r \rangle$. In other words, it is the smallest number r such that there are r vectors v_i so that $f = \sum_{i=1}^{s} v_i^{\otimes k}$.

4.2.2 Symmetric subrank of matrices

In Section 4.2.1, we introduced the symmetric subrank of tensors with the motivation in mind of using symmetric subrank as a method to upper bound the independence number of hypergraphs. It is natural to ask whether this method is better than using the subrank itself. It follows directly from the definition of the symmetric subrank that for any k-tensor f we have that $\mathrm{Q}_{\mathrm{s}}(f) \leq \mathrm{Q}(f)$. Can this inequality be strict? Over the finite field $\mathbb{F}_2$, we have seen in Example 4.2.3 that the inequality can be strict. In this and the following sections we will discuss relations and separations with the ordinary subrank. We obtain precise results under assumptions about the order, ground field and symmetry of the tensors.

In this section we consider tensors of order two. These we can simply think of as matrices via the identification $\sum_{i,j} f_{ij}\, e_i \otimes e_j \mapsto (f_{ij})_{ij}$. In the language of matrices the restriction order and symmetric restriction order are given as follows. For matrices $f \in \mathbb{F}^{n_1 \times n_2}$ and $g \in \mathbb{F}^{m_1 \times m_2}$ we have $f \leq g$ if there are matrices $A^{(i)} \in \mathbb{F}^{n_i \times m_i}$ such that $f = A^{(1)} g (A^{(2)})^T$. For matrices $f \in \mathbb{F}^{n \times n}$ and $g \in \mathbb{F}^{m \times m}$ we have $f \leq_{\mathrm{s}} g$ if there is a matrix $A \in \mathbb{F}^{n \times m}$ such that $f = AgA^T$. Note in particular how in this formulation we multiply on the left by A and on the right by the transpose of A. When A is invertible and $f = AgA^T$ the matrices f and g are often called *congruent*. However we will allow A to be non-invertible. The (symmetric) subrank of a matrix f is now the largest number r such that the $r \times r$ diagonal matrix $\langle r \rangle$ is a (symmetric) restriction of f.

First of all, as a basic fact that we will use later, we note that for any matrix f the subrank $\mathrm{Q}(f)$ equals the usual notion of matrix rank $\mathrm{rank}(f)$.

Lemma 4.2.4. *Let f be a matrix, then* $\mathrm{Q}(f) = \mathrm{rank}(f)$.

Proof. Clearly $\mathrm{Q}(f) \leq \mathrm{rank}(f)$. It is well-known that by Gaussian elimination we can find invertible matrices $A^{(i)}$ such that f is a diagonal matrix with $\mathrm{rank}(f)$ nonzero entries. Thus $\mathrm{Q}(f) \geq \mathrm{rank}(f)$. □

Lemma 4.2.5. *Let f be a $d \times d$ matrix over an arbitrary field $\mathbb{F}$ such that $f_{\ell,\ell} = 0$ for all $\ell \in [d]$ and $f_{i,j} = -f_{j,i}$ for all $i \neq j \in [d]$. Then $\mathrm{Q_s}(f) = 0$.*

Proof. For any matrix $B \in \mathbb{F}^{m \times d}$ let $g = BfB^T$. Then the diagonal entries g_{kk} are zero for all k. Indeed we have $g_{kk} = \sum_{i,j} B_{ki} B_{kj} f_{ij} = 0$ since $f_{ij} = -f_{ji}$ for all $i \neq j$. We conclude that $\mathrm{Q_s}(f) = 0$. □

In particular, if $\mathbb{F} \neq \mathbb{F}_2$, then the condition in Lemma 4.2.5 is equivalent to $f = -f^T$, that is, f being skew-symmetric.

Example 4.2.6. It is easy to find a $d \times d$ matrix f of full rank that satisfies the condition in Lemma 4.2.5. Then by Lemma 4.2.5 we have $\mathrm{Q_s}(f) = 0$ while $\mathrm{Q}(f) = d$. For example, for even d we may take f with entries $f_{i,d+1-i} = 1$ for all $1 \leq i \leq d/2$ and $f_{i,d+1-i} = -1$ for all $d/2 < i \leq d$ and all other entries equal to zero, that is,

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \ldots$$

Lemma 4.2.7. *Let f be a non-symmetric $d \times d$ matrix over an arbitrary field $\mathbb{F}$. Then $\mathrm{Q_s}(f) < d$.*

Proof. Suppose that $\mathrm{Q_s}(f) = d$. Then there is a matrix A such that $\langle d \rangle = AfA^T$. Since $\langle d \rangle$ has full rank, A must have full rank. We find that $f = A^{-1}\langle d \rangle (A^T)^{-1} = A^{-1}(A^{-1})^T$ and so f is symmetric. This is a contradiction. □

Example 4.2.8. Let C_{2k+1} be the directed cycle graph with vertex set $\{1, \ldots, 2k+1\}$ and edge set $\{(1,2), (2,3), \ldots, (2k+1, 1)\}$ and let f be the adjacency matrix of C_{2k+1} over any fixed field, so that f is the $(2k+1) \times (2k+1)$ matrix

$$f = \begin{pmatrix} 1 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 & 1 \\ 1 & 0 & 0 & 0 & \cdots & 0 & 1 \end{pmatrix}$$

We have $\mathrm{Q}(f) = \mathrm{rank}(f) = 2k+1$ by Lemma 4.2.4. On the other hand, $\mathrm{Q_s}(f) < 2k+1$ by Lemma 4.2.7.

4.2.3 Symmetric subrank of symmetric matrices

We have seen in the previous section that the symmetric subrank can be strictly smaller than the subrank for non-symmetric matrices. For *symmetric* matrices, we now prove that symmetric subrank and subrank are equal as long as the ground field is *quadratically*

closed[3], meaning that every element has a square root. Algebraically closed fields are in particular quadratically closed.

Theorem 4.2.9. *For any symmetric matrix f over a quadratically closed field $\mathbb{F} \neq \mathbb{F}_2$, $\mathrm{Q}(f) = \mathrm{Q_s}(f)$.*

It follows from Example 4.2.6 that the statement of Theorem 4.2.9 indeed fails over $\mathbb{F}_2$ if we let f be a full-rank anti-diagonal matrix.

The proof of Theorem 4.2.9 relies on the following theorem.

Theorem 4.2.10 (Ballantine [Bal68])**.** *Let $\mathbb{F}$ be a field with size at least 3 and f be a square matrix of size d over $\mathbb{F}$ that is not a nonzero skew-symmetric matrix. There is an invertible matrix B of size d such that BfB^T is a lower triangular matrix that has exactly $\mathrm{rank}(f)$ nonzero elements on its diagonal.*

Proof of Theorem 4.2.9. The symmetric matrix f is in particular not a nonzero skew-symmetric matrix, so we may apply Theorem 4.2.10 to find an invertible matrix B such that BfB^T is lower triangular with exactly $\mathrm{rank}(f)$ nonzero elements on its diagonal. Since f is symmetric, BfB^T is also symmetric. It follows that BfB^T is a diagonal matrix. Since the ground field is quadratically closed, there is a diagonal matrix C such that $CBfB^TC^T$ is a diagonal matrix with only zeroes and ones on the diagonal. Then clearly $\mathrm{Q_s}(f) \geq \mathrm{rank}(f) = \mathrm{Q}(f)$, which proves the claim. □

4.2.4 Symmetric subrank of symmetric tensors

In Section 4.2.3 we proved that the symmetric subrank and subrank coincide on symmetric matrices over the complex numbers. Extending the notion of a symmetric matrix, a tensor $f \in (\mathbb{F}^d)^{\otimes k}$ is called *symmetric* if for all $(i_1, \ldots, i_k) \in [d]^k$ and all permutations σ of $[k]$ we have $f_{i_1,\ldots,i_k} = f_{i_{\sigma(1)},\ldots,i_{\sigma(k)}}$. For symmetric k-tensors f of order $k \geq 3$, we find examples of strict inequality $\mathrm{Q_s}(f) < \mathrm{Q}(f)$ over finite fields.

Example 4.2.11. Let $f = e_1 \otimes e_2 \otimes e_3 + e_1 \otimes e_3 \otimes e_2 + e_2 \otimes e_1 \otimes e_3 + e_2 \otimes e_3 \otimes e_1 + e_3 \otimes e_1 \otimes e_2 + e_3 \otimes e_2 \otimes e_1 + e_1 \otimes e_1 \otimes e_1$, where $e_1, e_2, e_3 \in \mathbb{F}_2^3$ is the standard basis of $\mathbb{F}_2^3$. It is not hard to verify that $\mathrm{Q_s}(f) = 1$ while $\mathrm{Q}(f) = 2$.

Over the complex filed $\mathbb{C}$. Similar to Comon's question about tensor rank, which was recently answered negatively by Shitov [Shi18], we asked to find a symmetric tensor f satisfying $\mathrm{Q_s}(f) < \mathrm{Q}(f)$. Subsequently, we have been informated by Shitov [Shi22] that he can construct such an example f over $\mathbb{C}$.

Next, we prove that there is a general case where $\mathrm{Q_s}(f) = \mathrm{Q}(f)$. Namely, for symmetric complex tensors, if the subrank is maximal, then also the symmetric subrank is maximal, in the following sense:

Theorem 4.2.12. *Let $f \in (\mathbb{C}^d)^{\otimes k}$ be a symmetric tensor. If $\mathrm{Q}(f) = d$ then $\mathrm{Q_s}(f) = d$.*

[3] One could consider an alternative definition of symmetric subrank in which the symmetric restriction order is replaced by the following: let $f \leq_{\mathrm{s}}' g$ if and only if there is a matrix A and diagonal matrices $D_1, \ldots, D_k, E_1, \ldots, E_k$ such that $f = (E_1 \otimes \cdots \otimes E_k)(A \otimes \cdots \otimes A)(D_1 \otimes \cdots \otimes D_k)g$. Under this alternative symmetric restriction preorder, the subrank and (alternative) symmetric subrank become equal for all symmetric matrices over any field. [CGLM08] take a similar approach to this when dealing with the symmetric rank over the reals (which is not quadratically closed).

To prove Theorem 4.2.12 we use a simple corollary of the following theorem.

Theorem 4.2.13 (Belitskii and Sergeichuk [BS06])**.** *Let $f, f' \in (\mathbb{C}^d)^{\otimes k}$ be tensors of order k. If $A^1, \dots, A^k$ are invertible matrices of size d such that $f' = (A^{\pi(1)} \otimes \cdots \otimes A^{\pi(k)})f$ for all permutations $\pi \in \mathfrak{S}_k$. Then there is an invertible matrix B of size d such that $f' = (B \otimes \cdots \otimes B)f$.*

Corollary 4.2.14 (Corollary of Theorem 4.2.13)**.** *Let $f', f \in (\mathbb{C}^d)^{\otimes k}$ be symmetric tensors. If there are k invertible matrices $A^1, \dots, A^k$ of size d such that $f' = (A^1 \otimes \cdots \otimes A^k)f$. Then there is an invertible matrix B of size d such that $f' = (B \otimes \cdots \otimes B)f$.*

Proof of Corollary 4.2.14. For any permutation $\pi \in \mathfrak{S}_k$. We have

$$\begin{aligned}\sum_{j_1 \in [d], \dots, j_k \in [d]} A^{\pi(1)}_{i_1, j_1} \dots A^{\pi(k)}_{i_k, j_k} f_{j_1, \dots, j_k} &= \sum_{j_1 \in [d], \dots, j_k \in [d]} A^1_{i_{\pi^{-1}(1)}, j_1} \dots A^k_{i_{\pi^{-1}(k)}, j_k} f_{j_1, \dots, j_k} \\ &= f'_{i_{\pi^{-1}(1)}, \dots, i_{\pi^{-1}(k)}} = f'_{i_1, \dots, i_k}.\end{aligned}$$

Therefore $f' = (A^{\pi(1)} \otimes \cdots \otimes A^{\pi(k)})f$ for all $\pi \in \mathfrak{S}_k$. By using Theorem 4.2.13, the proof is completed. □

Proof of Theorem 4.2.12. Since $\mathrm{Q}(f) = d$, there are k matrices $A^1, \dots, A^k$ of size $d \times d$ such that $\langle d \rangle = (A^{(1)} \otimes \cdots \otimes A^{(k)})f$. Suppose that there is a matrix $A^{(i)}$ which is not invertible, then the rank of i-th flattening matrix of $(A^{(1)} \otimes \cdots \otimes A^{(k)})f$ is smaller than $d - 1$, that is, $\mathrm{rank}(\mathrm{flatten}_i((A^{(1)} \otimes \cdots \otimes A^{(k)})f) \leq d - 1$. But the rank of all flattenings of $\langle d \rangle$ are equal to d. Therefore all $A^{(1)}, \dots, A^{(k)}$ are invertible matrices. By the above corollary, there is an invertible matrix B such that $\langle d \rangle = (B \otimes \cdots \otimes B)f$, this implies $\mathrm{Q_s}(f) = d$. □

Interpretation in terms of homogeneous polynomials

There is a natural identification between symmetric tensors on the one hand and homogeneous polynomials on the other hand. A homogeneous polynomial is a polynomial whose monomials all have the same total degree k. Any symmetric k-tensor $f \in (\mathbb{F}^d)^{\otimes k}$ corresponds uniquely to a homogeneous polynomial of degree k in d variables $F \in \mathbb{F}[x_1, \dots, x_d]_k$ via the expression:

$$F(x_1, \dots, x_d) = \sum_{j_1, \dots, j_k \in [d]} f_{j_1, \dots, j_k} \cdot x_{j_1} \cdot x_{j_2} \cdots x_{j_k}.$$

We define the *symmetric subrank* of F, written $\mathrm{Q_s}(F)$, as the largest number $r \in \mathbb{N}$ such that there are d linear forms $\ell_1(y_1, \dots, y_r), \dots, \ell_d(y_1, \dots, y_r)$ in r variables $y_1, \dots, y_r$, such that

$$F\left(\ell_1(y_1, \dots, y_r), \dots, \ell_d(y_1, \dots, y_r)\right) = \sum_{i=1}^{r} y_i^k.$$

To phrase it differently, the symmetric subrank $\mathrm{Q_s}(F)$ is the largest $r \in \mathbb{N}$ such that there is a matrix $A = (a_{ij})_{i,j} \in \mathbb{F}^{d \times r}$ such that $F(A \cdot Y) = \sum_{i=1}^{r} y_i^k$, where $Y = (y_1, \dots, y_r)$

and $A \cdot Y = (a_{11}y_1 + \cdots + a_{1r}y_r, \ldots, a_{d1}y_1 + \cdots + a_{dr}y_r)$. The symmetric subrank for homogeneous polynomials and for symmetric tensors coincide via the above identification, in the sense that $\mathrm{Q_s}(F) = \mathrm{Q_s}(f)$.

In a similar way, the symmetric rank of a symmetric tensor has a natural interpretation in terms of the associated homogeneous polynomial [IK99]. This notion is also called the Waring rank.

Also the notion of symmetric restriction of symmetric tensors carries over to homogeneous polynomials, as follows. Let $F \in \mathbb{F}[x_1, \ldots, x_d]_k$ and $G \in \mathbb{F}[y_1, \ldots, y_{d'}]_k$ be homogeneous polynomials of degree k in d and d' variables, respectively. We say that F is a symmetric restriction of G, and write $F \leq_s G$, if there is a matrix $A \in \mathbb{F}^{d' \times d}$ such that $F = G(A \cdot X)$, where, as before, $X = (x_1, \ldots, x_d)$ and $A \cdot X$ is defined as before. The symmetric restriction of symmetric tensors and for homogeneous polynomials coincide, in the sense that $F \leq_s G$ if and only if $f \leq_s g$, where f and g are the symmetric tensors associated to F and G, respectively.

4.3 Asymptotic symmetric subrank

In Section 4.2, we introduced the symmetric subrank guided by the motivation of using this tensor parameter to upper bound the independence number of hypergraphs. In many hypergraph independence problems (e.g. the generalized multidimensional Szemerédi problem and its special cases: cap set, corner, etc.), the hypergraph under consideration has a power structure (under the strong product $\boxtimes$, which is simply the tensor product on the adjacency tensor). In other words, the parameter of interest in those problems is their Shannon capacity of hypergraph that corresponds to these problems, which is introduced in Section 3.2.

In this asymptotic context, and with upper bounding the Shannon capacity in mind, we introduce and study the asymptotic symmetric subrank. We define the *asymptotic symmetric subrank* of a tensor $f \in (\mathbb{F}^d)^{\otimes k}$ as

$$\tilde{\mathrm{Q}}_{\mathrm{s}}(f) := \limsup_{n \to \infty} \mathrm{Q_s}(f^{\otimes n})^{1/n}.$$

(The fact that we are using the lim sup rather than lim or sup is a technicality which in most relevant cases simplifies as we discuss below.) For any tensor $f \in (\mathbb{F}^d)^{\otimes k}$, since we have the basic inqualities $\mathrm{Q_s}(f) \leq \mathrm{Q}(f) \leq d$, we also have that $\tilde{\mathrm{Q}}_{\mathrm{s}}(f) \leq \tilde{\mathrm{Q}}(f) \leq d$.

Note that, because of the earlier Example 4.2.6, this lim sup cannot generally be replaced by a limit.[4] However, we will be interested in the adjacency tensors of hypergraphs which have the special property that the coefficients on the main diagonal are all one. In that case we can replace the lim sup by a limit or supremum as follows:

Proposition 4.3.1. *Let $f \in (\mathbb{F}^d)^{\otimes k}$ be a tensor such that there is an $i \in [d]$ with $f_{i,\ldots,i} = 1$. Then $\tilde{\mathrm{Q}}_{\mathrm{s}}(f) = \sup_n \mathrm{Q_s}(f^{\otimes n})^{1/n} = \lim_{n\to\infty} \mathrm{Q_s}(f^{\otimes n})^{1/n}$.*

[4] For the usual subrank, the *asymptotic subrank* of the tensor $f \in (\mathbb{F}^d)^{\otimes k}$ was defined by Strassen as the limit $\tilde{\mathrm{Q}}(f) = \lim_{n\to\infty} \mathrm{Q}(f^{\otimes n})^{1/n}$, which, since Q is super-multiplicative and $\mathrm{Q}(f) \geq 1$ if $f \neq 0$, equals the supremum $\sup_n \mathrm{Q}(f^{\otimes n})^{1/n}$ (Fekete's lemma). For the *symmetric* subrank, we have to be more careful about how we define the asymptotic symmetric subrank. For example, in Example 4.2.6 we gave a matrix f for which $f^{\otimes n}$ is symmetric if n is even and skew-symmetric if n is odd, and so $\mathrm{Q_s}(f^{\otimes n}) = 2^n$ if n is even, and $\mathrm{Q_s}(f^{\otimes n}) = 0$ when n is odd. Thus, the limit $\lim_{n\to\infty} \mathrm{Q_s}(f^{\otimes n})^{1/n}$ might not exist.

Proof. Let $B \in \mathbb{R}^{1\times d}$ be the $1 \times d$ matrix with $B_{1,i} = 1$ and the other entries equal to 0. Then $\langle 1\rangle = (B \otimes \cdots \otimes B)f$. Therefore $\mathrm{Q_s}(f) \geq 1$. The symmetric subrank is super-multiplicative under tensor product. Thus, by Fekete's lemma, we find the required statement that $\tilde{\mathrm{Q}}_\mathrm{s}(f) = \sup_n \mathrm{Q_s}(f^{\otimes n})^{1/n} = \lim_{n\to\infty} \mathrm{Q_s}(f^{\otimes n})^{1/n}$. □

The important property of $\tilde{\mathrm{Q}}_\mathrm{s}$ is that it directly gives an upper bound on the Shannon capacity of hypergraphs.

Proposition 4.3.2. *Let $H = (V, E)$ be directed k-uniform hypergraph on n vertices. Let $\mathbb{F}$ be any field. Let $f \in (\mathbb{F}^n)^{\otimes k}$ be a tensor such that, for every $e \in [n]^k$ if $e \notin E$, then $f_{e_1,\ldots,e_k} = 0$, and for every $i \in [n]$, $f_{i,\ldots,i} = 1$. Then $\Theta(H) \leq \tilde{\mathrm{Q}}_\mathrm{s}(f)$.*

Proof. By the definition of f, we have that $f^{\otimes n}$ satisfies the condition of Proposition 4.2.2 for the hypergraph $H^{\boxtimes n}$. Therefore $\Theta(H) = \sup_n(\alpha(H^{\boxtimes n}))^{1/n} \leq \sup_n(\mathrm{Q_s}(A_H^{\otimes n}))^{1/n} = \tilde{\mathrm{Q}}_\mathrm{s}(A_H)$. □

4.3.1 Asymptotic symmetric subrank of matrices

We conjecture that the asymptotic symmetric subrank of a k-tensor with $k \geq 3$ can be strictly smaller than the asymptotic subrank. This cannot happen for $k = 2$. In that case we prove that there is no strict inequality, again using Theorem 4.2.10.

Theorem 4.3.3. *For any matrix f over a quadratically closed field $\mathbb{F} \neq \mathbb{F}_2$ we have $\tilde{\mathrm{Q}}(f) = \tilde{\mathrm{Q}}_\mathrm{s}(f)$.*

Proof. We will use Theorem 4.2.10. We may assume that f is a $d \times d$ matrix. Let $r = \mathrm{rank}(f)$. Then $\tilde{\mathrm{Q}}(f) = \mathrm{rank}(f) = r$. Suppose that f is a skew-symmetric matrix. Then we have $\mathrm{Q_s}(f) = 0$ by Lemma 4.2.5. The matrix $f^{\otimes n}$ is symmetric if n is even and skew-symmetric if n is odd. Then by Theorem 4.2.9 we have

$$\mathrm{Q_s}(f^{\otimes n}) = \begin{cases} r^n \text{ if } n \text{ is even,} \\ 0 \text{ otherwise.} \end{cases}$$

Therefore $\tilde{\mathrm{Q}}_\mathrm{s}(f) = r$. Suppose that f is not skew-symmetric. By Theorem 4.2.10, there is an invertible matrix B and a lower-triangular matrix L such that $BfB^T = L$. Then $\mathrm{Q_s}(f) = \mathrm{Q_s}(L)$ and so $\tilde{\mathrm{Q}}_\mathrm{s}(f) = \tilde{\mathrm{Q}}_\mathrm{s}(L)$. There is a principal submatrix A of L of size r that has exactly r nonzero elements on its diagonal. Then $A^{\otimes n}$ is a submatrix of $L^{\otimes n}$. We choose $n = rk$ for some $k \in \mathbb{N}_{\geq 1}$. Then the submatrix of $A^{\otimes n}$ with rows and columns indexed by the elements in $[r]^n$ of type $(n/r, \ldots, n/r)$ is diagonal and has size

$$\binom{n}{n/r, \ldots, n/r} \geq r^{n-o(n)}.$$

We conclude that $\tilde{\mathrm{Q}}_\mathrm{s}(L) \geq r$. □

It follows from Example 4.2.6 that the statement of Theorem 4.3.3 is false over $\mathbb{F}_2$ by taking f to be an anti-diagonal matrix with ones on the antidiagonal.

4.3.2 Asymptotic symmetric subrank of symmetric tensors

For symmetric tensors we prove that the asymptotic symmetric subrank is equal the asymptotic subrank (as long as the field satisfies mild closedness and characteristic conditions):

Theorem 4.3.4. *Let f be a symmetric k-tensor over an algebraically closed field of characteristic at least $k+1$. Then $\tilde{\mathrm{Q}}(f) = \tilde{\mathrm{Q}}_{\mathrm{s}}(f)$.*

In particular, Theorem 4.3.4 holds for any symmetric tensor over the field of complex numbers.

In fact we prove a much more general asymptotic statement about the restriction preorder $\leq$ and the symmetric restriction preorder on symmetric tensors. We define the asymptotic restiction preorder $\lesssim$ on tensors f, g by writing $f \lesssim g$ if and only if $f^{\otimes n} \leq g^{\otimes n+o(n)}$. Similarly we define the asymptotic symmetric restriction preorder $\lesssim_{\mathrm{s}}$ on tensors f, g by writing $f \lesssim_{\mathrm{s}} g$ if and only if $f^{\otimes n} \leq_{\mathrm{s}} g^{\otimes n+o(n)}$.

Theorem 4.3.5. *For symmetric k-tensors f, g over an algebraically closed field of characteristic at least $k+1$ we have $f \lesssim g$ if and only if $f \lesssim_{\mathrm{s}} g$.*

It will also follow from our proof that on symmetric tensors (over an appropriate field) the asymptotic rank and symmetric asymptotic rank are equal:

Theorem 4.3.6. *Let f be a symmetric k-tensor over an algebraically closed field of characteristic at least $k+1$. Then $\mathrm{R}_{\mathrm{s}}(f) \leq 2^{k-1}\,\mathrm{R}(f)$ and in particular $\tilde{\mathrm{R}}(f) = \tilde{\mathrm{R}}_{\mathrm{s}}(f)$.*

For $k = 3$ the same relation between symmetric rank and rank for symmetric tensors was found in [Kay12].

The above three theorems are related to Comon's conjecture [CGLM08], which says that rank and symmetric rank coincide on symmetric tensors. Shitov [Shi18] gave a counterexample to Comon's conjecture. Our Theorem 4.3.4, Theorem 4.3.5 and Theorem 4.3.6 can be interpreted as saying that "Comon's conjecture" is true *asymptotically*, not only for rank (Theorem 4.3.6), but also for subrank (Theorem 4.3.4) and the restriction preorder (Theorem 4.3.5).

The proofs for all of the above will follow from three basic lemmas that we will discuss now. A crucial role will be played by the following k-tensor.

Definition 4.3.7 (fully symmetric k-tensor)**.** For any $k \in \mathbb{N}$ let $\mathfrak{S}_k$ be the symmetric group on k elements and define the k-tensor $h = \sum_{\pi \in S_k} e_{\pi(1)} \otimes \cdots \otimes e_{\pi(k)}$. We will call h the *fully symmetric k-tensor*.

For example, for $k = 3$, the tensor h is given by $h = e_1 \otimes e_2 \otimes e_3 + e_1 \otimes e_3 \otimes e_2 + e_2 \otimes e_1 \otimes e_3 + e_2 \otimes e_3 \otimes e_1 + e_3 \otimes e_1 \otimes e_2 + e_3 \otimes e_2 \otimes e_1$. The tensor h allows us to transform any restriction to a symmetric restriction:

Lemma 4.3.8. *Let f and g be symmetric k-tensors over a field of characteristic at least $k+1$. If $f \geq g$, then $f \otimes h \geq_{\mathrm{s}} g \otimes h$, and hence also $f \otimes h \geq_{\mathrm{s}} g$, where h is the fully symmetric tensor.*

Proof. Let $A_1, \ldots, A_k$ be linear maps such that $(A_1 \otimes \cdots \otimes A_k)f = g$. Let e_i^* denote the elements of the basis dual to the standard basis e_i. Define the linear map $B = \sum_i A_i \otimes e_i e_i^*$. Then

$$(B^{\otimes k})(f \otimes h) = k!((A_1 \otimes \cdots \otimes A_k)f) \otimes h.$$

Dividing by $k!$ proves the claim. □

In particular, Lemma 4.3.8 says that, if $f^{\otimes n} \geq \langle r \rangle$, then $f^{\otimes n} \otimes h \geq_{\mathrm{s}} \langle r \rangle$ for every $n \in \mathbb{N}$. Note that h is a fixed tensor that is independent of n. Our next goal is to prove that for every f there is a constant $c \in \mathbb{N}$ depending on f such that $f^{\otimes c} \geq_{\mathrm{s}} h$. This is true in the following sense.

Recall that for any subset $S \subseteq [k]$ that is not empty and not $[k]$, any k-tensor $f \in V_1 \otimes \cdots \otimes V_k$ can be flattened into a 2-tensor $(\bigotimes_{i \in S} V_i) \otimes (\bigotimes_{i \in [k] \setminus S} V_i)$. For a k-tensor f we call the ranks of these flattenings the flattening ranks of f.

Lemma 4.3.9. *Let f be a symmetric k-tensor over an algebraically closed field. Suppose that some flattening rank of f is at least 2. Then there is a $c \in \mathbb{N}$ such that $f^{\otimes c} \geq_{\mathrm{s}} h$.*

To prepare for the proof of Lemma 4.3.9 we prove the following lemma.

Lemma 4.3.10. *Let f be a symmetric k-tensor over an algebraically closed field. There exists a basis transformation $A \in \mathbb{F}^{d \times d}$ such that the support $S = \mathrm{supp}(A^{\otimes k} f) \subseteq [d]^k$ of f after applying the transformation A satisfies $(i, \ldots, i) \notin S$ for every $1 \leq i \leq d-1$.*

Proof. Suppose that $f \in (\mathbb{F}^d)^{\otimes k}$. If no element of the form $(i, \ldots, i)$ appears in S, then we are done. Otherwise, we may assume that $(d, \ldots, d)$ appears, so that the tensor f is of the form $f = f_1 e_1^{\otimes k} + f_2 e_2^{\otimes k} + \cdots + f_d e_d^{\otimes k} +$ other terms, for some coefficients f_i with $f_d \neq 0$.

We apply to f the invertible linear map that maps e_i to e_i for $1 \leq i \leq d-1$ and maps e_d to $e_d + \varepsilon_1 e_1 + \cdots + \varepsilon_{d-1} e_{d-1}$ for some $\varepsilon_i \in \mathbb{F}$. This gives a tensor $g \in (\mathbb{F}^d)^{\otimes k}$ that is isomorphic to f and of the form $g = (f_1 + \varepsilon_1^k f_d) e_1^{\otimes k} + \cdots + (f_{d-1} + \varepsilon_{d-1}^k f_d) e_{d-1}^{\otimes k} +$ other terms. Since f_d is nonzero and the ground field is algebraically closed, there are values for the ε_i such that $f_i + \varepsilon_i^k f_i$ is zero for every $1 \leq i \leq d-1$, in which case $(i, \ldots, i)$ does not appear in the support of g for every $1 \leq i \leq d-1$. □

Proof of Lemma 4.3.9. Let $f \in (\mathbb{F}^d)^{\otimes k}$. By Lemma 4.3.10 we may assume that $(i, \ldots, i)$ does not appear in the support $S = \mathrm{supp}(f) \subseteq [d]^k$ of f for $1 \leq i \leq d-1$. For every element $s \in S$ we define its type $(y_1, \ldots, y_d)$ by letting y_i be the number of times that i appears in s. Let Y be the set of types of elements of S. Since some flattening rank of f is at least 2, we cannot have that $S = \{(d, \ldots, d)\}$. Thus without loss of generality there is a type $y \in Y$ that satisfies $1 \leq y_1 \leq k-1$ and such that for every other type $y' \in Y$ it holds that $y'_1 \leq y_1$ (maximality assumption).

Let $R \subseteq [d]^k$ be the set of all k-tuples in $[d]^k$ of type y. Let A be the $|R| \times k$ matrix with rows given by the elements of R, in some arbitrary order. Let C be the set of columns of A. Note that in any $s \in S$ the element 1 can appear at most y_1 times by our maximality assumption.

We claim that $f^{\otimes |R|}$ restricts symmetrically to the fully symmetric k-tensor h by zeroing out all basis elements that are not in C. To prove this we need to show that for any choice

of k elements $v_1, \dots, v_k$ in C, if for every i we have that $((v_1)_i, \dots, (v_k)_i) \in S$, then $v_1, \dots, v_k$ are all different.

By construction of C, for any y_1 distinct elements $v_1, \dots, v_{y_1}$ of C there is an $1 \leq i \leq |R|$ such that $(v_1)_i = \cdots = (v_{y_1})_i = 1$. Thus also for any y_1 (not necessarily distinct) elements $v_1, \dots, v_{y_1}$ of C there is an $1 \leq i \leq |R|$ such that $(v_1)_i = \cdots = (v_{y_1})_i = 1$.

Let $v_1, \dots, v_k$ be an arbitrary collection of elements of C. Suppose that $v_1 = v_2$. By the previous argument we know that there is an $1 \leq i \leq |R|$ such that $(v_2)_i = \cdots = (v_{y_1+1})_i = 1$. From the assumption $v_1 = v_2$ it follows that $(v_1)_i = (v_2)_i = \cdots = (v_{y_1+1})_i = 1$. However, we picked the type $(y_1, \dots, y_d)$ such that y_1 is maximal and $y_1 \leq k-1$. The element 1 appears at least $y_1 + 1$ times in $((v_1)_i, \dots, (v_k)_i)$. Therefore $((v_1)_i, \dots, (v_k)_i)$ is not in S. □

Proof of Theorem 4.3.5. Suppose that $f \gtrsim g$. This means that $f^{\otimes m+o(m)} \geq g^{\otimes m}$. We know from Lemma 4.3.9 that there is a constant $c \in \mathbb{N}$, depending only on f, such that $f^{\otimes c} \geq_{\mathrm{s}} h$. By Lemma 4.3.8 we then have

$$f^{\otimes m+o(m)} \otimes f^{\otimes c} \geq_{\mathrm{s}} f^{\otimes m+o(m)} \otimes h \geq_{\mathrm{s}} g^{\otimes m}.$$

This means $f \gtrsim_{\mathrm{s}} g$, which proves the claim. □

Although essentially Theorem 4.3.4 and Theorem 4.3.6 can be proven abstractly from Theorem 4.3.5, we will give the (simple) proofs separately in terms of the above lemmas for the convenience of the reader and to get the precise statement of Theorem 4.3.6:

Proof of Theorem 4.3.4. Suppose that $\mathrm{Q}(f^{\otimes n}) \geq r$. Then $f^{\otimes n} \geq \langle r \rangle$. By Lemma 4.3.9 there is a constant $c \in \mathbb{N}$, depending only on f, such that $f^{\otimes c} \geq_{\mathrm{s}} h$. By Lemma 4.3.8 we then have

$$f^{\otimes n+c} \geq_{\mathrm{s}} f^{\otimes n} \otimes f^{\otimes c} \geq_{\mathrm{s}} f^{\otimes n} \otimes h \geq_{\mathrm{s}} \langle r \rangle.$$

Thus $\mathrm{Q}_{\mathrm{s}}(f^{\otimes n+c}) \geq r$, which implies the claim. □

Proof of Theorem 4.3.6. Suppose that $\mathrm{R}(f) \leq r$. Then $f \leq \langle r \rangle$. Let $s = \mathrm{R}_{\mathrm{s}}(h)$ be the symmetric rank of the fully symmetric tensor h and note that s is a constant depending only on k, the order of f. In fact, $s \leq 2^{k-1}$, which follows from the known identity

$$h = \frac{1}{2^{k-1}} \sum_{\varepsilon_i = \pm 1} \Big(\prod_{i=2}^{k} \varepsilon_i \Big) (e_1 + \varepsilon_2 e_2 + \varepsilon_3 e_3 + \cdots + \varepsilon_k e_k)^{\otimes k}$$

in which the sum goes over $\varepsilon_2, \dots, \varepsilon_k = \pm 1$. We refer to [GW09, Lemma B.2.3] for a proof of this identity. See also [LT10, Proposition 11.6]. Then

$$\langle rs \rangle = \langle r \rangle \otimes \langle s \rangle \geq_{\mathrm{s}} \langle r \rangle \otimes h \geq_{\mathrm{s}} f.$$

Thus $\mathrm{R}_{\mathrm{s}}(f) \leq rs$, which implies the first claim. Then, since s is constant, it follows that $\mathrm{R}(f^{\otimes n}) \leq \mathrm{R}_{\mathrm{s}}(f^{\otimes n}) \leq \mathrm{R}(f^{\otimes n}) s$ for every $n \in \mathbb{N}$, which implies the second claim. □

4.4 Asymptotic spectrum of symmetric tensors

In Section 4.2 we introduced the symmetric subrank and in Section 4.3 we introduced the asymptotic symmetric subrank, both motivated by the problem of upper bounding the independence number of hypergraphs (with the asymptotic symmetric subrank in particular being relevant for capacity-type questions, where the hypergraphs at hand have a power structure). We proved several equalities and separations for these parameters.

In this section we continue our analysis of the asymptotic symmetric subrank in a general fashion that also allows us to discuss the asymptotic symmetric rank and the asymptotic symmetric restriction preorder (which we will define).

At the core of this section is the duality theory of Strassen introduced and studied in [Str86, Str88, Str88, Str91, Tob91, Bür90] (see also [CVZ18] and [Zui18]) that gives a *dual formulation* for the (non-symmetric) asymptotic subrank, asymptotic rank and asymptotic restriction preorder in terms of the *asymptotic spectrum of tensors*. The asymptotic subrank of $f \in \mathbb{F}^{n_1} \otimes \cdots \otimes \mathbb{F}^{n_k}$ is defined as $\tilde{\mathrm{Q}}(f) = \lim_{n\to\infty} \mathrm{Q}(f^{\otimes n})^{1/n}$, the asymptotic rank is defined as $\tilde{\mathrm{R}}(f) = \lim_{n\to\infty} \mathrm{R}(f^{\otimes n})^{1/n}$ and the asymptotic restriction preorder is defined by $f \lesssim g$ if and only if $f^{\otimes n} \leq g^{\otimes(n+o(n))}$. As an application of the results of Section 4.3.2 we prove a strong connection between this theory and the natural symmetric variation.

The asymptotic spectrum of tensors (for any fixed $k \in \mathbb{N}$ and field $\mathbb{F}$) is defined as the set X of all real-valued maps from k-tensors over $\mathbb{F}$ to the nonnegative reals that are additive under the direct sum, multiplicative under the tensor product, monotone under the restriction preorder and normalized to 1 on the diagonal tensor $\langle 1 \rangle$ of size one. The duality theory says that: the asymptotic rank equals the pointwise maximum over all elements in the asymptotic spectrum of tensors, the asymptotic subrank equals the pointwise minimum over all elements in the asymptotic spectrum of tensors, and the asymptotic restriction preorder is characterized by $f \lesssim g$ if and only if for every ϕ in the asymptotic spectrum X it holds that $\phi(f) \leq \phi(g)$.

4.4.1 Asymptotic spectrum duality

We introduce the asymptotic spectrum of *symmetric* tensors as the natural symmetric variation on Strassen's asymptotic spectrum of tensors, to give a duality theory for the asymptotic symmetric (sub)rank and restriction preorder. We have defined the asymptotic symmetric subrank before. The asymptotic symmetric rank is similarly defined as $\tilde{\mathrm{R}}_{\mathrm{s}}(f) = \lim_{n\to\infty} \mathrm{R}_{\mathrm{s}}(f^{\otimes n})^{1/n}$ and the asymptotic symmetric restriction preorder is defined by $f \lesssim_{\mathrm{s}} g$ if and only if $f^{\otimes n} \leq_{\mathrm{s}} g^{\otimes(n+o(n))}$.

We define the asymptotic spectrum of symmetric tensors (for any fixed $k \in \mathbb{N}$ and field $\mathbb{F}$) as the set X_{s} of all real-valued maps from symmetric k-tensors over $\mathbb{F}$ to the nonnegative reals that are additive under the direct sum, multiplicative under the tensor product, monotone under the *symmetric* restriction preorder, and normalized to 1 on the diagonal tensor $\langle 1 \rangle$. It follows readily from the general part of the theory in [Str88] (see also [Zui18]) that the asymptotic spectrum of symmetric tensors X_{s} gives a dual formulation for the asymptotic symmetric subrank, asymptotic symmetric rank and asymptotic symmetric restriction preorder.

Theorem 4.4.1. *Let $\mathbb{F}$ be an algbraically closed field of characteristic at least $k+1$. Let X_{s} be the asymptotic spectrum of symmetric k-tensors. Let f and g be symmetric k-tensors. Then*

$$\tilde{\mathrm{Q}}_{\mathrm{s}}(f) = \min_{\phi \in X_{\mathrm{s}}} \phi(f),$$
$$\tilde{\mathrm{R}}_{\mathrm{s}}(f) = \max_{\phi \in X_{\mathrm{s}}} \phi(f),$$
$$f \lesssim_{\mathrm{s}} g \iff \forall \phi \in X_{\mathrm{s}},\ \phi(f) \leq \phi(g).$$

We will not give the proof of Theorem 4.4.1 as it follows along the same lines as the original proof in [Str88] (see also [Zui18]). The bulk of the proof is to show that the symmetric restriction preorder is a so-called "good preorder" ([Str88]) or Strassen preorder ([Zui18]). The only non-standard ingredient for the proof is the fact that for every nonzero symmetric k-tensor f either f is equivalent to $\langle 1 \rangle$ or $\tilde{\mathrm{Q}}_{\mathrm{s}}(f) > 1$, which follows from Theorem 4.3.4 and the fact that this property holds for $\tilde{\mathrm{Q}}$.

4.4.2 Surjective restriction from the asymptotic spectrum

The results of Section 4.3.2 answer a structural question: how are the asymptotic spectrum of tensors X and the asymptotic spectrum of symmetric tensors X_{s} related? One relation is clear: for every element $\phi \in X$ the restriction of ϕ to symmetric tensors is an element of X_{s}. We thus have the restriction map $r : X \to X_{\mathrm{s}}$ that maps $\phi \in X$ to the restriction of ϕ to symmetric tensors. We prove:

Theorem 4.4.2. *The restriction map $r : X \to X_{\mathrm{s}}$ is surjective.*

Theorem 4.4.2 has two readings: (1) if we understand what the elements are of the asymptotic spectrum of tensors X, then we also understand what the elements are of the asymptotic spectrum of symmetric tensors X_{s} by restriction, and (2) for any element $\psi \in X_{\mathrm{s}}$ there is an *extension* $\phi \in X$ such that ϕ restricts to ψ.

Theorem 4.4.2 follows from our Theorem 4.3.5 together with an application of the following powerful theorem from the theory of asymptotic spectra. The theorem uses the notion of a good preorder or Strassen preorder for which we refer the reader to the literature.

Theorem 4.4.3 ([Str88], [Zui18, Corollary 2.18])**.** *Let S be a semiring with a Strassen preorder P. Let T be a subsemiring of S. Then the restriction map from the asymptotic spectrum of S to the asymptotic spectrum of T is surjective.*

Proof of Theorem 4.4.2. We give a sketch of the proof. The proof is an application of Theorem 4.4.3. Let S be the semiring of k-tensors and let P be the asymptotic restriction preorder. This is a Strassen preorder. Let T be the subsemiring of S of symmetric k-tensors. Then Theorem 4.4.3 implies that the restriction map from the asymptotic spectrum of S with the asymptotic restriction preorder to the asymptotic spectrum of T with the asymptotic restriction preorder is surjective. Since the asymptotic restriction preorder on symmetric tensors coincides with the asymptotic symmetric restriction preorder by Theorem 4.4.2, the claim follows. □

To summarize what we have just seen, the asymptotic spectrum of tensors X and the asymptotic spectrum of symmetric tensors X_{s} are tightly related since the restriction map from the first to the second is surjective. What are the elements of X and X_{s}? A long line of work [Str86, Str88, Str88, Str91, Str05, Tob91, Bür90, CVZ18, CLZ20] has been devoted to this question. Our best understanding is for the case that the ground field $\mathbb{F}$ is the complex numbers[5] and that is what we will focus our discussion on here and in the next section.

The known elements in X (over the complex numbers) are a family of functions called the *quantum functionals*. These were introduced in [CVZ18] and are based on an information-theoretic and representation-theoretic study of powers of tensors. The quantum functionals more precisely form a continuous family F^θ indexed by probability distributions θ on $[k]$. This family includes the flattening ranks, but also includes more interesting functions that are properly real-valued which reveal asymptotic information that the flattening ranks do not reveal. It is possible but not known whether the quantum functionals are all elements of X. Proving this is a central open problem of the theory. In particular, the quantum functionals being all elements of X would imply that the matrix multiplication exponent ω equals 2, which would be a breakthrough result in complexity theory.

We may restrict the quantum functionals to symmetric tensors to find an infinite family of elements in X_{s}. Since we do not know whether the quantum functionals are all elements of X, we can, however, not conclude from Theorem 4.4.2 that their restriction gives all elements of X_{s}.

What we will do in the next section is give a natural construction of a single element in X_{s} following the same ideas as for the construction of the quantum functionals but applied directly to the symmetric restriction preorder. This single element we call the *symmetric quantum functional*. What we then find is that this symmetric quantum functional on symmetric tensors in fact equals the *uniform* quantum functional $F^{(1/k,\ldots,1/k)}$. Thus we do not find a new element in X_{s}, but we do find a different description of the uniform quantum functional restricted to symmetric tensors, and this might be algorithmically beneficial. This symmetric quantum functional is the *pointwise smallest* element among all elements in X_{s} that we currently know, and from previous work it follows that it equals the asymptotic slice rank (on symmetric tensors). Having discussed the plan we will now go into the details in the next section.

4.5 Symmetric quantum functional

In Section 4.4 we introduced the asymptotic spectrum of symmetric tensors X_{s} and proved a duality theorem for the asymptotic symmetric (sub)rank and restriction preorder in terms of it.

We use the ideas of the construction of the quantum functionals $F^\theta \in X$ from [CVZ18] to construct the *symmetric quantum functional* $F \in X_{\mathrm{s}}$ over the field of complex numbers. Let us from now on fix the base field to be the field of complex numbers. In fact we will take a more general approach and define the symmetric quantum functional not just for symmetric tensors but for arbitrary tensors.

[5] It is known that the asymptotic spectrum can only depend on the characteristic of the field [Str88].

Before recalling the definition of the quantum functionals F^θ and giving the new definition of the symmetric quantum functional F, here is what we will find. For symmetric tensors we will show that:

Theorem 4.5.1. *On symmetric tensors $F = F^{(1/k,\dots,1/k)}$.*

This gives an alternative description of the uniform quantum functional $F^{(1/k,\dots,1/k)}$, which may have algorithmic benefits.

In particular on symmetric tensors the symmetric quantum functional is in the asymptotic spectrum of symmetric tensors X_{s}.

Theorem 4.5.2. *On symmetric tensors we have $F \in X_{\mathrm{s}}$.*

For general tensors we find the following.

Theorem 4.5.3. *On arbitrary tensors we have $F \geq F^{(1/k,\dots,1/k)}$.*

In particular, since $F^{(1/k,\dots,1/k)} \geq \tilde{\mathrm{Q}}$ (because every quantum functional F^θ is in the asymptotic spectrum of tensors X), we also find $F \geq \tilde{\mathrm{Q}}$ on arbitrary tensors. However, via a known connection from [CVZ18] between the quantum functionals and the asymptotic slice rank (the pointwise minimum $\min_\theta F^\theta$ equals the asymptotic slice rank), we find that F, as a method to upper bound the Shannon capacity of hypergraphs, suffers from the induced matching barrier.

4.5.1 From quantum functionals to symmetric quantum functional

Before defining the quantum functionals and symmetric quantum functional and giving the proofs of the above, we must introduce some standard notation. Let $\mathcal{H}$ be a complex finite-dimensional Hilbert space with dimension $\dim(\mathcal{H}) = d$. Thus $\mathcal{H} \cong \mathbb{C}^d$. Recall a *state* or *density operator* on $\mathcal{H}$ is a positive semidefinite linear map $\rho : \mathcal{H} \to \mathcal{H}$ with $\mathrm{tr}(\rho) = 1$. Let $\mathrm{D}(\mathcal{H})$ be the set of states on $\mathcal{H}$. For $\rho \in \mathrm{D}(\mathcal{H})$, let $\mathrm{spec}(\rho) = (\lambda_1, \dots, \lambda_d)$ be the sequence of eigenvalues of ρ, ordered non-increasingly, that is, $\lambda_1 \geq \cdots \geq \lambda_d$. Since $\mathrm{tr}(\rho) = 1$, the sequence of eigenvalue of ρ is a probability distribution. It thus makes sense to define $H(\mathrm{spec}(\rho)) := -\sum_{j=1}^d \lambda_j \log \lambda_j$.

Given a state ρ on $\mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_k$, the jth *marginal* is the element $\rho_j = \mathrm{tr}_{\mathcal{H}_1 \dots \mathcal{H}_{j-1} \mathcal{H}_{j+1} \dots \mathcal{H}_k}(\rho)$ obtained from ρ by a partial trace. The jth marginal is itself a state, that is, $\rho_j \in \mathcal{S}(\mathcal{H}_j)$. Consider a nonzero element $f \in \mathcal{H}^{\otimes k}$. Then $\rho(f) = \frac{ff^\dagger}{\|f\|^2} \in \mathrm{D}(\mathcal{H}^{\otimes k})$, where $f^\dagger$ denotes the conjugate transpose of f, and we can consider the jth marginal $\rho_j(f) \in \mathcal{S}(\mathcal{H})$. Let $\mathrm{GL}(d)$ denote the set of invertible matrices acting on $\mathcal{H}$. For a tensor $f \in \mathcal{H}^{\otimes k}$, let $\overline{\mathrm{GL}(d) \cdot f}$ be the Euclidean closure (or equivalently Zariski closure) of the orbit $\{(g \otimes \cdots \otimes g) f : g \in \mathrm{GL}(d)\}$.

We begin with the definition of the symmetric quantum functional.

Definition 4.5.4 (Symmetric quantum functional)**.** Let $f \in \mathcal{H}^{\otimes k}$ be nonzero. We define the *symmetric quantum functional* F by $F(f) = 2^{E(f)}$ where

$$E(f) = \max\{H(p) : p \in \Delta(f)\},$$

where we define the subset $\Delta(f) \subseteq \mathbb{R}^d$, for $d = \dim(\mathcal{H})$, as

$$\Delta(f) = \Big\{\mathrm{spec}\Big(\frac{\rho_1(s) + \cdots + \rho_k(s)}{k}\Big) : s \in \overline{\mathrm{GL}(d) \cdot f} \setminus \{0\}\Big\}.$$

From the work of [NM84] and [Bri87] it follows that $\Delta(f)$ is a convex polytope.

The definition of the symmetric quantum functional F is inspired by the family of quantum functionals F^θ. Our main results about the symmetric quantum functional give precise relations between F and F^θ.

Definition 4.5.5 (Quantum functionals)**.** Let $\theta \in \mathcal{P}([k])$ and let $f \in \mathcal{H}^{\otimes k}$. The quantum functionals are defined by $F^\theta(f) = 2^{E^\theta(f)}$ where

$$E^\theta(f) = \max\left\{\sum_{i=1}^{s} \theta(i) H(\rho_i(s)) : s \in \overline{\mathrm{GL}(d)^{\times k} \cdot f} \setminus \{0\}\right\}$$

where $\mathrm{GL}(d)^{\times k} \cdot f = \{(g_1 \otimes \cdots \otimes g_k) \cdot f : g_1, \ldots, g_k \in \mathrm{GL}(d)\}$.

There is an asymptotic connection between the quantum functionals and the slice rank, which we will be using.

Theorem 4.5.6 ([CVZ18])**.** *For any $f \in \mathcal{H}^{\otimes k}$ the limit $\lim_{n\to\infty} \mathrm{SR}(f^{\otimes n})^{1/n}$ exists and equals the minimization $\min_{\theta \in \mathcal{P}([k])} F^\theta(f)$.*

4.5.2 Properties and relations

Now we are ready to state the precise results on the symmetric quantum functional. These results in particular imply the three main results that we stated above in Theorem 4.5.1, Theorem 4.5.2 and Theorem 4.5.3.

First of all, we prove that the symmetric quantum functional is at least the uniform quantum functional, and we show that the latter can be obtained as the regularization of the former:

Theorem 4.5.7. *Let $f \in \mathcal{H}^{\otimes k}$ be any tensor. Let $\theta = \left(\frac{1}{k}, \ldots, \frac{1}{k}\right)$. Then*

$$\lim_{n\to\infty} \mathrm{SR}(f^{\otimes n})^{1/n} \leq F^\theta(f) \leq F(f) \text{ and } \lim_{n\to\infty} F(f^{\otimes n})^{1/n} = \inf_n F(f^{\otimes n})^{1/n} = F^\theta(f).$$

Second, on symmetric tensors we prove the following even stronger connection between the symmetric quantum functional and the uniform quantum functional:

Theorem 4.5.8. *Let $f \in \mathcal{H}^{\otimes k}$ be a symmetric tensor. Then*

$$\lim_{n\to\infty} \mathrm{SR}(f^{\otimes n})^{1/n} = F^{(1/k,\ldots,1/k)}(f) = F(f).$$

Third, from the equality $F = F^{(1/k,\ldots,1/k)}$ on symmetric tensors (Theorem 4.5.8), and the known properties of $F^{(1/k,\ldots,1/k)}$, we directly obtain all of the following properties of the symmetric quantum functional F:

Corollary 4.5.9. *For any symmetric $f \in (\mathbb{C}^d)^{\otimes k}$ and $g \in (\mathbb{C}^e)^{\otimes k}$, and any $r \in \mathbb{N}$, we have*

1. $F(\langle r \rangle) = r$
2. $F(f \oplus g) = F(f) + F(g)$

3. $F(f \otimes g) = F(f)F(g)$

4. *if* $f \leq_{\mathrm{s}} g$ *then* $F(f) \leq F(g)$.

Therefore, the symmetric quantum functional belongs to the asymptotic spectrum of symmetric tensors X_{s}, which we discussed in Section 4.4.

We will now give the proofs of the above Theorem 4.5.7 and Theorem 4.5.8. We will need another characterization of $\Delta(f)$ from representation theory. Let λ be a partition of nk into at most d parts. We denote this by $\lambda \vdash_d nk$. Then $\bar{\lambda} := \lambda/nk = (\lambda_1/nk, \ldots, \lambda_d/nk)$ is a probability distribution on $[d]$. The symmetric group $\mathfrak{S}_{nk}$ acts on $(\mathcal{H}^{\otimes k})^{\otimes n}$ by permuting the tensor legs, that is, $\pi \cdot (v_1 \otimes \cdots \otimes v_{nk}) = v_{\pi^{-1}(1)} \otimes \cdots \otimes v_{\pi^{-1}(nk)}$ for $\pi \in \mathfrak{S}_{nk}$. The general linear group $\mathrm{GL}(d)$ acts on $(\mathcal{H}^{\otimes k})^{\otimes n}$ via the diagonal embedding $\mathrm{GL}(d) \to \mathrm{GL}(d)^{\times nk} : g \mapsto (g, \ldots, g)$, that is, $g \cdot v = (g \otimes \cdots \otimes g)v$ for $g \in \mathrm{GL}(d), v \in (\mathcal{H}^{\otimes k})^{\otimes n}$. The Schur–Weyl duality gives a decomposition of the space $(\mathcal{H}^{\otimes k})^{\otimes n}$ into direct sum of irreducible $\mathfrak{S}_{nk} \times \mathrm{GL}(d)$ representations. More precisely,

$$(\mathcal{H}^{\otimes k})^{\otimes n} \cong \bigoplus_{\lambda \vdash_d nk} [\lambda] \otimes \mathbb{S}_\lambda(\mathcal{H}),$$

where $\mathbb{S}_\lambda(\mathcal{H})$ is an irreducible representation of $\mathrm{GL}(d)$ and $[\lambda]$ is an irreducible representation of $\mathfrak{S}_{nk}$. Let $P_\lambda : (\mathcal{H}^{\otimes k})^{\otimes n} \to (\mathcal{H}^{\otimes k})^{\otimes n}$ be the equivariant projector onto the isotypical component of type λ, that is, onto the subspace of $(\mathcal{H}^{\otimes k})^{\otimes n}$ which isomorphic to $\mathbb{S}_\lambda(\mathcal{H}) \otimes [\lambda]$. Based on [Bri87], [Fra02], [Str05] (and also [Wal14, Section 2.1] and [Zui18, Chapter 6]) we have the following characterization of $\Delta(f)$.

Lemma 4.5.10. *The polytope* $\Delta(f)$ *is the Euclidean closure of the set*

$$\left\{ \frac{\lambda}{nk} : \exists n \in \mathbb{N}_{\geq 1}, \lambda \vdash_d nk, P_\lambda f^{\otimes n} \neq 0 \right\}.$$

Proof. The proof of the Lemma 4.5.10 can be found in the Appendix A.1. □

Proof of Theorem 4.5.7. We decompose $\mathcal{H}^{\otimes n}$ into a direct sum of irreducible $\mathfrak{S}_n \times \mathrm{GL}(d)$ representations as

$$\mathcal{H}^{\otimes n} \cong \bigoplus_{\lambda \vdash_d n} [\lambda] \otimes \mathbb{S}_\lambda(\mathcal{H}). \tag{4.2}$$

Let P_λ be the equivariant projector onto the isotypical component of type λ. The uniform quantum functional $F^{(\frac{1}{k}, \ldots, \frac{1}{k})}(f)$ has another characterization as follows [CVZ18]:

$$F^{(\frac{1}{k}, \ldots, \frac{1}{k})}(f) = \sup\Big\{ \Big(\prod_{i=1}^{k} \dim[\lambda^i] \Big)^{1/kn} : \exists n \in \mathbb{N}_{\geq 1},\ \lambda^i \vdash_d n, (P_{\lambda^1} \otimes \cdots \otimes P_{\lambda^k}) f^{\otimes n} \neq 0 \Big\}.$$

For the symmetric quantum functional, using the characterization of $\Delta(f)$ from representation theory, we have

$$F(f) = \sup \left\{ (\dim[\lambda])^{1/kn} : \exists n \in \mathbb{N}_{\geq 1}, \lambda \vdash kn, P_\lambda f^{\otimes n} \neq 0 \right\}.$$

We may write $(\mathcal{H}^{\otimes n})^{\otimes k}$ as a direct sum of irreducibles under the action of $\mathfrak{S}_{nk}$ as

$$\left(\mathcal{H}^{\otimes n}\right)^{\otimes k} \cong \bigoplus_{\lambda \vdash_d kn} ([\lambda])^{\oplus m_\lambda} \tag{4.3}$$

where $m_\lambda = \dim\left(\mathbb{S}_\lambda(\mathcal{H})\right)$. We view $\mathfrak{S}_n^{\times k}$ naturally as a subgroup of $\mathfrak{S}_{nk}$. For any $\lambda \vdash_d kn$ the restriction of $[\lambda]$ to the action of $\mathfrak{S}_n^{\times k}$ decomposes further as a direct sum of irreducibles under the action of $\mathfrak{S}_n^{\times k}$, so that

$$[\lambda] \cong \bigoplus_{\lambda^1 \vdash_d n, \ldots, \lambda^k \vdash_d n} \left([\lambda^1] \otimes \cdots \otimes [\lambda^k]\right)^{\oplus c_{\lambda^1,\ldots,\lambda^k}} \tag{4.4}$$

where $c_{\lambda^1,\ldots,\lambda^k}$ are multiplicities. Let $\lambda^1 \vdash_d n, \ldots, \lambda^k \vdash_d n$. Then $[\lambda^1] \otimes \cdots \otimes [\lambda^k]$ is irreducible representation of $\mathfrak{S}_n \times \cdots \times \mathfrak{S}_n$. This gives us the finer decomposition into irreducibles under the action of $\mathfrak{S}_n^{\times k}$ as

$$\left(\mathcal{H}^{\otimes n}\right)^{\otimes k} \cong \bigoplus_{\lambda^1 \vdash_d n, \ldots, \lambda^k \vdash_d n} \left([\lambda^1] \otimes \cdots \otimes [\lambda^k]\right)^{\oplus m_{\lambda^1,\ldots,\lambda^k}} \tag{4.5}$$

where $m_{\lambda^1,\ldots,\lambda^k} = \prod_{i=1}^k \dim\left(\mathbb{S}_{\lambda_i}(\mathcal{H})\right)$.

For any n and $\lambda^1 \vdash_d n, \ldots, \lambda^k \vdash_d n$ such that $\left(P_{\lambda^1} \otimes \cdots \otimes P_{\lambda^k}\right) f^{\otimes n} \neq 0$ the equivariant projection of $f^{\otimes n}$ on

$$\left([\lambda^1] \otimes \cdots \otimes [\lambda^k]\right)^{\oplus m_{\lambda^1,\ldots,\lambda^k}}$$

is non-zero. From (4.4) we know that there is a $\lambda \vdash_d kn$ such that $[\lambda^1] \otimes \cdots \otimes [\lambda^k]$ is a subspace of $[\lambda]$. For this λ it holds that $P_\lambda f^{\otimes n} \neq 0$ and $\dim[\lambda] \geq \prod_{i=1}^k \dim\left([\lambda^i]\right)$. This implies $F(f) \geq F^{(\frac{1}{k},\ldots,\frac{1}{k})}(f)$.

For any tensor $s \in \mathcal{H}^{\otimes k}$, it follows from a standard property of the von Neumann entropy [NC11, Theorem 11.10] that

$$H\left(\frac{\rho_1(s) + \cdots + \rho_k(s)}{k}\right) \leq \frac{H(\rho_1(s)) + \cdots + H(\rho_k(s))}{k} + \log k.$$

This implies $F(f) \leq kF^{(\frac{1}{k},\ldots,\frac{1}{k})}(f)$. Thus we have proven that

$$F^{(\frac{1}{k},\ldots,\frac{1}{k})}(f) \leq F(f) \leq kF^{(\frac{1}{k},\ldots,\frac{1}{k})}(f)$$

holds for every tensor f. In particular, applying this to the tensor power $f^{\otimes n}$ we have

$$F^{(\frac{1}{k},\ldots,\frac{1}{k})}(f^{\otimes n}) \leq F(f^{\otimes n}) \leq kF^{(\frac{1}{k},\ldots,\frac{1}{k})}(f^{\otimes n}).$$

Since $F^{(\frac{1}{k},\ldots,\frac{1}{k})}$ is multiplicative [CVZ18], we have

$$F^{(\frac{1}{k},\ldots,\frac{1}{k})}(f) \leq F(f^{\otimes n})^{1/n} \leq k^{1/n} F^{(\frac{1}{k},\ldots,\frac{1}{k})}(f).$$

Taking $n \to \infty$, we obtain $\lim_{n\to\infty} F(f^{\otimes n})^{1/n} = F^{(\frac{1}{k},\ldots,\frac{1}{k})}(f)$.

Finally, since F is sub-multiplicative (see in Appendix A.2), the limit $\lim_{n\to\infty} F(f^{\otimes n})^{1/n}$ equals the infimum $\inf_n F(f^{\otimes n})^{1/n}$ by Fekete's lemma. $\square$

Proof of Theorem 4.5.8. Let S be the set of symmetric tensors in $\overline{(\mathrm{GL}(d)^{\times k}) \cdot f} \setminus \{0\}$. Since f is a symmetric tensor, for any matrix A the tensor $(A \otimes \cdots \otimes A)f$ is also a symmetric tensor. Therefore $\overline{\mathrm{GL}(d) \cdot f} \setminus \{0\} \subseteq S$. Moreover, if s is a symmetric tensor then all marginal density matrices are equal: $\rho_1(s) = \cdots = \rho_k(s)$. Thus, for any $\theta \in \mathcal{P}([k])$, we have $E^\theta(s) = \rho_1(s)$. This implies $F(f) \leq F^\theta(f)$ since both $F(f)$ and $F^\theta(f)$ are given by the supremum of the same function and for $F(f)$ the supremum is taken over a smaller set than for $F^\theta(f)$. By Theorem 4.5.7 we have $F(f) = F^\theta(f)$ with $\theta = (\frac{1}{k}, \ldots, \frac{1}{k})$. Moreover, from the Proposition 4.5.6 we have $\lim_{n\to\infty} \mathrm{SR}(f^{\otimes n})^{1/n} = \min_{\theta \in \mathcal{P}([k])} F^\theta(f) \geq F(f)$, which implies $\lim_{n\to\infty} \mathrm{SR}(f^{\otimes n})^{1/n} = F(f)$. This proves the claim. □

4.6 Conclusion

In this chapter, we introduced the symmetric subrank of tensors and proved precise relations and separations between subrank and symmetric subrank. Then, we showed that for symmetric tensors the subrank and the symmetric subrank are asymptotically equal. This proves the asymptotic subrank analogon of a conjecture known as Comon's conjecture in the theory of tensors. This result allows us to prove a strong connection between the general and the symmetric versions of an asymptotic duality theorem of Strassen. Finally, we introduced a representation-theoretic method to asymptotically bound the symmetric subrank called the symmetric quantum functional in analogy with the quantum functionals, then studied the relations between these functionals. Nevertheless, the symmetric quantum functional cannot give better bounds than the quantum functionals which itself suffers from the induced matching barrier and cannot be used to make progress on corner problem. But we hope that future improved asymptotic upper bounds on the symmetric subrank can still overcome the induced matching barrier. In particular, we leave it as an open question to define a good symmetric version of Strassen's support functionals.

Chapter 5

Efficient bounds on quantum capacities

This chapter is based on joint work with Omar Fawzi and Ala Shayeghi [FST21].

5.1 Introduction

The optimal rates for many quantum information processing tasks of interest can be characterized in terms of a regularized divergence between quantum channels. For a divergence $\mathbf{D}$ defined on quantum states, the corresponding channel divergence is defined by maximizing the divergence between the channel outputs over the set of possible inputs. There are two natural variants: for quantum channels $\mathcal{N}$ and $\mathcal{M}$ the non-stabilized divergence is given by only allowing input states ρ in the input space of $\mathcal{N}$ and $\mathcal{M}$

$$\overline{\mathbf{D}}(\mathcal{N}\|\mathcal{M}) = \sup_{\rho} \mathbf{D}(\mathcal{N}(\rho)\|\mathcal{M}(\rho)) \ ,$$

whereas the stabilized version allows arbitrary input states

$$\mathbf{D}(\mathcal{N}\|\mathcal{M}) = \sup_{\rho} \mathbf{D}((\mathcal{I}\otimes\mathcal{N})(\rho)\|(\mathcal{I}\otimes\mathcal{M})(\rho)) \ ,$$

where $\mathcal{I}$ is the identity channel. The most well-known example illustrating these two variants is when $\mathbf{D}$ is the trace distance, then $\overline{\mathbf{D}}(\mathcal{N}\|\mathcal{M})$ is the superoperator trace norm and $\mathbf{D}(\mathcal{N}\|\mathcal{M})$ is the diamond norm, and it is known that we can have $\overline{\mathbf{D}}(\mathcal{N}\|\mathcal{M}) \ll \mathbf{D}(\mathcal{N}\|\mathcal{M})$ [KSV02].

When analyzing tasks in the independent and identically distributed limit, an important divergence $\mathbf{D}$ is the Umegaki divergence D defined by $\mathrm{D}(\rho\|\sigma) = \mathrm{tr}(\rho\log\rho) - \mathrm{tr}(\rho\log\sigma)$. For example, in asymmetric hypothesis testing between channels $\mathcal{N}$ and $\mathcal{M}$, the Stein exponent is characterized using D, namely in terms of the *regularized* channel divergence $\mathrm{D}^{\mathrm{reg}}(\mathcal{N}\|\mathcal{M}) := \sup_n \frac{1}{n}\mathrm{D}(\mathcal{N}^{\otimes n}\|\mathcal{M}^{\otimes n})$ [WBHK20, WW19, FFRS20]. In fact, it turns out that the channel divergence D is in general non-additive [FFRS20], in which case we have $\mathrm{D}^{\mathrm{reg}}(\mathcal{N}\|\mathcal{M}) > \mathrm{D}(\mathcal{N}\|\mathcal{M})$.

Another example is the Holevo information of a quantum channel $\mathcal{N}$, which is given by $\chi(\mathcal{N}) = \min_\sigma \overline{\mathrm{D}}(\mathcal{N}\|\mathcal{T}_\sigma)$ where $\mathcal{T}_\sigma$ is the replacer channel that outputs σ for any input density operator [OPW97]. The Holevo-Schumacher-Westmorland theorem (see e.g., [Wil13])

states that the classical capacity of $\mathcal{N}$ is given by $\chi^{\mathrm{reg}}(\mathcal{N}) := \sup_n \frac{1}{n}\chi(\mathcal{N}^{\otimes n})$ and the regularization is needed for some channels, as shown by [Has09].

The objective of this chapter is to provide efficient ways of computing, or more specifically upper bounding, such regularized channel divergences. In order to achieve this, we use the recently introduced *#-Rényi divergence* [FF21b]. The #-Rényi divergence of order $\alpha > 1$, denoted by $\mathrm{D}^{\#}_{\alpha}$, between two positive semidefinite operators is defined as

$$\begin{aligned}
\mathrm{D}^{\#}_{\alpha}(\rho\|\sigma) &:= \frac{1}{\alpha-1}\log \mathrm{Q}^{\#}_{\alpha}(\rho\|\sigma),\\
\mathrm{Q}^{\#}_{\alpha}(\rho\|\sigma) &:= \min_{A\geq 0} \mathrm{tr}(A) \quad \text{s.t.} \quad \rho \leq \sigma \#_{\frac{1}{\alpha}} A\,,
\end{aligned}$$

where $\#_\alpha$ denotes the α-geometric mean of two positive definite matrices ρ and σ. This divergence has several desirable computational and operational properties such as an efficient semidefinite programming representation for states and channels, and a chain rule property. An important property of this new divergence is that its regularization is equal to the sandwiched (also known as the minimal) quantum Rényi divergence. Let $\mathcal{N}_{X\to Y}$ and $\mathcal{M}_{X\to Y}$ be two quantum channels. The channel divergence corresponding to $\mathrm{D}^{\#}_{\alpha}$ can be expressed in terms of a convex optimization program [FF21b] as follows.

$$\begin{aligned}
\mathrm{D}^{\#}_{\alpha}(\mathcal{N}\|\mathcal{M}) &:= \frac{1}{\alpha-1}\log \mathrm{Q}^{\#}_{\alpha}(\mathcal{N}\|\mathcal{M}),\\
\mathrm{Q}^{\#}_{\alpha}(\mathcal{N}\|\mathcal{M}) &:= \min_{A_{XY}\geq 0} \|\mathrm{tr}_Y(A_{XY})\|_\infty \quad \text{s.t.} \quad (J^{\mathcal{N}}_{XY}) \leq (J^{\mathcal{M}}_{XY})\#_{1/\alpha} A_{XY}\ ,
\end{aligned}$$

where $\|.\|_\infty$ denotes the operator norm and $J^{\mathcal{N}}_{XY}, J^{\mathcal{M}}_{XY}$ are choi matrices of $\mathcal{N}, \mathcal{M}$, respectively.

As a first application, we consider the task of approximating the *regularized sandwiched Rényi divergence*. For $\alpha \in (1,\infty)$, the regularized sandwiched α-Rényi divergence between channels $\mathcal{N}_{X\to Y}$ and $\mathcal{M}_{X\to Y}$ is defined as

$$\widetilde{\mathrm{D}}^{\mathrm{reg}}_{\alpha}(\mathcal{N}\|\mathcal{M}) := \lim_{k\to\infty} \frac{1}{k}\widetilde{\mathrm{D}}_{\alpha}(\mathcal{N}^{\otimes k}\|\mathcal{M}^{\otimes k})\,.$$

Ref. [FF21b] provided a converging hierarchy of upper bounds on the regularized divergence between channels:

$$\frac{1}{k}\mathrm{D}^{\#}_{\alpha}(\mathcal{N}^{\otimes k}\|\mathcal{M}^{\otimes k}) - \nu(d,k,\alpha) \leq \widetilde{\mathrm{D}}^{\mathrm{reg}}_{\alpha}(\mathcal{N}\|\mathcal{M}) \leq \frac{1}{k}\mathrm{D}^{\#}_{\alpha}(\mathcal{N}^{\otimes k}\|\mathcal{M}^{\otimes k})\ , \tag{5.1}$$

where $d = \dim X \dim Y$ and $\nu(d,k,\alpha) = \frac{1}{k}\frac{\alpha}{\alpha-1}(d^2+d)\log(k+d)$. Here $\widetilde{\mathrm{D}}_\alpha$ is the *sandwiched Rényi divergence* [MDS+13], [WWY14] of order $\alpha \in (1,\infty)$ (see Section 5.2 for definition). The regularized sandwiched Rényi divergence between channels can be used to obtain improved characterization of many information processing tasks such as channel discrimination [FF21a, FF21b]. However, the sandwiched Rényi divergence between channels is non-additive in general [FFRS20] and it is unclear whether its regularization can be computed efficiently.

From 5.1, $\widetilde{\mathrm{D}}^{\mathrm{reg}}_{\alpha}(\mathcal{N}\|\mathcal{M})$ can be approximated by $\frac{1}{k}\mathrm{D}^{\#}_{\alpha}(\mathcal{N}^{\otimes k}\|\mathcal{M}^{\otimes k})$ with arbitrary accuracy for sufficiently large k in finite time. Namely, if we take $k = \lceil \frac{8\alpha d^3}{(\alpha-1)\epsilon}\rceil$ then we

have

$$|\widetilde{\mathrm{D}}_\alpha^{\mathrm{reg}}(\mathcal{N}\|\mathcal{M}) - \frac{1}{k}\mathrm{D}_\alpha^{\#}(\mathcal{N}^{\otimes k}\|\mathcal{M}^{\otimes k})| \leq \epsilon\,.$$

However, the size of the program for computing $\mathrm{D}_\alpha^{\#}(\mathcal{N}^{\otimes k}\|\mathcal{M}^{\otimes k})$ grows exponentially with k.

New results in this chapter

We exploit the symmetries of the resulting hierarchy of optimization programs to obtain a concise representation and solve it efficiently. Specifically, for quantum channels $\mathcal{N}, \mathcal{M}$, we show in Theorem 5.4.3 that the permutation symmetry of the optimization program defining $\mathrm{D}_\alpha^{\#}(\mathcal{N}^{\otimes k}\|\mathcal{M}^{\otimes k})$ can be used to transform it into a semidefinite program with $\mathrm{poly}(k)$ variables and constraints compared to the straightforward representation which is of size exponential in k. However, as we will see, a direct implementation of this transformation would require an exponential time computation. In Theorem 5.4.6, we provide an algorithm which performs this transformation in $\mathrm{poly}(k)$ time, for fixed input and output dimensions. As a first application, we consider the task of approximating the regularized sandwiched Rényi divergence between two channels. Note that the sandwiched Rényi divergence (see Section 5.2 for the definition) is in general non-additive [FFRS20], and it is not known whether its regularization is efficiently computable. Ref. [FF21b] shows that the regularized quantity can be approximated up to arbitrary accuracy by $\frac{1}{k}\mathrm{D}_\alpha^{\#}(\mathcal{N}^{\otimes k}\|\mathcal{M}^{\otimes k})$, for sufficiently large k. Our results imply that the regularized sandwiched Rényi divergence between two channels can be approximated up to an accuracy $\epsilon \in (0, 1]$, in time that is polynomial in $1/\epsilon$ (for fixed input/output dimensions). Furthermore, when the channels admit additional group symmetries, we present a general approach to combine these symmetries with the intrinsic permutation invariance to further simplify the problem. As an example demonstrating the potential of this approach, in Section 5.4.2, we apply our method to generalized amplitude damping channels and we show how a very simple symmetry of these channels leads to considerable reductions in the size of the convex optimization program for computing the channel divergence (see Table 5.2).

In Section 5.5, we present a procedure for efficiently computing improved strong converse bounds on the classical capacity of quantum channels by considering the generalized Upsilon-information [WFT19] induced by the $\mathrm{D}^{\#}$ Rényi divergences. To illustrate our method, we apply it to the amplitude damping channel (see Table 5.3 for a comparison with the best previously known bounds). Even though the improvements we obtain for the classical capacity are very small for this channel, the amplitude damping channel $\mathcal{A}_p$ is one of the current challenges as far as the classical capacity is concerned. In particular, it remains open whether $\chi^{\mathrm{reg}}(\mathcal{A}_p) = \chi(\mathcal{A}_p)$. Finally, in Section 5.6, we use our method for computing improved upper bounds on the two-way assisted quantum capacity of channels by considering the generalized Theta-information [FF21a] induced by the $\mathrm{D}^{\#}$ divergences and apply it to the amplitude damping channel, as an example (see Figure 5.2 for a comparison with the best previously known bounds).

Outline of the chapter. We first briefly recall quantum divergences in Section 5.2. Then, in Section 5.3 we present several tools to represent the convex program have symmetry. Next, we provide efficient ways of computing the regularized channel divergences

in Section 5.4. In Section 5.5, we present a procedure for efficiently computing improved strong converse bounds on the classical capacity of quantum channels. Lastly, we use our method for computing improved upper bounds on the two-way assisted quantum capacity of channels in Section 5.6.

5.2 Quantum divergences and Channel divergences

A functional $\mathbf{D} : \mathrm{D}(\mathcal{H}) \times \mathscr{P}(\mathcal{H}) \to \mathbb{R}$ is a *generalized quantum divergence* [PV10, SW13] if it satisfies the data-processing inequality

$$\mathbf{D}(\mathcal{N}(\rho)\|\mathcal{N}(\sigma)) \leq \mathbf{D}(\rho\|\sigma).$$

Let $\rho \in \mathrm{D}(\mathcal{H})$ and $\sigma \in \mathscr{P}(\mathcal{H})$ such that $\rho \ll \sigma$. The *sandwiched Rényi divergence* [MDS+13], [WWY14] of order $\alpha \in (1, \infty)$ is defined as

$$\widetilde{\mathrm{D}}_\alpha(\rho\|\sigma) := \frac{1}{\alpha - 1} \log \operatorname{tr}\left[\left(\sigma^{\frac{1-\alpha}{2\alpha}} \rho \sigma^{\frac{1-\alpha}{2\alpha}}\right)^\alpha\right].$$

The *geometric Rényi divergence* [PR98, Mat15, Tom15, HM17, FF21a] of order α is defined as

$$\widehat{\mathrm{D}}_\alpha(\rho\|\sigma) := \frac{1}{\alpha - 1} \log \operatorname{tr}\left[\sigma^{1/2} \left(\sigma^{-1/2} \rho \sigma^{-1/2}\right)^\alpha \sigma^{1/2}\right].$$

The *max divergence* is defined as

$$\mathrm{D}_{\max}(\rho\|\sigma) := \log \inf\{\lambda > 0 : \rho \leq \lambda\sigma\}.$$

The inverses in these formulations are generalized inverses, i.e., the inverse on the support. When $\rho \ll \sigma$ does not hold, these quantities are set to ∞. Recently, in [FF21b], the authors introduced an interesting quantum Rényi divergence called *#-Rényi divergence*. To define this divergence, we recall the geometric mean of two positive definite matrices.

For $\alpha \in (0, 1)$, the α-geometric mean of two positive definite matrices ρ and σ is defined as

$$\rho \#_\alpha \sigma = \rho^{1/2} (\rho^{-1/2} \sigma \rho^{-1/2})^\alpha \rho^{1/2}.$$

The α-geometric mean has the following properties (see Refs. [KA80] and [FF21b]):

1. Monotonicity: $A \leq C$ and $B \leq D$ implies $A \#_\alpha B \leq C \#_\alpha D$.
2. Transformer inequality: $M(A \#_\alpha B)M^* \leq (MAM^*) \#_\alpha (MBM^*)$, with equality if M is invertible.
3. $(aA) \#_\alpha (bB) = a(b/a)^\alpha (A \#_\alpha) B$, for any $a > 0$ and $b \geq 0$.
4. Joint-concavity/sub-additivity: for any $A_i, B_i \geq 0$ we have

$$\sum_i A_i \#_\alpha B_i \leq \left(\sum_i A_i\right) \#_\alpha \left(\sum_i B_i\right) \quad .$$

5. Direct sum: for any $A_1, A_2, B_1, B_2 \geq 0$, we have

$$(A_1 \oplus A_2)\#_\alpha(B_1 \oplus B_2) = (A_1\#_\alpha B_1) \oplus (A_2\#_\alpha B_2) \ ,$$

where $A_1 \oplus A_2$ form a block diagonal matrix $\begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix}$.

The #-*Rényi divergence* [FF21b] of order α between two positive semidefinite operators is defined as

$$\begin{aligned} \mathrm{D}^{\#}_\alpha(\rho\|\sigma) &:= \frac{1}{\alpha-1}\log \mathrm{Q}^{\#}_\alpha(\rho\|\sigma), \\ \mathrm{Q}^{\#}_\alpha(\rho\|\sigma) &:= \min_{A\geq 0} \mathrm{tr}(A) \quad \text{s.t.} \quad \rho \leq \sigma \#_{\frac{1}{\alpha}} A. \end{aligned}$$

We note that the above convex program may be expressed as a semidefinite program when α is a rational number [FS17, Sag13]. The order between these divergences is summarized in the proposition below.

Proposition 5.2.1. *For any $\rho, \sigma \in \mathscr{P}(\mathcal{H})$ and $\alpha \in (1, 2]$, we have*

$$\mathrm{D}(\rho\|\sigma) \leq \widetilde{\mathrm{D}}_\alpha(\rho\|\sigma) \leq \mathrm{D}^{\#}_\alpha(\rho\|\sigma) \leq \widehat{\mathrm{D}}_\alpha(\rho\|\sigma) \leq \mathrm{D}_{\max}(\rho\|\sigma).$$

For a quantum channel $\mathcal{N}_{X'\to Y}$, a subchannel $\mathcal{M}_{X'\to Y}$ and a generalized quantum divergence $\mathbf{D}$ the corresponding channel divergence [LKDW18] is defined as

$$\mathbf{D}(\mathcal{N}\|\mathcal{M}) := \sup_{\rho_X \in \mathrm{D}(X)} \mathbf{D}(\mathcal{N}_{X'\to Y}(\phi_{XX'})\|\mathcal{M}_{X'\to Y}(\phi_{XX'})) \ ,$$

where $\phi_{XX'}$ is a purification of ρ_X. For $\mathbf{D} = \mathrm{D}^{\#}_\alpha$, the channel divergence can be expressed in terms of a convex optimization program [FF21b] as follows.

$$\mathrm{D}^{\#}_\alpha(\mathcal{N}\|\mathcal{M}) := \frac{1}{\alpha-1}\log \mathrm{Q}^{\#}_\alpha(\mathcal{N}\|\mathcal{M}), \tag{5.2}$$

$$\mathrm{Q}^{\#}_\alpha(\mathcal{N}\|\mathcal{M}) := \min_{A_{XY}\geq 0} \|\mathrm{tr}_Y(A_{XY})\|_\infty \quad \text{s.t.} \quad (J^{\mathcal{N}}_{XY}) \leq (J^{\mathcal{M}}_{XY})\#_{1/\alpha} A_{XY} \ , \tag{5.3}$$

where $\|.\|_\infty$ denotes the operator norm.

The generalization of $\mathrm{D}^{\#}_\alpha$ to channels is subadditive under tensor products [FF21b]: For any $\alpha \in (1, \infty)$, quantum channels $\mathcal{N}_1, \mathcal{N}_2$, and subchannels $\mathcal{M}_1, \mathcal{M}_2$, we have

$$\mathrm{D}^{\#}_\alpha(\mathcal{N}_1 \otimes \mathcal{N}_2\|\mathcal{M}_1 \otimes \mathcal{M}_2) \leq \mathrm{D}^{\#}_\alpha(\mathcal{N}_1\|\mathcal{M}_1) + \mathrm{D}^{\#}_\alpha(\mathcal{N}_2\|\mathcal{M}_2)\,.$$

5.3 Tools for efficiently representing structured convex programs

In this section, we provide the necessary mathematical background on how symmetries in a convex optimization problem can be utilized to represent the program efficiently, we refer the interested reader to references such as [LM11] for more information.

5.3.1 Matrix $*$-algebra background

A subset $\mathcal{A}$ of the set of all $n \times n$ complex matrices is said to be a *matrix $*$-algebra* over $\mathbb{C}$, if it contains the identity operator and is closed under addition, scalar multiplication, matrix multiplication, and taking the conjugate transpose. For our applications, the structure in the optimization programs we consider will allow us to assume that the variables live in such an algebra. A map $\varphi : \mathcal{A} \to \mathcal{B}$ between two matrix $*$-algebras $\mathcal{A}$ and $\mathcal{B}$ is called a $*$-isomorphism if

- φ is a linear bijection,
- $\varphi(AB) = \varphi(A)\varphi(B)$ for all $A, B \in \mathcal{A}$,
- $\varphi(A^*) = \varphi(A)^*$ for all $A \in \mathcal{A}$.

The matrix algebras $\mathcal{A}$ and $\mathcal{B}$ are called *isomorphic* and we write $\mathcal{A} \cong \mathcal{B}$. Note that, by the second property above, $*$-isomorphisms preserve positive semidefiniteness. From a standard result in the theory of matrix $*$-algebra, we get the following structure theorem.

Theorem 5.3.1 (Theorem 1,[Gij05])**.** *Let $\mathcal{A} \subseteq \mathbb{C}^{n\times n}$ be a matrix $*$-algebra. There are numbers t, $m_1, \ldots, m_t$ such that there is a $*$-isomorphism ϕ between $\mathcal{A}$ and a direct sum of complete matrix algebras*

$$\phi \ : \ \mathcal{A} \to \bigoplus_{i=1}^{t} \mathbb{C}^{m_i \times m_i} \ . \tag{5.4}$$

In other words, under the mapping ϕ, all the elements of $\mathcal{A}$ have a common block-diagonal structure. Moreover, this is the finest such decomposition for a generic element of $\mathcal{A}$. We remark that the $*$-isomorphism ϕ can be computed in polynomial time in the dimension of the matrix $*$-algebra $\mathcal{A}$ (see e.g., Theorem 2.7 in Ref. [LM11] and the following discussion, or Ref. [Gij05]).

Regular $*$-representation

In general, computing the block-diagonal decomposition above and the corresponding mapping is a non-trivial procedure. In this section, we introduce a simpler $*$-isomorphism which embeds $\mathcal{A}$ into $\mathbb{C}^{m\times m}$, where $m = \dim \mathcal{A}$.

Let $\mathcal{A}$ be a matrix $*$-algebra of dimension m and $\mathcal{C} = \{C_1, \ldots, C_m\}$ be an orthonormal basis for $\mathcal{A}$ with respect to the Hilbert-Schmidt inner product. Let L be the linear map defined for every $A \in \mathcal{A}$ by the left-multiplication by A. Consider the matrix representation of L with respect to the orthonormal basis $\mathcal{C}$. For every $A \in \mathcal{A}$, $L(A)$ is represented by an $m \times m$ complex matrix given by $L(A)_{ij} = \langle C_i, AC_j \rangle$, for every $i, j \in [m]$. The map $L : \mathcal{A} \to \mathbb{C}^{m\times m}$, is called the *regular $*$-representation* of $\mathcal{A}$ associated with the orthonormal basis $\mathcal{C}$. Since L is a linear map, it is completely specified by its image for the elements of the basis $\mathcal{C}$. Let $(p^t_{rs})_{r,s,t\in[m]}$ be the *multiplication parameters* of $\mathcal{A}$ with respect to the basis $\mathcal{C}$ defined by $C_rC_s = \sum_{t=1}^{m} p^t_{rs} C_t$. Then, $L(C_r)_{ij} = p^i_{rj}$, for every $r \in [m]$.

Theorem 5.3.2 ([KPS07]). *Let $\mathcal{L}$ be the matrix $*$-algebra generated by the matrices $L(C_1), \ldots, L(C_m)$. Then the map ψ defined as*

$$\psi : \mathcal{A} \to \mathcal{L} \ , \ \psi(C_r) = L(C_r) \ , \ r \in [m] \ , \tag{5.5}$$

is a $*$-isomorphism.

Note that under the $*$-isomorphism ψ of Theorem 5.3.2, for $\mathcal{A} \subseteq \mathbb{C}^{n\times n}$, the matrix dimensions are reduced from $n\times n$ to $m\times m$, whereas the $*$-isomorphism ϕ of Theorem 5.3.1 provides a fine block-diagonal decomposition into t blocks where the block matrix i is of size $m_i \times m_i$, satisfying $m = m_1^2 + \ldots + m_t^2$.

5.3.2 Representative set of group action

Let G be a finite group acting on a finite dimensional complex vector space $\mathcal{H}$. As we have seen in Section 2.5, the space $\mathcal{H}$ can be decomposed as $\mathcal{H} = \mathcal{H}_1 \oplus \cdots \oplus \mathcal{H}_t$ such that each $\mathcal{H}_i$ is a direct sum $\mathcal{H}_{i,1} \oplus \cdots \oplus \mathcal{H}_{i,m_i}$ of irreducible G-modules with the property that $\mathcal{H}_{i,j} \cong \mathcal{H}_{i',j'}$ if and only if $i = i'$.

For each $i \in [t]$ and $j \in [m_i]$, let $u_{i,j} \in \mathcal{H}_{i,j}$ be a nonzero vector such that for each i and all $j, j' \in [m_i]$, there is a bijective G-equivariant map from $\mathcal{H}_{i,j}$ to $\mathcal{H}_{i,j'}$ that maps $u_{i,j}$ to $u_{i,j'}$. For $i \in [t]$, we define a matrix U_i as $[u_{i,1}, \ldots, u_{i,m_i}]$, with $u_{i,j}$ forming the j-th column of U_i. The matrix set $\{U_1, \ldots, U_t\}$ obtained in this way is called a *representative* for the action of G on $\mathcal{H}$. The columns of the matrices U_i can be viewed as elements of the dual space $\mathcal{H}^*$ (by taking the standard inner product). Then each U_i is an ordered set of linear functions on $\mathcal{H}$.

Since $\mathcal{H}_{i,j}$ is the linear space spanned by $G \cdot u_{i,j}$ (for each i, j), we have

$$\mathcal{H} = \bigoplus_{i=1}^{t} \bigoplus_{j=1}^{m_i} \mathbb{C}G \cdot u_{i,j} \, ,$$

where $\mathbb{C}G = \left\{ \sum_{g\in G} \alpha_g g : \alpha_g \in \mathbb{C} \right\}$ denotes the complex group algebra of G. Moreover, by Schur's lemma, one has

$$\dim \mathrm{End}^G(\mathcal{H}) = \dim \mathrm{End}^G \left(\bigoplus_{i=1}^{t} \bigoplus_{j=1}^{m_i} \mathcal{H}_{i,j} \right) = \sum_{i=1}^{t} m_i^2 \, . \tag{5.6}$$

Remark 5.3.3. It is straightforward to see that $\mathrm{End}^G(\mathcal{H})$ corresponds to the subset of G-invariant matrices and has the structure of a matrix $*$-algebra. For $\mathcal{A} = \mathrm{End}^G(\mathcal{H})$, the structural parameters of Theorem 5.3.1 have a representation theoretic interpretation. In particular, the number of the direct summands t corresponds to the number of isomorphism classes of irreducible G-submodules and m_i is the multiplicity of the irreducible G-submodules in class i.

Note that with the action of the finite group G on the space $\mathcal{H}$, any inner product $\langle \, , \rangle$ on $\mathcal{H}$ gives rise to a G-invariant inner product $\langle \, , \rangle_G$ on $\mathcal{H}$ via the rule $\langle x, y\rangle_G := \frac{1}{|G|} \sum_{g\in G} \langle g \cdot x, g \cdot y\rangle$. Let $\langle \, , \rangle$ be a G-invariant inner product on $\mathcal{H}$ and $\{U_1, \ldots, U_t\}$ be a representative

for the action of G on $\mathcal{H}$. Consider the linear map $\phi : \mathrm{End}^G(\mathcal{H}) \to \bigoplus_{i=1}^t \mathbb{C}^{m_i \times m_i}$ defined as

$$\phi(A) \coloneqq \bigoplus_{i=1}^{t} \left(\langle A u_{i,j'}, u_{i,j}\rangle\right)_{j,j'=1}^{m_i} \ , \ \forall A \in \mathrm{End}^G(\mathcal{H}) \ . \tag{5.7}$$

For $i \in [t]$ and $A \in \mathrm{End}^G(\mathcal{H})$, we denote the matrix $\left(\langle A u_{i,j'}, u_{i,j}\rangle\right)_{j,j'=1}^{m_i}$ corresponding to the i-th block of $\phi(A)$ by $[\![\phi(A)]\!]_i$.

Lemma 5.3.4 (Proposition 2.4.4, [Pol19])**.** *The linear map ϕ of Eq. (5.7) is bijective and for every $A \in \mathrm{End}^G(\mathcal{H})$, we have $A \geq 0$ if and only if $\phi(A) \geq 0$. Moreover, there is a unitary matrix U such that*

$$U^* A U = \bigoplus_{i=1}^{t} \bigoplus_{j=1}^{d_i} [\![\phi(A)]\!]_i \ , \ \forall A \in \mathrm{End}^G(\mathcal{H}) \ ,$$

where $d_i = \dim(\mathcal{H}_{i,1})$, for every $i \in [t]$.

Lemma 5.3.4 plays a very important role in our symmetry reductions. Note that $\dim(\mathrm{End}^G(\mathcal{H})) = \sum_{i=1}^t m_i^2$ can be significantly smaller than $\dim \mathcal{H}$. Moreover, by this lemma, for any $A \in \mathrm{End}^G(\mathcal{H})$, the task of checking whether A is a positive semidefinite matrix can be reduced to checking if the smaller $m_i \times m_i$ matrices $[\![\phi(A)]\!]_i$ are positive semidefinite, for every $i \in [t]$. The mapping ϕ in Eq. (5.7) is a special case of the $*$-isomorphism of Theorem 5.3.1, where $\mathcal{A}$ is the matrix $*$-algebra $\mathrm{End}^G(\mathcal{H})$.

5.3.3 Representation theory of the symmetric group

Fix $k \in \mathbb{N}$ and a finite-dimensional vector space $\mathcal{H}$ with $\dim(\mathcal{H}) = d$. We consider the natural action of the symmetric group $\mathfrak{S}_k$ on $\mathcal{H}^{\otimes k}$ by permuting the indices, i.e.,

$$\pi \cdot (h_1 \otimes \cdots \otimes h_k) = h_{\pi^{-1}(1)} \otimes \cdots \otimes h_{\pi^{-1}(k)} \, , h_i \in \mathcal{H} \, , \forall \pi \in \mathfrak{S}_k \, .$$

Based on classical representation theory of the symmetric group, we describe a representative set for the action of $\mathfrak{S}_k$ on $\mathcal{H}^{\otimes k}$. The concepts and notation we introduce in this section will be used throughout this chapter.

A *partiton* λ of k is a sequence $(\lambda_1, \ldots, \lambda_d)$ of natural numbers with $\lambda_1 \geq \ldots \lambda_d > 0$ and $\lambda_1 + \cdots + \lambda_d = k$. The number d is called the *height* of λ. We write $\lambda \vdash_d k$ if λ is a partition of k with height d. Let $\mathrm{Par}(d, k) \coloneqq \{\lambda : \lambda \vdash_d k\}$. The *Young shape* $Y(\lambda)$ of λ is the set

$$Y(\lambda) \coloneqq \{(i, j) \in \mathbb{N}^2 : 1 \leq j \leq d, 1 \leq i \leq \lambda_j\} \, .$$

Following the French notation [Pro07], for an index $j_0 \in [d]$, the j_0-th *row* of $Y(\lambda)$ is set of elements (i, j_0) in $Y(\lambda)$. Similarly, fixing an element $i_0 \in [\lambda_1]$, the i_0-th *column* of $Y(\lambda)$ is set of elements (i_0, j) in $Y(\lambda)$. We label the elements in $Y(\lambda)$ from 1 to k according the lexicographic order on their positions. Then the *row stabilizer* R_λ of λ is the group of permutations π of $Y(\lambda)$ with $\pi(L) = L$ for each row L of $Y(\lambda)$. Similarly, the *column*

stabilizer C_λ of λ is the group of permutations π of $Y(\lambda)$ with $\pi(L) = L$ for each column L of $Y(\lambda)$.

For $\lambda \vdash_d k$, a λ-*tableau* is a function $\tau : Y(\lambda) \to \mathbb{N}$. A λ-tableau is *semistandard* if the entries are non-decreasing in each row and strictly increasing in each column. Let $T_{\lambda,d}$ be the collection of semistandard λ-tableaux with entries in $[d]$. We write $\tau \sim \tau'$ for λ-tableaux τ, τ' if $\tau' = \tau r$ for some $r \in R_\lambda$. Let $e_1, \ldots, e_d$ be the standard basis of $\mathcal{H}$. For any $\tau \in T_{\lambda,d}$, define $u_\tau \in \mathcal{H}^{\otimes k}$ as

$$u_\tau := \sum_{\tau' \sim \tau} \sum_{c \in C_\lambda} \operatorname{sgn}(c) \bigotimes_{y \in Y(\lambda)} e_{\tau'(c(y))} \,. \tag{5.8}$$

Here the Young shape $Y(\lambda)$ is ordered by concatenating its rows. Then the matrix set

$$\{U_\lambda : \lambda \vdash_d k\} \quad \text{with } U_\lambda = [u_\tau : \tau \in T_{\lambda,d}] \tag{5.9}$$

is a representative for the natural action of $\mathfrak{S}_k$ on $\mathcal{H}^{\otimes k}$ [LPS17, Section 2.1]. Moreover, we have

$$|\mathrm{Par}(d,k)| \leq (k+1)^d \text{ and } |T_{\lambda,d}| \leq (k+1)^{d(d-1)/2} \,, \forall \lambda \in \mathrm{Par}(d,k) \ . \tag{5.10}$$

5.4 Efficient approximation of the regularized divergence of channels

For $\alpha \in (1, \infty)$, the regularized sandwiched α-Rényi divergence between channels $\mathcal{N}_{X \to Y}$ and $\mathcal{M}_{X \to Y}$ is defined as

$$\widetilde{\mathrm{D}}^{\mathrm{reg}}_\alpha(\mathcal{N} \| \mathcal{M}) := \lim_{k \to \infty} \frac{1}{k} \widetilde{\mathrm{D}}_\alpha(\mathcal{N}^{\otimes k} \| \mathcal{M}^{\otimes k}) \ . \tag{5.11}$$

The regularized sandwiched Rényi divergence between channels can be used to obtain improved characterization of many information processing tasks such as channel discrimination [FF21a, FF21b]. However, the sandwiched Rényi divergence between channels is non-additive in general [FFRS20] and it is unclear whether its regularization can be computed efficiently. Ref. [FF21b] provides a converging hierarchy of upper bounds on the regularized divergence between channels:

Theorem 5.4.1 ([FF21b])**.** *Let $\alpha \in (1, \infty)$ and $\mathcal{N}, \mathcal{M}$ be completely positive maps from $\mathscr{L}(X)$ to $\mathscr{L}(Y)$. Then for any $k \geq 1$,*

$$\frac{1}{k} \mathrm{D}^{\#}_\alpha(\mathcal{N}^{\otimes k} \| \mathcal{M}^{\otimes k}) - \frac{1}{k} \frac{\alpha}{\alpha - 1} (d^2 + d) \log(k + d) \leq \widetilde{\mathrm{D}}^{\mathrm{reg}}_\alpha(\mathcal{N} \| \mathcal{M}) \leq \frac{1}{k} \mathrm{D}^{\#}_\alpha(\mathcal{N}^{\otimes k} \| \mathcal{M}^{\otimes k}) \ ,$$

where $d = \dim X \dim Y$.

We note that $\frac{1}{k} \mathrm{D}^{\#}_\alpha(\mathcal{N}^{\otimes k} \| \mathcal{M}^{\otimes k})$ is decreasing in k (since the $\mathrm{D}^{\#}$ channel divergence is subadditive). Moreover, $\mathrm{D}^{\#}_\alpha(\mathcal{N}^{\otimes k} \| \mathcal{M}^{\otimes k})$ can be written in terms of a convex program as ([FF21b])

$$\frac{1}{\alpha - 1} \log \min_{A_{X^{\otimes k} Y^{\otimes k}} \geq 0} \| \mathrm{tr}_{Y^{\otimes k}}(A_{X^{\otimes k} Y^{\otimes k}}) \|_\infty \quad \text{s.t.} \quad (J^{\mathcal{N}^{\otimes k}}) \leq (J^{\mathcal{M}^{\otimes k}}) \#_{1/\alpha} A_{X^{\otimes k} Y^{\otimes k}} . \tag{5.12}$$

Therefore, Theorem 5.4.1 establishes that $\widetilde{\mathrm{D}}_\alpha^{\mathrm{reg}}(\mathcal{N}\|\mathcal{M})$ can be approximated by $\frac{1}{k}\mathrm{D}_\alpha^{\#}(\mathcal{N}^{\otimes k}\|\mathcal{M}^{\otimes k})$ with arbitrary accuracy for sufficiently large k in finite time. Namely, if we take $k = \lceil \frac{8\alpha d^3}{(\alpha-1)\epsilon} \rceil$ then we have

$$|\widetilde{\mathrm{D}}_\alpha^{\mathrm{reg}}(\mathcal{N}\|\mathcal{M}) - \frac{1}{k}\mathrm{D}_\alpha^{\#}(\mathcal{N}^{\otimes k}\|\mathcal{M}^{\otimes k})| \leq \epsilon \,.$$

However, the size of Program (5.12) grows exponentially with k.

5.4.1 Exploiting symmetries to simplify the problem

In this section, we will show how the symmetries of Program (5.12) can be used to simplify this optimization problem and solve it in time polynomial in k. We first focus on the natural symmetries arising due to invariance under permutation of physical systems. In Section 5.4.2, we show how additional symmetries can be utilized to further simplify the problem. Our approach can be summarized as follows: First, we show that program (5.12) is invariant with respect to the action of the symmetric group. Using this observation, we show that the program can be transformed into an equivalent program with polynomially many constraints, each of polynomial size in k. In order to show this, we use the block-diagonal decomposition given by Lemma 5.3.4. A naive implementation of this transformation, however, involves exponential time computations. We show that the simplified form of the program can be directly computed in poly(k) time.

Recall that, for every $\pi \in \mathfrak{S}_k$, we consider the action of π on k copies of a finite dimensional Hilbert space $\mathcal{H}$ as

$$\pi \cdot (h_1 \otimes \cdots \otimes h_k) = h_{\pi^{-1}(1)} \otimes \cdots \otimes h_{\pi^{-1}(k)} \quad , \quad h_i \in \mathcal{H}\,, \forall i \in [k] \ .$$

Let $P_X(\pi)$ and $P_Y(\pi)$ be the permutation matrices corresponding to the action of π on $X^{\otimes k}$ and $Y^{\otimes k}$, respectively. Note that the action of π on $(X \otimes Y)^{\otimes k}$ corresponds to the simultaneous permutation of the X and Y tensor factors and the corresponding permutation matrix, when the subsystems are reordered as $X^{\otimes k} \otimes Y^{\otimes k}$, is given by $P_{X\otimes Y}(\pi) = P_X(\pi) \otimes P_Y(\pi)$.

The following lemma shows that the feasible region of the convex program (5.12) may be restricted to the permutation invariant algebra of operators on $X^{\otimes k} \otimes Y^{\otimes k}$ without changing the optimal value.

For a linear operator $X \in \mathscr{L}(\mathcal{H}^{\otimes k})$, we define its *group average* operator denoted $\overline{X}$ as

$$\overline{X} := \frac{1}{|\mathfrak{S}_k|} \sum_{\pi \in \mathfrak{S}_k} P_{\mathcal{H}}(\pi) X P_{\mathcal{H}}(\pi)^* \ .$$

Lemma 5.4.2. *The convex program of Eq.* (5.12) *has an optimal solution* $A \in \mathrm{End}^{\mathfrak{S}_k}\left(X^{\otimes k} \otimes Y^{\otimes k}\right)$.

Proof. It is straightforward to check that by Slater's condition the optimal value is achieved by a feasible solution. We will prove that for every feasible solution A, the corresponding group-average operator $\overline{A}$ is a feasible solution with an objective value not greater than the original value.

To simplify the notation, let $\Pi(\pi) \coloneqq P_{X\otimes Y}(\pi)$. We have

$$\begin{aligned}
\overline{A} \,\#_{1/\alpha}\, J^{\mathcal{M}^{\otimes k}} &= \left(\frac{1}{|\mathfrak{S}_k|}\sum_{\pi\in\mathfrak{S}_k}\Pi(\pi)A\Pi(\pi)^*\right)\#_{1/\alpha}\left(\frac{1}{|\mathfrak{S}_k|}\sum_{\pi\in\mathfrak{S}_k}\Pi(\pi)J^{\mathcal{M}^{\otimes k}}\Pi(\pi)^*\right) && (5.13)\\
&\geq \sum_{\pi\in\mathfrak{S}_k}\left(\frac{1}{|\mathfrak{S}_k|}\Pi(\pi)A\Pi(\pi)^*\right)\#_{1/\alpha}\left(\frac{1}{|\mathfrak{S}_k|}\Pi(\pi)J^{\mathcal{M}^{\otimes k}}\Pi(\pi)^*\right) && (5.14)\\
&= \frac{1}{|\mathfrak{S}_k|}\sum_{\pi\in\mathfrak{S}_k}\Pi(\pi)\left(A\#_{1/\alpha}J^{\mathcal{M}^{\otimes k}}\right)\Pi(\pi)^* && (5.15)\\
&\geq \frac{1}{|\mathfrak{S}_k|}\sum_{\pi\in\mathfrak{S}_k}\Pi(\pi)\left(J^{\mathcal{N}^{\otimes k}}\right)\Pi(\pi)^* && (5.16)\\
&= J^{\mathcal{N}^{\otimes k}}\ , && (5.17)
\end{aligned}$$

where Eq. (5.13) holds since $J^{\mathcal{M}^{\otimes k}} \in \mathrm{End}^{\mathfrak{S}_k}\left(X^{\otimes k}\otimes Y^{\otimes k}\right)$, inequality (5.14) follows from the joint-concavity property of the geometric mean, Eq. (5.15) is a consequence of properties 2 and 3 of the geometric mean, inequality (5.16) holds by feasibility of A, and finally, Eq. (5.17) follows since $J^{\mathcal{N}^{\otimes k}} \in \mathrm{End}^{\mathfrak{S}_k}\left(X^{\otimes k}\otimes Y^{\otimes k}\right)$.

For the objective function, note that since $\Pi(\pi) = P_X(\pi)\otimes P_Y(\pi)$, we have

$$\mathrm{tr}_{Y^{\otimes k}}\left(\Pi(\pi)A\Pi(\pi)^T\right) = P_X(\pi)\,\mathrm{tr}_{Y^{\otimes k}}(A)\,P_X(\pi)^T\ .$$

Therefore, by the triangle inequality and the unitary invariance of the operator norm, we have

$$\begin{aligned}
\|\mathrm{tr}_{Y^{\otimes k}}\left(\overline{A}\right)\|_\infty &= \left\|\mathrm{tr}_{Y^{\otimes k}}\left(\frac{1}{|\mathfrak{S}_k|}\sum_{\pi\in\mathfrak{S}_k}\Pi(\pi)A\Pi(\pi)^T\right)\right\|_\infty\\
&= \left\|\frac{1}{|\mathfrak{S}_k|}\sum_{\pi\in\mathfrak{S}_k}P(\pi_X)\mathrm{tr}_{Y^{\otimes k}}(A)P(\pi_X)^T\right\|_\infty\\
&\leq \frac{1}{|\mathfrak{S}_k|}\sum_{\pi\in\mathfrak{S}_k}\left\|P(\pi_X)\mathrm{tr}_{Y^{\otimes k}}(A)P(\pi_X)^T\right\|_\infty\\
&= \|\mathrm{tr}_{Y^{\otimes k}}(A)\|_\infty.
\end{aligned}$$

This concludes the proof. □

Recall that in the convex program (5.12), the number of the variables and the size of the PSD constraints grow exponentially with k. Using the observation made in Lemma 5.4.2, we show that this optimization problem can be transformed into a form having a number of variables and constraints that is polynomial in k. Before doing so, we introduce some notation.

Let $\mathcal{H} \in \{X, Y, X\otimes Y\}$ and $d_{\mathcal{H}} \coloneqq \dim\mathcal{H}$. The algebra of $\mathfrak{S}_k$-invariant operators on $\mathcal{H}^{\otimes k}$ is given by

$$\mathrm{End}^{\mathfrak{S}_k}\left(\mathcal{H}^{\otimes k}\right) = \left\{A\in\mathscr{L}(\mathcal{H}^{\otimes k}) : P_{\mathcal{H}}(\pi)\,A\,P_{\mathcal{H}}(\pi)^* = A,\ \forall\pi\in\mathfrak{S}_k\right\}.$$

Let $\phi_{\mathcal{H}}$ denote the linear map defined in Eq. (5.7) that maps the elements of $\mathrm{End}^{\mathfrak{S}_k}\left(\mathcal{H}^{\otimes k}\right)$ into block-diagonal form:

$$\begin{aligned}\phi_{\mathcal{H}} : \mathrm{End}^{\mathfrak{S}_k}\left(\mathcal{H}^{\otimes k}\right) &\to \bigoplus_{\lambda\in\mathrm{Par}(d_{\mathcal{H}},k)} \mathbb{C}^{m_\lambda^{\mathcal{H}}\times m_\lambda^{\mathcal{H}}} \\ A &\mapsto \bigoplus_{\lambda\in\mathrm{Par}(d_{\mathcal{H}},k)} \left(\langle Au_\gamma, u_\tau\rangle\right)_{\tau,\gamma\in T_{\lambda,d_{\mathcal{H}}}} \ . \end{aligned} \tag{5.18}$$

In this decomposition, the number of blocks and the size of the blocks are bounded by a polynomial in k. In particular, we have

$$t^{\mathcal{H}} := |\mathrm{Par}(d_{\mathcal{H}},k)| \le (k+1)^{d_{\mathcal{H}}} \ , \tag{5.19}$$

$$m_\lambda^{\mathcal{H}} := |T_{\lambda,d_{\mathcal{H}}}| \le (k+1)^{d_{\mathcal{H}}(d_{\mathcal{H}}-1)/2} \ , \quad \forall\lambda\in\mathrm{Par}(d_{\mathcal{H}},k) \ . \tag{5.20}$$

From Eqs. (5.19) and (5.20), we get

$$m^{\mathcal{H}} := \dim\left[\mathrm{End}^{\mathfrak{S}_k}(\mathcal{H}^{\otimes k})\right] \le (k+1)^{d_{\mathcal{H}}^2} \tag{5.21}$$

Theorem 5.4.3. *The channel divergence* $\mathrm{D}_\alpha^{\#}(\mathcal{N}^{\otimes k}\|\mathcal{M}^{\otimes k})$ *can be formulated as a convex program with* $\mathrm{O}\left(k^{d^2}\right)$ *variables and* $\mathrm{O}\left(k^d\right)$ *PSD constraints involving matrices of size at most* $(k+1)^{d(d-1)/2}$*, where* $d = d_X d_Y$*.*

Proof. By Lemma 5.4.2 and Property 2 of the α-geometric mean, after a permutation of the X and Y tensor factors, the formulation (5.12) for $\mathrm{D}_\alpha^{\#}(\mathcal{N}^{\otimes k}\|\mathcal{M}^{\otimes k})$ can be written as

$$\frac{1}{\alpha-1}\log \quad \min_{A,y} \quad y \tag{5.22}$$

$$\text{s.t.} \quad \mathrm{tr}_{Y^{\otimes k}}(A) \le y\,\mathrm{id}_{X^{\otimes k}} \ , \tag{5.23}$$

$$\left(J^{\mathcal{N}}\right)^{\otimes k} \le \left(J^{\mathcal{M}}\right)^{\otimes k} \#_{1/\alpha} A \ , \tag{5.24}$$

where $A \in \mathscr{P}\left((X\otimes Y)^{\otimes k}\right) \cap \mathrm{End}^{\mathfrak{S}_k}\left((X\otimes Y)^{\otimes k}\right)$ and $y\in\mathbb{R}$.

For $\mathcal{H}\in\{X, X\otimes Y\}$, let $\phi_{\mathcal{H}} : \mathrm{End}^{\mathfrak{S}_k}(\mathcal{H}^{\otimes k}) \to \bigoplus_{i=1}^{t^{\mathcal{H}}} \mathbb{C}^{m_i^{\mathcal{H}}\times m_i^{\mathcal{H}}}$ be the bijective linear map defined in Eq. (5.18) which block-diagonalizes the corresponding invariant algebra, where to simplify the notation, the blocks are indexed by $i\in[t^{\mathcal{H}}]$ instead of $\lambda\in\mathrm{Par}(d_{\mathcal{H}},k)$. For $Z\in\mathrm{End}^{\mathfrak{S}_k}(\mathcal{H}^{\otimes k})$, we denote the i-th block of $\phi_{\mathcal{H}}(Z)$ by $[\![\phi_{\mathcal{H}}(Z)]\!]_i$. Note that by Lemma 5.3.4, $\phi_{\mathcal{H}}$ preserves positive semidefiniteness. Therefore, since $\mathrm{tr}_{Y^{\otimes k}}(A)$, $\mathrm{id}_{X^{\otimes k}} \in \mathrm{End}^{\mathfrak{S}_k}(X^{\otimes k})$, the constraint (5.23) can be mapped by ϕ_X into the direct sum form. By Lemma 5.3.4 and Property 2 of the α-geometric mean, we have $\phi_{X\otimes Y}\left(\left(J^{\mathcal{M}}\right)^{\otimes k}\#_{1/\alpha}A\right) = \phi_{X\otimes Y}\left(\left(J^{\mathcal{M}}\right)^{\otimes k}\right)\#_{1/\alpha}\phi_{X\otimes Y}(A)$. Therefore, by Property 5 of the α-geometric mean (direct sum property), the constraint (5.24) can be decomposed into constraints involving the smaller diagonal blocks as well. The transformed convex program can be written as

$$\begin{aligned}
&\frac{1}{\alpha-1}\log \quad \min \quad y \\
&\qquad \text{s.t.} \quad \left[\!\!\left[\left(\phi_X\circ\mathrm{tr}_{Y^{\otimes k}}\circ\phi_{X\otimes Y}^{-1}\right)(\oplus_l A_l)\right]\!\!\right]_j \le y\,\mathrm{id}_{m_j^X} \ , && j\in\left[t^X\right] \\
&\qquad\qquad \left[\!\!\left[\phi_{X\otimes Y}\left(\left(J^{\mathcal{N}}\right)^{\otimes k}\right)\right]\!\!\right]_i \le \left[\!\!\left[\phi_{X\otimes Y}\left(\left(J^{\mathcal{M}}\right)^{\otimes k}\right)\right]\!\!\right]_i \#_{1/\alpha} A_i \ , && i\in\left[t^{X\otimes Y}\right] \\
&\qquad\qquad A_i \in \mathscr{P}\left(\mathbb{C}^{m_i^{X\otimes Y}}\right), && i\in\left[t^{X\otimes Y}\right]
\end{aligned}$$

The statement of the theorem follows since for $\mathcal{H} \in \{X, X \otimes Y\}$, by Eq. (5.19), we have $t^{\mathcal{H}} \leq (k+1)^{d_{\mathcal{H}}}$ and by Eq. (5.20), for every $i \in \left[t^{\mathcal{H}}\right]$, we have $m_i^{\mathcal{H}} \leq (k+1)^{d_{\mathcal{H}}(d_{\mathcal{H}}-1)/2}$. $\square$

Note that a direct implementation of the transformation mapping the convex program (5.12) into the polynomial-size form of Theorem 5.4.3 involves exponential computations. Next, we show how to do this efficiently.

A basis for the invariant subspace. The canonical basis of the matrix $*$-algebra $\mathrm{End}^{\mathfrak{S}_k}\left(\mathcal{H}^{\otimes k}\right)$ consists of zero-one incidence matrices of orbits of the group action on pairs (see [KPS07, LM11] for more information). In particular, let the standard basis of $\mathcal{H}^{\otimes k}$ be indexed by $i \in \left[(d_{\mathcal{H}})^k\right]$. Then the orbit of the pair $(i,j) \in \left[(d_{\mathcal{H}})^k\right]^2$ under the action of the group $\mathfrak{S}_k$ is given by

$$O(i,j) = \{(\pi(i), \pi(j)) : \pi \in \mathfrak{S}_k\},$$

where $\pi(i)$ is the index of the basis vector $P_{\mathcal{H}}(\pi)|i\rangle$. With this notation, for every $A \in \mathrm{End}^{\mathfrak{S}_k}\left(\mathcal{H}^{\otimes k}\right)$, and every $\pi \in \mathfrak{S}_k$, we have $A_{ij} = A_{\pi^{-1}(ij)} = A_{\pi^{-1}(i),\pi^{-1}(j)}$. The set $\left[(d_{\mathcal{H}})^k\right]^2$ decompose into orbits $O_1^{\mathcal{H}}, \ldots, O_{m^{\mathcal{H}}}^{\mathcal{H}}$ under the action of $\mathfrak{S}_k$. For each $r \in [m^{\mathcal{H}}]$, we construct a zero-one matrix $C_r^{\mathcal{H}}$ of size $(d_{\mathcal{H}})^k \times (d_{\mathcal{H}})^k$ given by

$$(C_r^{\mathcal{H}})_{ij} = \begin{cases} 1 & \text{if } (i,j) \in O_r^{\mathcal{H}}\,, \\ 0 & \text{otherwise.} \end{cases} \tag{5.25}$$

The set $\mathcal{C}^{\mathcal{H}} = \{C_1^{\mathcal{H}}, \ldots, C_{m^{\mathcal{H}}}^{\mathcal{H}}\}$ forms an orthogonal basis of $\mathrm{End}^{\mathfrak{S}_k}\left(\mathcal{H}^{\otimes k}\right)$ with $m^{\mathcal{H}} \leq (k+1)^{d_{\mathcal{H}}^2}$.

Enumerating all orbits. For each $r = 1, \ldots, m^{\mathcal{H}}$, we need to compute a representative element of $O_r^{\mathcal{H}}$. In order to do so, we define a matrix $E^{(i,j)} \in \mathbb{Z}_{\geq 0}^{d_{\mathcal{H}} \times d_{\mathcal{H}}}$

$$(E^{(i,j)})_{a,b} \coloneqq |\{v \in [k] : i_v = a, j_v = b\}|\ ,\ \forall a, b \in [d_{\mathcal{H}}]\,. \tag{5.26}$$

By the construction in Eq. (5.26), for two pairs $(i,j), (i',j') \in [d_{\mathcal{H}}]^k \times [d_{\mathcal{H}}]^k$, we have $(i',j') = (\pi(i), \pi(j))$, for some $\pi \in \mathfrak{S}_k$ if and only if $E^{(i,j)} = E^{(i',j')}$. Therefore, there is a one-to-one correspondence between the orbits $\left\{O_r^{\mathcal{H}}\right\}_{r \in [m^{\mathcal{H}}]}$ and $E \in \mathbb{Z}_{\geq 0}^{d_{\mathcal{H}} \times d_{\mathcal{H}}}$ such that $\sum_{a,b} E_{a,b} = k$. Therefore, we can determine a representative element for every $O_r^{\mathcal{H}}$ in $\mathrm{poly}(k)$ time by listing all non-negative integer solutions of the equation $\sum_{a,b \in [d_{\mathcal{H}}]} E_{a,b} = k$.

Any matrix in $\mathrm{End}^{\mathfrak{S}_k}(\mathcal{H}^{\otimes k})$ can be written in the basis $\mathcal{C}^{\mathcal{H}}$ as

$$M(z) \coloneqq \sum_{r=1}^{m^{\mathcal{H}}} z_r C_r^{\mathcal{H}}\,,\ \text{for some } z \in \mathbb{C}^{[m^{\mathcal{H}}]}\,. \tag{5.27}$$

Using the representative matrix for the action of $\mathfrak{S}_k$ on the space $\mathcal{H}^{\otimes k}$ in Eq. (5.9), we get

$$\phi_{\mathcal{H}}(M(z)) = \sum_{r=1}^{m^{\mathcal{H}}} z_r \phi_{\mathcal{H}}\left(C_r^{\mathcal{H}}\right) = \sum_{r=1}^{m^{\mathcal{H}}} z_r \bigoplus_{\lambda \vdash_{d_{\mathcal{H}}} k} U_\lambda^T C_r^{\mathcal{H}} U_\lambda\,. \tag{5.28}$$

Note that U_λ is real matrix for all $\lambda \in \mathrm{Par}(d_{\mathcal{H}}, k)$.

We show that, for every $r \in [m^{\mathcal{H}}]$, $\phi_{\mathcal{H}}(C_r^{\mathcal{H}})$ can be computed in $\mathrm{poly}(k)$ time. In order to do so, we show how to efficiently compute each block $U_\lambda^T C_r^{\mathcal{H}} U_\lambda$ indexed by $\lambda \in \mathrm{Par}(d_{\mathcal{H}}, k)$. This in fact boils down to efficiently computing $u_\tau^T C_r^{\mathcal{H}} u_\gamma$, for every $\tau, \gamma \in T_{\lambda, d_{\mathcal{H}}}$. We note that u_τ, u_γ, and $C_r^{\mathcal{H}}$ all have exponential size in k.

For $\mathcal{H} \in \{X, Y, X \otimes Y\}$, let $W_{\mathcal{H}} \coloneqq \mathcal{H} \otimes \mathcal{H}$. For every $p = (i, j) \in [d_{\mathcal{H}}]^2$, define

$$a_p \coloneqq e_i \otimes e_j \in W_{\mathcal{H}} \,,$$

where $\{e_i\}_{i \in [d_{\mathcal{H}}]}$ is the standard basis of $\mathcal{H}$. Then the set $\mathcal{B} \coloneqq \left\{a_p : p \in [d_{\mathcal{H}}]^2\right\}$ is a basis of $W_{\mathcal{H}}$. Let $\mathcal{B}^* \coloneqq \left\{a_p^* : p \in [d_{\mathcal{H}}]^2\right\}$ be the corresponding dual basis for $W_{\mathcal{H}}^*$.

Using the natural identification of $\left[(d_{\mathcal{H}})^k\right]^2$ and $([d_{\mathcal{H}}]^2)^k$, for every $r \in [m^{\mathcal{H}}]$, we map $O_r^{\mathcal{H}} \subseteq \left[(d_{\mathcal{H}})^k\right]^2$ to $\mathsf{O}_r^{\mathcal{H}} \subseteq ([d_{\mathcal{H}}]^2)^k$. Then corresponding to each operator $C_r^{\mathcal{H}}$, we define

$$\mathsf{C}_r^{\mathcal{H}} \coloneqq \sum_{(p_1, \ldots, p_k) \in \mathsf{O}_r^{\mathcal{H}}} a_{p_1} \otimes \cdots \otimes a_{p_k} \in W_{\mathcal{H}}^{\otimes k} \,. \tag{5.29}$$

Note that $\mathsf{C}_r^{\mathcal{H}}$ can be obtained from $\mathrm{vec}\left(C_r^{\mathcal{H}}\right)$ by applying the permutation operator which maps $\left(\mathcal{H}^{\otimes k}\right)^{\otimes 2}$ to $\left(\mathcal{H}^{\otimes 2}\right)^{\otimes k}$. For every $(p_1, \ldots, p_k) \in \left[(d_{\mathcal{H}})^2\right]^k$, let

$$m(p_1, \ldots, p_k) \coloneqq a_{p_1}^* \cdots a_{p_k}^* \in \mathcal{O}_k(W_{\mathcal{H}}) \tag{5.30}$$

be a degree k monomial expressed in the basis $\mathcal{B}^*$. Note that, for a fixed $r \in [m^{\mathcal{H}}]$, $m(p_1, \ldots, p_k)$ is the same monomial, for every $(p_1, \ldots, p_k) \in \mathsf{O}_r^{\mathcal{H}}$. We denote this monomial by $m\left(\mathsf{O}_r^{\mathcal{H}}\right)$. Moreover, $\left\{\mathsf{O}_r^{\mathcal{H}}\right\}_{r \in [m^{\mathcal{H}}]}$ partitions $\left[(d_{\mathcal{H}})^2\right]^k$ into disjoint subsets. Therefore, there is a bijection between $\left\{\mathsf{O}_i^{\mathcal{H}}\right\}_{i \in [m^{\mathcal{H}}]}$ and the set of degree k monomials expressed in the basis $\mathcal{B}^*$.

Let $\zeta : (W_{\mathcal{H}}^*)^{\otimes k} \to \mathcal{O}_k(W_{\mathcal{H}})$ be the linear map defined as

$$\zeta(w_1^* \otimes \cdots \otimes w_k^*) \coloneqq w_1^* \cdots w_k^* \,, \ \forall w_1^*, \ldots, w_k^* \in W_{\mathcal{H}}^* \,.$$

To simplify the notation we write $\overline{w} = \zeta(w)$, for every $w \in (W_{\mathcal{H}}^*)^{\otimes k}$.

For every $\lambda \in \mathrm{Par}(d_{\mathcal{H}}, k)$ and $\tau, \gamma \in T_{\lambda, d_{\mathcal{H}}}$, define the polynomial $f_{\tau,\gamma} \in \mathbb{C}[x_{i,j} : i, j \in [d_{\mathcal{H}}]]$ by

$$f_{\tau,\gamma}(X) \coloneqq \sum_{\substack{\tau' \sim \tau \\ \gamma' \sim \gamma}} \sum_{c, c' \in C_\lambda} \mathrm{sgn}(cc') \prod_{y \in Y(\lambda)} x_{\tau' c(y), \gamma' c'(y)} \,, \tag{5.31}$$

for $X = (x_{i,j})_{i,j=1}^{d_{\mathcal{H}}} \in \mathbb{C}^{d_{\mathcal{H}} \times d_{\mathcal{H}}}$. Refs. [LPS17, Proposition 3] and [Gij09, Theorem 7] show that the polynomial in Eq. (5.31) can be computed (i.e., expressed as a linear combination of monomials in variables $x_{i,j}$) in polynomial time.

Lemma 5.4.4. *For every $\lambda \in \mathrm{Par}(d_{\mathcal{H}}, k)$ and every $\tau, \gamma \in T_{\lambda, d_{\mathcal{H}}}$, expressing the polynomial $f_{\tau,\gamma}(X)$ as a linear combination of monomials can be done in* $\mathrm{poly}(k)$ *time, for fixed $d_{\mathcal{H}}$.*

We use this to prove the following lemma:

Lemma 5.4.5 (Lemma 2, [LPS17]). *Let $\lambda \in \mathrm{Par}(d_{\mathcal{H}}, k)$, $\tau, \gamma \in T_{\lambda, d_{\mathcal{H}}}$, and $r \in [m^{\mathcal{H}}]$. Then $u_\tau^T C_r^{\mathcal{H}} u_\gamma$ can be computed in polynomial time in k, for fixed $d_{\mathcal{H}}$.*

Proof. The proof can be found in [LPS17], but we include a concise proof for the reader's convenience. For every $r \in [m^{\mathcal{H}}]$, it is straightforward to see that $u_\tau^T C_r^{\mathcal{H}} u_\gamma = (u_\tau \otimes u_\gamma)^T \mathrm{vec}\left(C_r^{\mathcal{H}}\right)$. Therefore, by a permutation of the tensor factors, we get $u_\tau^T C_r^{\mathcal{H}} u_\gamma = w\, \mathsf{C}_r^{\mathcal{H}}$, for $w \in (W_{\mathcal{H}}^*)^{\otimes k}$ given by

$$w = \sum_{\substack{\tau' \sim \tau \\ \gamma' \sim \gamma}} \sum_{c, c' \in C_\lambda} \mathrm{sgn}(cc') \bigotimes_{y \in Y(\lambda)} (A)_{\tau'(c(y)), \gamma'(c'(y))}\,,$$

where $A \in (W^*)^{d_{\mathcal{H}} \times d_{\mathcal{H}}}$ with $(A)_{x,y} = a^*_{(x,y)}$. Then

$$\begin{aligned}
\sum_{r \in [m^{\mathcal{H}}]} \left(u_\tau^T C_r^{\mathcal{H}} u_\gamma\right) m(\mathsf{O}_r^{\mathcal{H}}) &= \sum_{r \in [m^{\mathcal{H}}]} \left(w\, \mathsf{C}_r^{\mathcal{H}}\right) m(\mathsf{O}_r^{\mathcal{H}}) \\
&= \sum_{(p_1, \ldots, p_k) \in ([d_{\mathcal{H}}]^2)^k} \left(w\left(a_{p_1} \otimes \cdots \otimes a_{p_k}\right)\right) a^*_{p_1} \cdots a^*_{p_k} \\
&= \overline{w} = \sum_{\substack{\tau' \sim \tau \\ \gamma' \sim \gamma}} \sum_{c, c' \in C_\lambda} \mathrm{sgn}(cc') \prod_{y \in Y(\lambda)} (A)_{\tau'(c(y)), \gamma'(c'(y))} \\
&= f_{\tau,\gamma}(A)\,.
\end{aligned}$$

Therefore, $u_\tau^T C_r^{\mathcal{H}} u_\gamma$ is exactly the coefficient of the monomial $m(\mathsf{O}_r^{\mathcal{H}})$ in $f_{\tau,\gamma}(A)$, which by Lemma 5.4.4 can be computed in $\mathrm{poly}(k)$ time. □

Theorem 5.4.6. *There exists an algorithm which given as input $J^{\mathcal{M}}$, $J^{\mathcal{N}}$, and $k \in \mathbb{N}$, outputs in $\mathrm{poly}(k)$ time (for fixed $\dim(X \otimes Y)$) the description of a convex program of size described in Theorem 5.4.3 for computing $\mathrm{D}_\alpha^{\#}(\mathcal{N}^{\otimes k} \| \mathcal{M}^{\otimes k})$.*

Proof. For $\mathcal{H} \in \{X \otimes Y, X\}$, let $\left\{O_r^{\mathcal{H}}\right\}_{r \in [m^{\mathcal{H}}]}$ denote the set of orbits of pairs and $\left\{C_r^{\mathcal{H}}\right\}_{r \in [m^{\mathcal{H}}]}$ denote the canonical basis of $\mathrm{End}^{\mathfrak{S}_k}\left(\mathcal{H}^{\otimes k}\right)$ defined in Eq. (5.25). For every $r \in [m^{X \otimes Y}]$, we define $D_r \coloneqq \mathrm{tr}_{Y^{\otimes k}}\left(C_r^{X \otimes Y}\right)$. Note that $D_r \in \mathrm{End}^{\mathfrak{S}_k}\left(X^{\otimes k}\right)$. Then by Theorem 5.4.3, $\mathrm{D}_\alpha^{\#}(\mathcal{N}^{\otimes k} \| \mathcal{M}^{\otimes k})$ can be formulated as the following convex program:

$$\begin{aligned}
&\frac{1}{\alpha - 1} \log \min \quad y \\
&\quad \text{s.t.} \sum_{r=1}^{m^{X \otimes Y}} z_r \left[\!\left[\phi_X(D_r)\right]\!\right]_j \leq y\, \mathrm{id}_{m_j^X}\ , && j \in \left[t^X\right] \\
&\qquad \left[\!\left[\phi_{X \otimes Y}\left(\left(J^{\mathcal{N}}\right)^{\otimes k}\right)\right]\!\right]_i \leq \left[\!\left[\phi_{X \otimes Y}\left(\left(J^{\mathcal{M}}\right)^{\otimes k}\right)\right]\!\right]_i \#_{1/\alpha} \sum_{r=1}^{m^{X \otimes Y}} z_r \left[\!\left[\phi_{X \otimes Y}(C_r^{X \otimes Y})\right]\!\right]_i\ , && i \in \left[t^{X \otimes Y}\right] \\
&\qquad \sum_{r=1}^{m^{X \otimes Y}} z_r \left[\!\left[\phi_{X \otimes Y}(C_r^{X \otimes Y})\right]\!\right]_i \geq 0\ , && i \in \left[t^{X \otimes Y}\right] \\
&\qquad y, z_r \in \mathbb{R}\ , && r \in \left[m^{X \otimes Y}\right]
\end{aligned}$$

Here, we use the notation introduced in Theorem 5.4.3. By Lemma 5.4.5, the block diagonal matrices $\phi_{X\otimes Y}(C_r^{\mathcal{H}})$ can be computed in poly(k) time, for every $r \in [m^{\mathcal{H}}]$. Therefore, to complete the proof it suffices to show how to expand $\left(J^{\mathcal{N}}\right)^{\otimes k}, \left(J^{\mathcal{M}}\right)^{\otimes k}$ in the basis $\left\{C_r^{X\otimes Y}\right\}_{r\in[m^{X\otimes Y}]}$ and D_r in the basis $\left\{C_r^{X}\right\}_{r\in[m^{X}]}$, for every $r \in [m^{X\otimes Y}]$.

For $\left(J^{\mathcal{M}}\right)^{\otimes k} \in \mathrm{End}^{\mathfrak{S}_k}\left((X\otimes Y)^{\otimes k}\right)$, if we take an arbitrary representative element $(p_1, \ldots, p_k)$ of $\mathsf{O}_r^{X\otimes Y}$, for every $r \in \left[m^{X\otimes Y}\right]$, and define

$$z_r \coloneqq \prod_{t=1}^{k} (J^{\mathcal{M}})_{p_t} \,,$$

then we have $\left(J^{\mathcal{M}}\right)^{\otimes k} = \sum_{r=1}^{m^{X\otimes Y}} z_r C_r^{X\otimes Y}$. The same method can be used for $\left(J^{\mathcal{N}}\right)^{\otimes k}$.

Recall that, for every $r \in \left[m^{X\otimes Y}\right]$, we have

$$C_r^{X\otimes Y} = \sum_{(i,j)\in O_r^{X\otimes Y}} |i\rangle\langle j| \,,$$

where $i = \left(i_1^X i_1^Y \cdots i_k^X i_k^Y\right)$ and $j = \left(j_1^X j_1^Y \cdots j_k^X j_k^Y\right)$. For any representative element (i,j) of $O_r^{X\otimes Y}$ if $i^Y = (i_1^Y \cdots i_k^Y) \neq j^Y = j_1^Y \cdots j_k^Y$ then $\mathrm{tr}_{Y^{\otimes k}}\left(|i\rangle\langle j|\right) = 0$. Therefore,

$$D_r = \sum_{\substack{(i,j)\in O_r^{X\otimes Y} \\ i^Y = j^Y}} |i^X\rangle\langle j^X|.$$

Moreover, for any representative element (i,j) of $O_r^{X\otimes Y}$, we can determine the orbit O_t^X that contains (i^X, j^X) in poly(k) time. So if we define $\alpha \coloneqq |\{\pi \in \mathfrak{S}_k : \pi(i^X) = i^X, \pi(j^X) = j^X\}|$, then

$$D_r = \alpha\, C_t^X \,.$$

Furthermore, we have $\alpha = \prod_{a,b\in[d_X]} [(E^{(i^X,j^X)})_{a,b}]!$ with $E^{(i^X,j^X)} \in \mathbb{Z}_{\geq 0}^{d_X\times d_X}$ defined in Eq. (5.26). This concludes the proof. □

Alternatively, the regular $*$-representation approach can be used to show that the convex program (5.12) can be computed in poly(k) time. For $\mathcal{H} \in \{X, X\otimes Y\}$, let $\psi_{\mathcal{H}}$ be the regular $*$-representation of $\mathrm{End}^{\mathfrak{S}_k}\left(\mathcal{H}^{\otimes k}\right)$, defined explicitly in Theorem 5.3.2. We denote by $\{O_r^{\mathcal{H}}\}_{r\in[m^{\mathcal{H}}]}$ and $\{C_r^{\mathcal{H}}\}_{r\in[m^{\mathcal{H}}]}$, the orbits of pairs and the canonical basis of $\mathrm{End}^{\mathfrak{S}_k}(\mathcal{H}^{\otimes k})$, following the construction in Eq. (5.25). The convex program can be reformulated as

$$\begin{aligned}
\frac{1}{\alpha-1}\log \quad \min \quad & y \\
\text{s.t.} \quad & \textstyle\sum_{r=1}^{m^{X\otimes Y}} x_r\, \psi_X(D_r) \leq y \mathrm{id}_{m^X} \ , \\
& \psi_{X\otimes Y}\left(\left(J^{\mathcal{N}}\right)^{\otimes k}\right) \leq \psi_{X\otimes Y}\left(\left(J^{\mathcal{M}}\right)^{\otimes k}\right) \#_{1/\alpha} \textstyle\sum_{r=1}^{m^{X\otimes Y}} x_r\, \psi_{X\otimes Y}\left(C_r^{X\otimes Y}\right) \ , \\
& \textstyle\sum_r x_r\, \psi_{X\otimes Y}\left(C_r^{X\otimes Y}\right) \geq 0 \ , \\
& x_1, \ldots, x_{m^{X\otimes Y}}, y \in \mathbb{R} \ .
\end{aligned}$$

Recall that $\psi_{\mathcal{H}}\left(\mathrm{End}^{\mathfrak{S}_k}\left(\mathcal{H}^{\otimes k}\right)\right) \subseteq \mathbb{C}^{m^{\mathcal{H}} \times m^{\mathcal{H}}}$, where $m^{\mathcal{H}} \leq (k+1)^{d_{\mathcal{H}}^2}$.

Note that $\|C_r^{\mathcal{H}}\| := \sqrt{\langle C_r^{\mathcal{H}}, C_r^{\mathcal{H}}\rangle}$ equals the size of the orbit $O_r^{\mathcal{H}}$. Using the structure of the orbits, we can compute the multiplication parameters of $\mathrm{End}^{\mathfrak{S}_k}\left(\mathcal{H}^{\otimes k}\right)$ with respect to the orthogonal basis $\{C_1^{\mathcal{H}}, \ldots, C_{m^{\mathcal{H}}}^{\mathcal{H}}\}$ as

$$p_{rs}^t = \left|\left\{l \in \left[(d_{\mathcal{H}})^k\right] \ : \ (i,l) \in O_r^{\mathcal{H}} \ , \ (l,j) \in O_s^{\mathcal{H}}\right\}\right|,$$

where $(i,j) \in O_t^{\mathcal{H}}$. Here, p_{rs}^t does not depend on the choice of i and j. Let E^s, E^r, E^t be the matrices defined in Eq. (5.26) for orbits $O_s^{\mathcal{H}}, O_r^{\mathcal{H}}, O_t^{\mathcal{H}}$, respectively. The following proposition implies that p_{rs}^t can be computed in poly(k) time.

Proposition 5.4.7 ([Gij09])**.** *The numbers p_{rs}^t are given by*

$$p_{rs}^t = \sum_B \prod_{x,y=1}^{d_{\mathcal{H}}} \binom{(E^t)_{x,y}}{B_{x,1,y}, \ldots, B_{x,d_{\mathcal{H}},y}},$$

where the sum runs over all $B \in \mathbb{Z}_{\geq 0}^{d_{\mathcal{H}} \times d_{\mathcal{H}} \times d_{\mathcal{H}}}$ that satisfy $\sum_z B_{x,y,z} = (E^r)_{x,y}$, $\sum_x B_{x,y,z} = (E^s)_{y,z}$, $\sum_y B_{x,y,z} = (E^t)_{x,z}$ for all $x, y, z \in [d_{\mathcal{H}}]$ and $\sum_{x,y,z \in [d_{\mathcal{H}}]} B_{x,y,z} = k$.

Table 5.1 compares the reduction in the size of the matrices for different values of k, using both methods of regular $*$-representation and block-diagonal decomposition. The first column contains $\dim\left[\mathscr{L}(X^{\otimes k} \otimes Y^{\otimes k})\right]$, for $X = Y = \mathbb{C}^2$ and different values of k. The numbers in the second column correspond to the reduced matrix sizes using regular $*$-representation and the third column contains the block sizes in the block-diagonal decomposition. As illustrated by these examples, the size of the variables and the constraint matrices can be significantly reduced by using block-diagonalization. While the reduction obtained by using the regular $*$-representation is not as strong, it has the advantage that it is easy to compute using the explicit formula given in Eq. (5.5).

k	$\dim \mathscr{L}(X^{\otimes k} \otimes Y^{\otimes k})$	$\dim \mathrm{End}^{\mathfrak{S}_k}\left(X^{\otimes k} \otimes Y^{\otimes k}\right)$	Block sizes
2	256	136	$10, 6$
3	4096	816	$20, 20, 4$
4	65536	3876	$45, 35, 20, 15, 1$

Table 5.1: Dimensions of $\mathscr{L}(X^{\otimes k} \otimes Y^{\otimes k})$, $\mathrm{End}^{\mathfrak{S}_k}(X^{\otimes k} \otimes Y^{\otimes k})$, and the block sizes in the block-diagonal form with $X = Y = \mathbb{C}^2$.

5.4.2 Beyond permutation invariance

So far, we have only focused on the permutation symmetries of convex optimization problem (5.12) arising from considering multiple copies of quantum channels. In this section, we discuss how the group symmetries of the underlying channels may be used to further simplify these convex programs. In particular, we show how the symmetries of the channels can be combined with the permutation symmetry and expressed as invariance under the action of a single group. Theorem 5.3.1 is then used to simplify the programs.

Let G be a finite group, and denote by G^k, the k-fold direct product of G. Consider the group $H := G^k \rtimes_\gamma \mathfrak{S}_k$, an outer semi-direct product of G^k and $\mathfrak{S}_k$, defined as follows:

- The underlying set is the Cartesian product of the sets G^k and $\mathfrak{S}_k$, i.e., the set of ordered pairs (g, π), where $g = (g_1, g_2, \ldots, g_k) \in G^k$ and $\pi \in \mathfrak{S}_k$.
- $\gamma : \mathfrak{S}_k \to \mathrm{Aut}\left(G^k\right)$ is a group homomorphism given by

$$\gamma(\pi)\,(g_1, g_2, \ldots, g_k) = \left(g_{\pi(1)}, g_{\pi(2)}, \ldots, g_{\pi(k)}\right) \quad ,$$

  for every $\pi \in \mathfrak{S}_k$ and $g = (g_1, g_2, \ldots, g_k) \in G^k$.
- The group operation $*$ is defined for any pair $(g, \pi), (g', \pi') \in H$ as

$$(g', \pi') * (g, \pi) = (g'\gamma(\pi')(g), \pi'\pi).$$

Consider an arbitrary action of G on a finite dimensional Hilbert space $\mathcal{H}$, and the natural action of $\mathfrak{S}_k$ on $\mathcal{H}^{\otimes k}$ defined for every $\pi \in \mathfrak{S}_k$ as

$$\pi \cdot (h_1 \otimes \cdots \otimes h_k) = h_{\pi^{-1}(1)} \otimes \cdots \otimes h_{\pi^{-1}(k)} \quad , \quad h_i \in \mathcal{H}\,, \forall i \in [k] \ . \tag{5.32}$$

Then it is easy to check that the following defines an action of $H = G^k \rtimes_\gamma \mathfrak{S}_k$ on $\mathcal{H}^{\otimes k}$:

$$(g, \pi) \cdot (h_1 \otimes \cdots \otimes h_k) = g_1 \cdot h_{\pi^{-1}(1)} \otimes \cdots \otimes g_k \cdot h_{\pi^{-1}(k)} \quad , \quad h_i \in \mathcal{H}\,, \forall i \in [k] \ , \tag{5.33}$$

for all $\pi \in \mathfrak{S}_k$ and $g \in G^k$. In particular, we have

$$(g', \pi') \cdot ((g, \pi) \cdot (h_1 \otimes \cdots \otimes h_k)) = \left((g', \pi') * ((g, \pi)) \cdot (h_1 \otimes \cdots \otimes h_k)\right) \quad ,$$

for every $(g, \pi), (g', \pi') \in H$.

For $\mathcal{H} \in \{X, Y\}$, let $\rho_{\mathcal{H}} : G \to \mathrm{GL}(\mathcal{H})$ be the representation of G defined by its action on $\mathcal{H}$ and $\rho_{X\otimes Y} \coloneqq \rho_X \otimes \rho_Y$. Let $\sigma_{\mathcal{H}}$ denote the representation of G^k on $\mathcal{H}^{\otimes k}$ given by $\sigma_{\mathcal{H}}(g) \coloneqq \rho_{\mathcal{H}}(g_1) \otimes \ldots \otimes \rho_{\mathcal{H}}(g_k)$, for every $g \in G^k$. As before, denote by $P_{\mathcal{H}}$ the representation of $\mathfrak{S}_k$ on $\mathcal{H}^{\otimes k}$ defined by the action (5.32). Then the representation of $H = G^k \rtimes_\gamma \mathfrak{S}_k$ defined above on $\mathcal{H}^{\otimes k}$ is given by $\sigma_{\mathcal{H}}(g)P_{\mathcal{H}}(\pi)$, for every $(g, \pi) \in H$. Note that in (5.33), for $\mathcal{H} = X \otimes Y$, the action of (g, π) on $(X \otimes Y)^{\otimes k}$ corresponds to the simultaneous permutation of the X and Y tensor factors followed by applying $\rho_X(g_i) \otimes \rho_Y(g_i)$ on i-th $X \otimes Y$ tensor factor. When the subsystems are reordered as $X^{\otimes k} \otimes Y^{\otimes k}$, this action is simply given by $\sigma_X(g)P_X(\pi) \otimes \sigma_Y(g)P_Y(\pi)$. With the above notation, we are now ready to state the following proposition:

Proposition 5.4.8. *Let $\mathcal{N}_{X\to Y}$ and $\mathcal{M}_{X\to Y}$ be a quantum channels with Choi operators $J^{\mathcal{N}}, J^{\mathcal{M}} \in \mathrm{End}^G(X \otimes Y)$, for some finite group G. Then the convex program (5.12) has an optimal solution $A \in \mathrm{End}^H(X^{\otimes k} \otimes Y^{\otimes k})$, where $H = G^k \rtimes_\gamma \mathfrak{S}_k$.*

Proof. The proof is based on convexity and exactly follows the steps of the proof of Lemma 5.4.2, except the group average operator $\bar{A}$ is now obtained with respect to the group H. □

Next, we discuss the irreducible representations of $G^k \rtimes_\gamma \mathfrak{S}_k$ and the corresponding multiplicities for the action of H on $\mathcal{H}^{\otimes k}$, defined in Equation (5.33). First, we need to introduce some notations.

Suppose that G has t irreducible representations and let m_i denote the multiplicity of the i-th irreducible representation in the representation $\rho_{\mathcal{H}}$ of G on $\mathcal{H}$. Let $\mathcal{T}(k)$ be the collection of all t-tuples $(k_1, \ldots, k_t)$ of non-negative integers such that $\sum_{i=1}^{t} k_i = k$. For $(k_1, \ldots, k_t) \in \mathcal{T}(k)$ and $(\lambda^1, \ldots, \lambda^t)$ satisfying $\lambda^i \vdash_{m_i} k_i$, for every $i \in [t]$, we write $(\lambda^1, \ldots, \lambda^t) \vdash_{\boldsymbol{m}} (k_1, \ldots, k_t)$, where $\boldsymbol{m} = (m_1, \ldots, m_t)$. We then use a result from [Pol19].

Proposition 5.4.9 (Proposition 3.1.1, [Pol19])**.** *The irreducible representations of $H = G^k \rtimes_\gamma \mathfrak{S}_k$ are labeled by*

$$\{(k_1, \ldots, k_t), (\lambda^1, \ldots, \lambda^t) : (k_1, \ldots, k_t) \in \mathcal{T}(k), (\lambda^1, \ldots, \lambda^t) \vdash_{\boldsymbol{m}} (k_1, \ldots, k_t)\}.$$

and the corresponding multiplicities are $\prod_{i=1}^{t} |T_{\lambda^i, m_i}|$.

Note that $|\mathcal{T}(k)| = \binom{k+t-1}{t-1}$, where t, the number of irreducible representations of G, is a property of G and independent of k. Since G is a finite group, we have $t \leq |G|$. Moreover, for a fixed tuple $(k_1, \ldots, k_t) \in \mathcal{T}(k)$, by Inequality (5.19), we have the size of the set $\{(\lambda^1, \ldots, \lambda^t) : (\lambda^1, \ldots, \lambda^t) \vdash (k_1, \ldots, k_t)\}$ is at most $\prod_{i=1}^{t}(k_i + 1)^{m_i}$. Since $m_i \leq \dim(\mathcal{H})$, for every $i \in [t]$, the number of irreducible representations of H is polynomial in k. Since $|T_{\lambda^i, m_i}| \leq (k_i + 1)^{m_i(m_i-1)/2}$, the multiplicity of the corresponding irreducible representation of H is at most $\prod_{i=1}^{t} |T_{\lambda^i, m_i}| \leq \prod_{i=1}^{t}(k_i + 1)^{m_i(m_i-1)/2}$.

Application to the generalized amplitude damping channel

As an application, we consider the generalized amplitude damping (GAD) channel defined as

$$\mathcal{A}_{p,q}(\rho) = \sum_{i=1}^{4} A_i \rho A_i^*, \quad p, q \in [0, 1] \tag{5.34}$$

with the Kraus operators

$$\begin{aligned} A_1 &= \sqrt{1-q}(|0\rangle\langle 0| + \sqrt{1-p}|1\rangle\langle 1|), \quad & A_2 &= \sqrt{p(1-q)}(|0\rangle\langle 1|), \\ A_3 &= \sqrt{q}(\sqrt{1-p}|0\rangle\langle 0| + |1\rangle\langle 1|), & A_4 &= \sqrt{pq}|1\rangle\langle 0|. \end{aligned} \tag{5.35}$$

The GAD channel reduces to the conventional amplitude damping (AD) channel, when $q = 0$. In this case we have $X = Y = \mathbb{C}^2$. Let $N_{p,q}$ be the Choi matrix of $\mathcal{A}_{p,q}$. Note that for the Pauli Z operator given by

$$\mathrm{Z} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

we have, $(\mathrm{Z} \otimes \mathrm{Z}) N_{p,q} (\mathrm{Z} \otimes \mathrm{Z}) = N_{p,q}$ for all $p, q \in [0, 1]$. Let $G = \mathbb{Z}_2$ be the cyclic group of order 2 and define the group representation $\rho : G \to \mathrm{GL}(\mathbb{C}^2)$ given by $\rho(1) = \mathrm{Z}$. Then for the representation $\rho_{X \otimes Y}$ defined for every $g \in \mathbb{Z}_2$ as $\rho_{X \otimes Y}(g) = \rho(g) \otimes \rho(g)$, we have $N_{p,q} \in \mathrm{End}^G(X \otimes Y)$. The representation $\rho_{X \otimes Y}$ has two irreducible representations, which are both 1-dimensional (since G is an Abelian group). In this representation, the multiplicities are $(m_1, m_2) = (2, 2)$. Therefore, the multiplicities in the representation of $H = G^k \rtimes_\gamma \mathfrak{S}_k$ on $X^{\otimes k} \otimes Y^{\otimes k}$ are at most $(k_1 + 1)(k_2 + 1) \leq (k_1 + k_2 + 2)^2/4 = (k + 2)^2/4$.

Furthermore, since $t = 2$, we have $|\mathcal{T}(k)| = k$, and for any $(k_1, k_2) \in \mathcal{T}(k)$, the size of the set $\{(\lambda^1, \lambda^2) : (\lambda^1, \lambda^2) \vdash (k_1, k_2)\}$ is at most $(k_1 + 1)^2(k_2 + 1)^2 \leq (k + 2)^4/16$. Therefore the number of irreducible representations of H is at most $k(k+2)^4/16$. Since the dimension of the invariant subspace is equal to the sum of squares of the multiplicities of the irreducible representations, we have $\dim \mathrm{End}^H(X^{\otimes k} \otimes Y^{\otimes k}) \leq \left((k + 2)^2/4\right)^2 k(k+2)^4/16 = k(k + 2)^8/256$.

Therefore, in this example, by considering the additional Z symmetry discussed above, we can reduce the dimension of the invariant subspace from $\mathrm{O}\left(k^{16}\right)$ for the permutation action (see Eq. (5.21)) to $\mathrm{O}\left(k^9\right)$, when we combine the two symmetries. Moreover, the maximum block size is reduced from $\mathrm{O}\left(k^6\right)$ (see Eq. (5.20)) to $\mathrm{O}\left(k^2\right)$. This shows the potential of the approach introduced above for channels with stronger symmetries.

In the following table, we compare the dimensions of the $\mathfrak{S}_k$-invariant and H-invariant subspace of operators for $X = Y = \mathbb{C}^2$ and different values of k. We also list the number of irreducible representations and the maximum block size of the invariant operators in the block-diagonal form.

k	$\mathfrak{S}_k$			$G^k \rtimes \mathfrak{S}_k$		
	$\dim \mathrm{End}^{\mathfrak{S}_k}(\mathcal{H}^{\otimes k})$	max.block size	#-irreps	$\dim \mathrm{End}^{G^k \rtimes \mathfrak{S}_k}(\mathcal{H}^{\otimes k})$	max.block size	#-irreps
2	136	10	2	36	4	5
3	816	20	3	120	6	8
4	3876	45	5	330	9	14
5	15504	84	6	792	12	20
6	54264	140	9	1716	16	30
7	170544	224	11	3432	20	40
8	490314	360	15	6435	25	55
9	1307504	540	18	11440	30	70
10	3268760	770	23	19448	36	91

Table 5.2: The comparison of the reductions obtained by considering invariance under the action of $\mathfrak{S}_k$ and $G^k \rtimes \mathfrak{S}_k$ on $\mathcal{H}^{\otimes k}$, where $\mathcal{H} = \mathbb{C}^2 \otimes \mathbb{C}^2$.

We use our method for efficient computation of the #-Rényi divergence between multiple copies of channels to provide improved upper bounds on the regularized Umegaki divergence between the AD channel $\mathcal{A}_{0.3,0}$ and the GAD channel $\mathcal{A}_{p,0.9}$, over the range $p \in [0.4, 0.8]$. Note that the Umegaki divergence between these channels is known to be non-additive [FFRS20], i.e., $\mathrm{D}^{\mathrm{reg}}(\mathcal{A}_{0.3,0} \| \mathcal{A}_{p,0.9}) > \mathrm{D}(\mathcal{A}_{0.3,0} \| \mathcal{A}_{p,0.9})$. Figure 5.1 illustrates the improvement obtained using $\mathrm{D}_\alpha^{\#}$ on $k = 1$ and $k = 6$ copies compared to $\widehat{D}_\alpha$, for $\alpha = 2$. The convex programs are implemented in MATLAB using the CVX package [GB14] and the CVXQUAD package [FSP18], via the MOSEK solver [ApS19]. Computations in this experiment were done using Intel(R) Core i5-6300U with 16GB of RAM memory. The running time for $k = 6$ copies on our program is less than 45 minutes while the program without using symmetry cannot be carried out due to insufficient memory. We note that without using the symmetry reduction the matrices are of size 4096×4096.

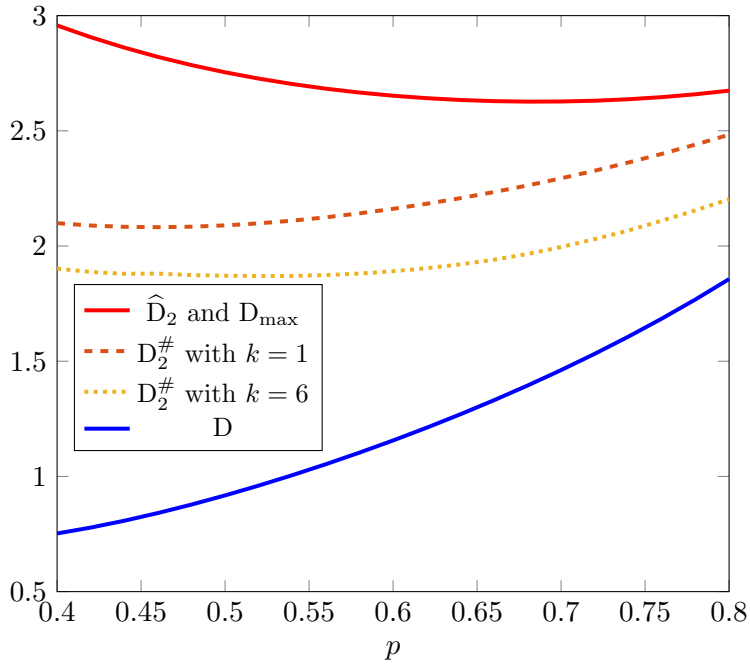

Figure 5.1: Comparison between different bounds on $\mathrm{D}^{\mathrm{reg}}(\mathcal{A}_{0.3,0}\|\mathcal{A}_{p,0.9})$ over the range $p \in [0.4, 0.8]$.

5.5 Efficient bounds on classical capacity of quantum channels

The *(unassisted) classical capacity* of a quantum channel is defined as the maximum rate at which classical information can be transmitted over the quantum channel in the asymptotic limit of many channel uses. For a quantum channel $\mathcal{N}$, the classical capacity is characterized by the regularized Holevo information [Hol98, SW97] as

$$\mathrm{C}(\mathcal{N}) = \lim_{k\to\infty} \frac{1}{k}\chi(\mathcal{N}^{\otimes k}) \ ,$$

where $\chi(\mathcal{N})$ is the Holevo capacity of the channel $\mathcal{N}$ defined as

$$\chi(\mathcal{N}) \coloneqq \max_{\mathcal{E}=\{p_i,\rho_i\}_i} \mathrm{H}\left(\sum_i p_i\rho_i\right) - \sum_i p_i \mathrm{H}\left(\mathcal{N}(\rho_i)\right) \ ,$$

where the maximization is over all quantum ensembles $\mathcal{E} = \{p_i, \rho_i\}_i$. Here, H denotes the von Neumann entropy, defined as $\mathrm{H}(\sigma) \coloneqq -\mathrm{tr}\left(\rho \log \rho\right)$, for every positive semidefinite operator ρ. Note that the Holevo information is in general non-additive [Has09].

We denote by $\mathcal{V}_{\mathrm{cb}}(X, Y)$ the set of constant bounded subchannels from $\mathscr{L}(X)$ to $\mathscr{L}(Y)$ defined as

$$\mathcal{V}_{\mathrm{cb}}(X, Y) \coloneqq \{\mathcal{M} \in \mathrm{CP}(X : Y) : \exists\, \sigma \in \mathrm{D}(Y) \text{ s.t. } \mathcal{M}_{X\to Y}(\rho) \leq \sigma, \forall \rho \in \mathrm{D}(X)\} \ .$$

Let $\mathcal{V}(X,Y) \coloneqq \{\mathcal{M} \in \mathrm{CP}(X:Y) : \beta(J_{XY}^{\mathcal{M}}) \leq 1\}$, with $\beta(J_{XY}^{\mathcal{M}})$ defined in terms of the following SDP

$$\beta(J_{XY}^{\mathcal{M}}) \coloneqq \min_{R_{XY},S_Y} \mathrm{tr}(S_Y) \quad \text{s.t.} \quad R_{XY} \pm (J_{XY}^{\mathcal{M}})^{T_Y} \geq 0 \ , \ \mathrm{id}_X \otimes S_Y \pm R_{XY}^{T_Y} \geq 0 \ ,$$

where $(\cdot)^{T_Y}$ denotes the partial transpose on system Y. Note that the set $\mathcal{V}(X,Y)$ is a convex subset of $\mathcal{V}_{\mathrm{cb}}(X,Y)$ containing all the constant channels [WFT19].

Let $\mathbf{D}$ be a generalized quantum divergence. For any quantum channel $\mathcal{N}_{X\to Y}$, define

$$\Upsilon(\mathbf{D},k)(\mathcal{N}) := \min_{\mathcal{M}\in\mathcal{V}(X^{\otimes k},Y^{\otimes k})} \mathbf{D}(\mathcal{N}^{\otimes k}\|\mathcal{M}).$$

The following proposition provides upper bounds on the classical capacity of a quantum channel.

Proposition 5.5.1 ([WFT19])**.** *Let $\mathbf{D}$ be a generalized quantum divergence. If $\mathbf{D}$ is bounded below by the Umegaki relative entropy on quantum states and the corresponding channel divergence is subadditive under tensor product of channels, then, for any $k \geq 1$,*

$$\mathrm{C}(\mathcal{N}) \leq \frac{1}{k}\Upsilon(\mathbf{D},k)(\mathcal{N})\,.$$

Proof. The proof can be found in [WFT19], but we include a concise proof for the reader's convenience. As shown in [OPW97] the Holevo information can be written as a divergence radius:

$$\begin{aligned}
\chi(\mathcal{N}) &= \min_{\sigma\in\mathrm{D}(Y)} \max_{\rho\in\mathrm{D}(X)} \mathrm{D}(\mathcal{N}(\rho)\|\sigma) \\
&= \min_{\mathcal{M}\in\mathcal{V}_{\mathrm{cb}}(X,Y)} \max_{\rho\in\mathrm{D}(X)} \mathrm{D}(\mathcal{N}(\rho)\|\mathcal{M}(\rho)) \\
&\leq \min_{\mathcal{M}\in\mathcal{V}(X,Y)} \max_{\rho\in\mathrm{D}(X)} \mathrm{D}(\mathcal{N}(\rho)\|\mathcal{M}(\rho)) \\
&\leq \min_{\mathcal{M}\in\mathcal{V}(X,Y)} \mathrm{D}(\mathcal{N}\|\mathcal{M})
\end{aligned}$$

where we used the fact that if $\sigma \leq \sigma'$ then $\mathrm{D}(\rho\|\sigma) \geq \mathrm{D}(\rho\|\sigma')$ and the fact that $\mathcal{V}(X,Y) \subseteq \mathcal{V}_{\mathrm{cb}}(X,Y)$. So, for $n,k \in \mathbb{N}$, we have

$$\begin{aligned}
\chi(\mathcal{N}^{\otimes nk}) &\leq \min_{\mathcal{M}\in\mathcal{V}(X^{\otimes nk},Y^{\otimes nk})} \mathrm{D}(\mathcal{N}^{\otimes nk}\|\mathcal{M}) \\
&\leq \min_{\mathcal{M}\in\mathcal{V}(X^{\otimes k},Y^{\otimes k})} \mathrm{D}(\mathcal{N}^{\otimes nk}\|\mathcal{M}^{\otimes n}),
\end{aligned}$$

where we used the fact that if $\mathcal{M} \in \mathcal{V}(X^{\otimes k},Y^{\otimes k})$, then $\mathcal{M}^{\otimes n} \in \mathcal{V}(X^{\otimes nk},Y^{\otimes nk})$. Since $\mathbf{D}$ is bounded below by D and subadditive under tensor product of channels, we have

$$\frac{1}{nk}\chi(\mathcal{N}^{\otimes nk}) \leq \min_{\mathcal{M}\in\mathcal{V}(X^{\otimes k},Y^{\otimes k})} \frac{1}{k}\,\mathbf{D}(\mathcal{N}^{\otimes k}\|\mathcal{M}) = \frac{1}{k}\Upsilon(\mathbf{D},k)(\mathcal{N}).$$

Taking the limit as $n \to \infty$, we get the desired result. $\square$

Note that by Proposition 5.2.1, for $\alpha \in (1,2]$, we have

$$\Upsilon(\widetilde{\mathrm{D}}_\alpha, k)(\mathcal{N}) \leq \Upsilon(\mathrm{D}^{\#}_\alpha, k)(\mathcal{N}) \leq \Upsilon(\widehat{\mathrm{D}}_\alpha, k)(\mathcal{N}) \leq \Upsilon(\mathrm{D}_{\max}, k)(\mathcal{N}).$$

Remark 5.5.2. If in addition the generalized quantum divergence $\mathbf{D}$ satisfies $\widetilde{\mathrm{D}}_\alpha \leq \mathbf{D}$, for some $\alpha \in (1,\infty)$, then $\frac{1}{k}\Upsilon(\mathbf{D}, k)(\mathcal{N})$ is a strong converse bound, i.e., above this communication rate, the error probability goes to 1.

Both $\mathrm{D}_{\max}$ and $\widehat{\mathrm{D}}_\alpha$ have the desired properties and were used in [WFT19] and [FF21a] to obtain bounds on the classical capacity. On the other hand, $\widetilde{\mathrm{D}}_\alpha$ is not always additive [FFRS20] so it cannot be used in general. The best-known general strong converse bound is given by $\frac{1}{k}\Upsilon(\widehat{\mathrm{D}}_\alpha, k)$, and it is SDP computable [FF21a]. For $\mathbf{D} = \mathrm{D}^{\#}_\alpha$, using the formulation of the channel divergence given in Eqs. (5.2) and (5.3), the converse bound of Proposition 5.5.1 can be written in terms of a convex program. For every $k \geq 1$, we have

$$\begin{aligned}
\Upsilon(\mathrm{D}^{\#}_\alpha, k)(\mathcal{N}) = \frac{1}{\alpha - 1}\log \quad \min \quad & \|\mathrm{tr}_{Y^{\otimes k}}(A)\|_\infty \\
\text{s.t.} \quad & J^{\mathcal{N}^{\otimes k}} \leq J^{\mathcal{M}} \#_{1/\alpha} A \ , \\
& R \pm (J^{\mathcal{M}})^{T_{Y^{\otimes k}}} \geq 0 \ , \\
& (I_{X^{\otimes k}} \otimes S) \pm R^{T_{Y^{\otimes k}}} \geq 0 \ , \\
& \mathrm{tr}(S) \leq 1 \ , \\
& A, J^{\mathcal{M}}, R \in \mathscr{P}(X^{\otimes k} \otimes Y^{\otimes k}) \, , \, S \in \mathscr{P}(Y^{\otimes k}) \ .
\end{aligned} \tag{5.36}$$

Note that the optimization problem in Eq. (5.36) does not scale well with k since the sizes of the constraint matrices grow exponentially fast. This bottleneck will be addressed in the next section.

5.5.1 Exploiting symmetries to simplify the problem

Using a similar argument as in Lemma 5.4.2, one may restrict the feasible region of the convex program (5.36) to the $\mathfrak{S}_k$-invariant subspace of operators.

Lemma 5.5.3. *For every $\alpha \in (1,\infty)$, the convex program (5.36) has an optimal solution $(A, R, J^{\mathcal{M}}, S)$, with $A, R, J^{\mathcal{M}} \in \mathrm{End}^{\mathfrak{S}_k}\left(X^{\otimes k} \otimes Y^{\otimes k}\right)$ and $S \in \mathrm{End}^{\mathfrak{S}_k}(Y^{\otimes k})$.*

Proof. It is straightforward to check that by Slater's condition the optimal value is achieved by a feasible solution. For an arbitrary feasible solution $(A, J^{\mathcal{M}}, R, S)$, we will prove that the corresponding group-average operators $(\overline{A}, \overline{J^{\mathcal{M}}}, \overline{R}, \overline{S})$ are feasible with an objective value not greater than the original value.

For brevity of notation, we write $\Pi(\pi) := P_{X \otimes Y}(\pi)$. The first constraint, $J^{\mathcal{N}^{\otimes k}} \leq \overline{J^{\mathcal{M}}} \#_{1/\alpha} \overline{A}$, follows from a similar argument as in Lemma 5.4.2. For the second constraint note that, for every $\pi \in \mathfrak{S}_k$, $\Pi(\pi)^* = \Pi(\pi)^T$, and we have

$$\left(\Pi(\pi) J^{\mathcal{M}} \Pi(\pi)^*\right)^{T_{Y^{\otimes k}}} = \left(\Pi(\pi) J^{\mathcal{M}} \Pi(\pi)^T\right)^{T_{Y^{\otimes k}}} = \Pi(\pi)(J^{\mathcal{M}})^{T_{Y^{\otimes k}}} \Pi(\pi)^T \ . \tag{5.37}$$

Therefore,

$$\left(\overline{J^{\mathcal{M}}}\right)^{T_{Y\otimes k}} = \left(\frac{1}{|\mathfrak{S}_k|}\sum_{\pi\in\mathfrak{S}_k}\Pi(\pi)J^{\mathcal{M}}\Pi(\pi)^*\right)^{T_{Y\otimes k}} = \frac{1}{|\mathfrak{S}_k|}\sum_{\pi\in\mathfrak{S}_k}\Pi(\pi)(J^{\mathcal{M}})^{T_{Y\otimes k}}\Pi(\pi)^* \ , \tag{5.38}$$

and the feasibility of $J^{\mathcal{M}}$ and R implies $-\overline{R} \leq \left(\overline{J^{\mathcal{M}}}\right)^{T_{Y\otimes k}} \leq \overline{R}$. Similarly, we get

$$\left(\overline{R}\right)^{T_{Y\otimes k}} = \left(\frac{1}{|\mathfrak{S}_k|}\sum_{\pi\in\mathfrak{S}_k}\Pi(\pi)\,R\,\Pi(\pi)^*\right)^{T_{Y\otimes k}} = \frac{1}{|\mathfrak{S}_k|}\sum_{\pi\in\mathfrak{S}_k}\Pi(\pi)(R)^{T_{Y\otimes k}}\Pi(\pi)^* \ , \tag{5.39}$$

and the feasibility of S and R implies $-\mathrm{id}_{X^{\otimes k}}\otimes\overline{S} \leq \left(\overline{R}\right)^{T_{Y\otimes k}} \leq \mathrm{id}_{X^{\otimes k}}\otimes\overline{S}$. Finally, the forth constraint holds since $\mathrm{tr}(\overline{S}) = \mathrm{tr}(S) \leq 1$.

For the objective function, using the same argument as in Lemma 5.4.2, we get $\|\mathrm{tr}_{Y^{\otimes k}}\left(\overline{A}\right)\|_\infty \leq \|\mathrm{tr}_{Y^{\otimes k}}(A)\|_\infty$. This concludes the proof. □

Next, we show that the convex program (5.36) may be reformulated so that it scales only polynomially with k.

Theorem 5.5.4. *Let $\mathcal{N}_{X\to Y}$ be a quantum channel. For every $k \geq 1$, the strong converse bound $\frac{1}{k}\Upsilon(\mathrm{D}^{\#}_{\alpha}, k)(\mathcal{N})$ of Proposition 5.5.1 can be formulated as a convex program with only $\mathrm{O}\left(k^{d^2}\right)$ variables and $\mathrm{O}\left(k^d\right)$ PSD constraints involving matrices of size at most $(k+1)^{d(d-1)/2}$, where $d = d_X d_Y$.*

Proof. Let Q denote the permutation matrix which maps $X^{\otimes k}\otimes Y^{\otimes k}$ to $(X\otimes Y)^{\otimes k}$. Then, by Lemma 5.5.3, the optimization problem (5.36) can be written as

$$\Upsilon(\mathrm{D}^{\#}_{\alpha}, k)(\mathcal{N}) = \frac{1}{\alpha-1}\ \log\ \min\ y \tag{5.40}$$

$$\text{s.t.}\quad \mathrm{tr}_{Y^{\otimes k}}(A) \leq y\ \mathrm{id}_{X^{\otimes k}}\ , \tag{5.41}$$

$$\left(J^{\mathcal{N}}\right)^{\otimes k} \leq J^{\mathcal{M}}\#_{1/\alpha}A\ , \tag{5.42}$$

$$R \pm (J^{\mathcal{M}})^{T_{Y\otimes k}} \geq 0\ , \tag{5.43}$$

$$Q(\mathrm{id}_{X^{\otimes k}}\otimes S)Q^T \pm R^{T_{Y\otimes k}} \geq 0\ , \tag{5.44}$$

$$\mathrm{tr}(S) \leq 1\ , \tag{5.45}$$

where $A, J^{\mathcal{M}}, R \in \mathrm{End}^{\mathfrak{S}_k}\left((X\otimes Y)^{\otimes k}\right)$ and $S \in \mathrm{End}^{\mathfrak{S}_k}\left(Y^{\otimes k}\right)$ are positive semidefinite operators and $y \in \mathbb{R}$.

Following the notation introduced in Theorem 5.4.6, for $\mathcal{H} \in \{X, Y, X\otimes Y\}$, let $\phi_{\mathcal{H}} : \mathrm{End}^{\mathfrak{S}_k}(\mathcal{H}^{\otimes k}) \to \bigoplus_{i=1}^{t^{\mathcal{H}}}\mathbb{C}^{m_i^{\mathcal{H}}\times m_i^{\mathcal{H}}}$ be the bijective linear map which block-diagonalizes the corresponding invariant algebra, where to simplify the notation, the blocks are indexed by $i \in [t^{\mathcal{H}}]$ instead of $\lambda \in \mathrm{Par}(d_{\mathcal{H}}, k)$. For $Z \in \mathrm{End}^{\mathfrak{S}_k}(\mathcal{H}^{\otimes k})$, we write $[\![\phi_{\mathcal{H}}(Z)]\!]_i$ to denote the i-th block of $\phi_{\mathcal{H}}(Z)$. Since $J^{\mathcal{N}^{\otimes k}}$, $J^{\mathcal{M}}$, A and R, $(J^{\mathcal{M}})^{T_{Y\otimes k}}$ are elements of $\mathrm{End}^{\mathfrak{S}_k}\left((X\otimes Y)^{\otimes k}\right)$, the constraints (5.42) and (5.43) can be mapped into the direct sum

form under $\phi_{X\otimes Y}$. Similarly, since $Q\left(\mathrm{id}_{X^{\otimes k}}\otimes S\right)Q^T$, $R^{T_{Y^{\otimes k}}}\in \mathrm{End}^{\mathfrak{S}_k}\left((X\otimes Y)^{\otimes k}\right)$, by properties 2 and 5 of the α-geometric mean, the constraint (5.44) can be decomposed into constraints involving the smaller diagonal blocks by applying $\phi_{X\otimes Y}$. Finally, since $\mathrm{tr}_{Y^{\otimes k}}(A)$, $\mathrm{id}_{X^{\otimes k}}\in \mathrm{End}^{\mathfrak{S}_k}(X^{\otimes k})$, the constraint (5.41) can be mapped by ϕ_X into the direct sum form. The transformed convex program is given by

$$
\begin{aligned}
\frac{1}{\alpha-1}\log \quad \min \quad & y \\
\text{s.t.} \quad & \left[\!\left[\left(\phi_X\circ \mathrm{tr}_{Y^{\otimes k}}\circ \phi^{-1}_{X\otimes Y}\right)\left(\oplus_l A_l\right)\right]\!\right]_j \leq y\,\mathrm{id}_{m_j^X} \ , \\
& \left[\!\left[\phi_{X\otimes Y}\left(\left(J^{\mathcal{N}}\right)^{\otimes k}\right)\right]\!\right]_i \leq J_i \,\#_{1/\alpha}\, A_i \ , \\
& R_i \pm \left[\!\left[\left(\phi_{X\otimes Y}\circ T_{Y^{\otimes k}}\circ \phi^{-1}_{X\otimes Y}\right)\left(\oplus_l A_l\right)\right]\!\right]_i \geq 0 \ , \\
& \left[\!\left[\phi_{X\otimes Y}\left(Q\left(\mathrm{id}_{X^{\otimes k}}\otimes \phi_Y^{-1}\left(\oplus_r S_r\right)\right)Q^T\right)\right]\!\right]_i \pm \left[\!\left[\left(\phi_{X\otimes Y}\circ T_{Y^{\otimes k}}\circ \phi^{-1}_{X\otimes Y}\right)\left(\oplus_l R_l\right)\right]\!\right]_i \geq 0 \ , \\
& \textstyle\sum_r \mathrm{tr}(S_r)\leq 1 \ , \\
& A_i, R_i, J_i \in \mathscr{P}\left(\mathbb{C}^{m_i^{X\otimes Y}}\right), S_r\in \mathscr{P}\left(\mathbb{C}^{m_r^Y}\right) \ ,
\end{aligned}
$$

for all $i\in\left[t^{X\otimes Y}\right]$, $j\in\left[t^X\right]$, and $r\in\left[t^Y\right]$. The statement of the theorem follows since for $\mathcal{H}\in\{X,Y,X\otimes Y\}$, we have $t^{\mathcal{H}}\leq (k+1)^{d_{\mathcal{H}}}$ and $m_i^{\mathcal{H}}\leq (k+1)^{d_{\mathcal{H}}(d_{\mathcal{H}}-1)/2}$, for every $i\in\left[t^{\mathcal{H}}\right]$. □

Finally, we show how to efficiently compute a formulation of $\frac{1}{k}\Upsilon(\mathrm{D}^{\#}_{\alpha},k)(\mathcal{N})$ as a convex program of polynomial size.

Theorem 5.5.5. *Let $\mathcal{N}_{X\to Y}$ be a quantum channel. There exists an algorithm which given as input $J^{\mathcal{N}}$ and $k\in\mathbb{N}$, outputs in* $\mathrm{poly}(k)$ *time (for fixed* $\dim(X\otimes Y)$*) the description of a convex program of size described in Theorem 5.5.4 for computing the strong converse bound $\frac{1}{k}\Upsilon(\mathrm{D}^{\#}_{\alpha},k)(\mathcal{N})$.*

Proof. As in the proof of Theorem 5.4.6, for $\mathcal{H}\in\{X,Y,X\otimes Y\}$, let $\left\{O_r^{\mathcal{H}}\right\}_{r\in[m^{\mathcal{H}}]}$ denote the set of orbits of pairs and $\left\{C_r^{\mathcal{H}}\right\}_{r\in[m^{\mathcal{H}}]}$ denote the canonical basis of $\mathrm{End}^{\mathfrak{S}_k}\left(\mathcal{H}^{\otimes k}\right)$ defined in Eq. (5.25). For every $r\in[m^{X\otimes Y}]$, we define $D_r\coloneqq \mathrm{tr}_{Y^{\otimes k}}\left(C_r^{X\otimes Y}\right)$. Note that $D_r\in\mathrm{End}^{\mathfrak{S}_k}\left(X^{\otimes k}\right)$. Then by Theorem 5.5.4, $\Upsilon(\mathrm{D}^{\#}_{\alpha},k)(\mathcal{N})$ can be formulated as the following convex program:

$$
\begin{aligned}
&\frac{1}{\alpha-1}\log\ \min\ y \\
&\qquad \text{s.t.}\ \sum_{r=1}^{m^{X\otimes Y}} z_r \left[\!\left[\phi_X\left(D_r\right)\right]\!\right]_j \leq y\,\mathrm{id}_{m_j^X}\ , \\
&\qquad \left[\!\left[\phi_{X\otimes Y}\left(\left(J^{\mathcal{N}}\right)^{\otimes k}\right)\right]\!\right]_i \leq \sum_{l=1}^{m^{X\otimes Y}} x_l \left[\!\left[\phi_{X\otimes Y}\left(C_l^{X\otimes Y}\right)\right]\!\right]_i \#_{1/\alpha} \sum_{r=1}^{m^{X\otimes Y}} z_r \left[\!\left[\phi_{X\otimes Y}\left(C_r^{X\otimes Y}\right)\right]\!\right]_i\ , \\
&\qquad \sum_{l=1}^{m^{X\otimes Y}} y_l \left[\!\left[\phi_{X\otimes Y}\left(C_l^{X\otimes Y}\right)\right]\!\right]_i \pm \sum_{r=1}^{m^{X\otimes Y}} z_r \left[\!\left[\phi_{X\otimes Y}\left(\left(C_r^{X\otimes Y}\right)^{T_{Y\otimes k}}\right)\right]\!\right]_i \geq 0\ , \\
&\qquad \sum_{s=1}^{m^Y} w_s \left[\!\left[\phi_{X\otimes Y}\left(Q\left(\mathrm{id}_{X^{\otimes k}}\otimes C_s^Y\right)Q^T\right)\right]\!\right]_i \pm \sum_{l=1}^{m^{X\otimes Y}} y_l \left[\!\left[\phi_{X\otimes Y}\left(\left(C_l^{X\otimes Y}\right)^{T_{Y\otimes k}}\right)\right]\!\right]_i \geq 0\ , \\
&\qquad \sum_{s=1}^{m^Y} w_s\,\mathrm{tr}(C_s^Y) \leq 1\ , \\
&\qquad \sum_{r=1}^{m^{X\otimes Y}} z_r \left[\!\left[\phi_{X\otimes Y}\left(C_r^{X\otimes Y}\right)\right]\!\right]_i \geq 0\ , \\
&\qquad \sum_{r=1}^{m^{X\otimes Y}} y_r \left[\!\left[\phi_{X\otimes Y}\left(C_r^{X\otimes Y}\right)\right]\!\right]_i \geq 0\ , \\
&\qquad \sum_{r=1}^{m^{X\otimes Y}} x_r \left[\!\left[\phi_{X\otimes Y}\left(C_r^{X\otimes Y}\right)\right]\!\right]_i \geq 0\ , \\
&\qquad \sum_{s=1}^{m^Y} w_s \left[\!\left[\phi_Y\left(C_s^Y\right)\right]\!\right]_t \geq 0\ , \\
&\qquad x_r, y_r, z_r, w_s, y \in \mathbb{R}, \quad \forall r \in [m^{X\otimes Y}], s \in [m^Y]\ ,
\end{aligned}
$$

where $j \in \left[t^X\right]$, $i \in \left[t^{X\otimes Y}\right]$ and $t \in \left[t^Y\right]$.

In Theorem 5.4.6, we showed how to efficiently compute $\phi_X(D_r)$, $\phi_{X\otimes Y}(C_r^{X\otimes Y})$, and $\phi_{X\otimes Y}(\left(J^{\mathcal{N}}\right)^{\otimes k})$. Note that $\phi_Y(C_s^Y)$ can be similarly computed in poly(k) time. Therefore, to complete the proof it suffices to show that $\phi_{X\otimes Y}\left(\left(C_r^{X\otimes Y}\right)^{T_{Y\otimes k}}\right)$, $\phi_{X\otimes Y}\left(Q^T(I_{X^{\otimes k}}\otimes C_r^Y)Q\right)$, and $\mathrm{tr}(C_s^Y)$ can computed in poly(k) time.

Recall that, for every $r \in \left[m^{X\otimes Y}\right]$, we have

$$
C_r^{X\otimes Y} = \sum_{(i,j)\in O_r^{X\otimes Y}} |i\rangle\langle j|\,,
$$

where $i = \left(i_1^X i_1^Y \cdots i_k^X i_k^Y\right)$ and $j = \left(j_1^X j_1^Y \cdots j_k^X j_k^Y\right)$. Therefore, we have

$$
\left(C_r^{X\otimes Y}\right)^{T_{Y\otimes k}} = \sum_{(i,j)\in O_r^{X\otimes Y}} |i_1^X j_1^Y \cdots i_k^X j_k^Y\rangle\langle j_1^X i_1^Y \cdots j_k^X i_k^Y| = C_{T(r)}^{X\otimes Y}\ ,
$$

where $T(r)$ denotes the index of the orbit given by

$$O^{X\otimes Y}_{T(r)} = \left\{ \left(i^X_1 j^Y_1 \cdots i^X_k j^Y_k , j^X_1 i^Y_1 \cdots j^X_k i^Y_k \right) \,:\, (i,j) \in O^{X\otimes Y}_r \right\} \quad .$$

Therefore, $\phi_{X\otimes Y}\left(\left(C^{X\otimes Y}_r \right)^{T_{Y\otimes k}} \right) = \phi_{X\otimes Y}\left(C^{X\otimes Y}_{T(r)} \right)$ can be computed efficiently.

For $r = 1, \ldots, m^{X\otimes Y}$, let (i,j) be an arbitrary representative element of $O^{X\otimes Y}_r$. Let

$$\alpha_r := (\mathrm{id}_{X^{\otimes k}})_{(i^X, j^X)} \cdot (C^Y_r)_{(i^Y, j^Y)} ,$$

where $i^X = \left(i^X_1 \ldots i^X_k \right)$, $i^Y = \left(i^Y_1 \ldots i^Y_k \right)$, and j^X and j^Y are defined in a similar way. Then we have $Q^T (I_{X^{\otimes k}} \otimes C^Y_r) Q = \sum_{r=1}^{m^{X\otimes Y}} \alpha_r C^{X\otimes Y}_r$, which implies that $\phi_{X\otimes Y}\left(Q^T (I_{X^{\otimes k}} \otimes C^Y_r) Q \right)$ can be computed in poly(k) time by Lemma 5.4.5.

Finally, for every $s \in [m^Y]$, we have

$$C^Y_s = \sum_{(i_1 \ldots i_k, j_1 \ldots j_k) \in O^Y_s} |i_1 \ldots i_k\rangle\langle j_1 \ldots j_k| \, .$$

Therefore, $\mathrm{tr}(C^Y_s) > 0$ iff $O^Y_s = \{(\pi(i), \pi(i)) : \pi \in \mathfrak{S}_k\}$, for some $i \in [d_Y]^k$. Let $s \in [m^Y]$ such that $\mathrm{tr}(C^Y_s) > 0$ and let $(i_1 \ldots i_k, i_1 \ldots i_k)$ be an arbitrary representative element of O^Y_s. For every $a \in [d_Y]$, define $\beta(a) := |\{v \in [k] : i_v = a\}|$, then $\mathrm{tr}(C^Y_s) = k! / \prod_{a \in [d_Y]} \beta(a)!$. □

As an example, $\Upsilon(\mathrm{D}^{\#}_2, 6)$ is computed for the amplitude damping (AD) channel $\mathcal{A}_{p,0}$, defined in Eq. (5.34), for different values of p. For this channel, the best previously known upper bound on the classical capacity $\mathrm{C}(\mathcal{A}_{p,0})$ for $p \in [0, 0.75]$ is given by quantity $\mathrm{C}_\beta(\mathcal{A}_{p,0}) = \log(1 + \sqrt{1-p})$ in [WFD17]. Table 5.3 shows that $\frac{1}{6}\Upsilon(\mathrm{D}^{\#}_2, 6)$ is a slightly improved upper bound compared to the bounds obtained using $\widehat{\mathrm{D}}_\alpha$ and $\mathrm{D}_{\max}$ which happen to coincide for the AD channel [FF21a] with the value $\log(1 + \sqrt{1-p})$. We remark that the best known upper bound for the AD channel $\mathcal{A}_{p,0}$ with $p \in [0.75, 1]$ is given by the entanglement-assisted classical capacity [BSST99] of the channel.

p	$\Upsilon(\mathrm{D}_{\max}, 1)$, $\Upsilon(\widehat{\mathrm{D}}_2, 1)$ and C_β	$\frac{1}{6}\Upsilon(\mathrm{D}^{\#}_2, 6)$
0.1	0.9626	**0.9615**
0.2	0.9218	**0.9201**
0.3	0.8770	**0.8745**
0.4	0.8274	**0.8239**
0.5	0.7716	**0.7670**
0.6	0.7071	**0.7014**
0.7	0.6302	**0.6234**
0.75	0.5850	**0.5777**

Table 5.3: Upper bounds on the classical capacity of the amplitude damping channel $\mathcal{A}_{p,0}$ with different parameters p.

5.6 Two-way assisted quantum capacity

In this section, we consider $\mathrm{D}_\alpha^\#$ in the framework of *generalized Theta-information* which was introduced in [FF21a]. As we will see, the generalized Theta-information induced by #-channel divergence gives efficiently computable strong converse bounds on the two-way-assisted quantum capacity, $\mathrm{Q}^{\leftrightarrow}(\mathcal{N})$, for any quantum channel $\mathcal{N}$.

The two-way assisted quantum capacity of a quantum channel $\mathcal{N}$ is the maximum rate at which quantum information can be transmitted reliably from a sender to a receiver, when the parties are allowed to perform arbitrary LOCC (short for local operations and classical communication) between consecutive channel uses [BDSW96]. While the two-way assisted quantum capacity for some specific channels such as the quantum erasure channel is known [BDS97], no general characterization of $\mathrm{Q}^{\leftrightarrow}(\mathcal{N})$ is known for an arbitrary quantum channel $\mathcal{N}$.

In [Rai99, Rai01], the authors relaxed the set LOCC to a larger class of operations known as PPT-preserving operations, which is the set of channels that are positive partial transpose preserving. A quantum channel $\mathcal{P}_{AB\to A'B'}$ is PPT-preserving if the linear map $T_{B'}\circ\mathcal{P}_{AB\to A'B'}\circ T_B$ is completely positive and trace-preserving [Rai01], where $T_{B'}$ and T_B denote the partial transpose map. For any quantum channel $\mathcal{N}$, we denote by $\mathrm{Q}^{\mathrm{PPT},\leftrightarrow}(\mathcal{N})$ the PPT-assisted quantum capacity of $\mathcal{N}$. In this case, the operations between the channel uses are allowed to be PPT-preserving operations. Because of the containment LOCC $\subset$ PPT [Rai01], we have the following inequality

$$\mathrm{Q}^{\leftrightarrow}(\mathcal{N}) \leq \mathrm{Q}^{\mathrm{PPT},\leftrightarrow}(\mathcal{N})\,,$$

for all quantum channels $\mathcal{N}$.

Inspired by the formulation of the Rains set [Rai01], in [FF21a] the authors introduced the set of subchannels given by the zero set of the Holevo-Werner bound [HW01] as

$$\Theta(X,Y) \coloneqq \{\mathcal{M}\in \mathrm{CP}(X:Y) : \exists R_{XY} \text{ s.t. } R_{XY} \pm (J_{XY}^{\mathcal{M}})^{T_Y} \geq 0, \mathrm{tr}_Y(R_{XY}) \leq \mathrm{id}_X\}\,.$$

Let $\mathbf{D}$ be a generalized divergence. For any quantum channel $\mathcal{N}_{X\to Y}$, define

$$R_\Theta(\mathbf{D},k)(\mathcal{N}) \coloneqq \mathbf{D}(\mathcal{N}^{\otimes k}\|\mathcal{M}^{\otimes k})\,,$$

where $\mathcal{M} = \mathrm{argmin}_{\mathcal{M}\in\Theta(X,Y)}\,\mathbf{D}(\mathcal{N}\|\mathcal{M})$.

For any quantum channel $\mathcal{N}$, by [FF21a, Theorem 17], [FF21b, Proposition 5.9], [BW18, Corollary 5] and the relation between the divergences in Proposition 5.2.1, the following holds:

Proposition 5.6.1. *Let $\mathcal{N}$ be a quantum channel. For any $\alpha\in(1,2]$ and $k\geq 1$,*

$$\mathrm{Q}^{\leftrightarrow}(\mathcal{N}) \leq \mathrm{Q}^{\mathrm{PPT},\leftrightarrow}(\mathcal{N}) \leq \mathrm{Q}^{\mathrm{PPT},\leftrightarrow,\dagger}(\mathcal{N}) \leq \frac{1}{k}R_\Theta(\mathrm{D}_\alpha^\#,k)(\mathcal{N}) \leq R_\Theta(\widehat{\mathrm{D}}_\alpha,1)(\mathcal{N}) \leq R_\Theta(\mathrm{D}_{\max},1)(\mathcal{N})\,,$$

where $\mathrm{Q}^{\mathrm{PPT},\leftrightarrow,\dagger}(\mathcal{N})$ is the strong converse capacity corresponding to $\mathrm{Q}^{\mathrm{PPT},\leftrightarrow}(\mathcal{N})$.

The squashed entanglement of the channel $\mathcal{N}$ introduced in [TGW14] is known to be a converse bounds for $\mathrm{Q}^{\mathrm{PPT},\leftrightarrow}(\mathcal{N})$. However, it remains open whether it is a strong converse and the quantity itself is NP-hard to compute [Hua14]. Using a similar method

as in Section 5.5, we can show that $R_\Theta(\mathrm{D}^\#_\alpha, k)(\mathcal{N})$ can be computed in poly(k) time for any quantum channel $\mathcal{N}$.

As an example, $R_\Theta(\mathrm{D}^\#_\alpha, 6)$ is computed for the qubit amplitude damping channel $\mathcal{A}_{p,0}$, defined in Eq. (5.34), for values of $p \in [0, 1]$. The comparison between the two-way/PPT assisted quantum capacity is given in Figure 5.2. The bound $\frac{1}{6}R_\Theta(\mathrm{D}^\#_2, 6)$ demonstrates an improvement compared to the best previously known strong converse bound given by $R_\Theta(\widehat{D}_2, 1)$.

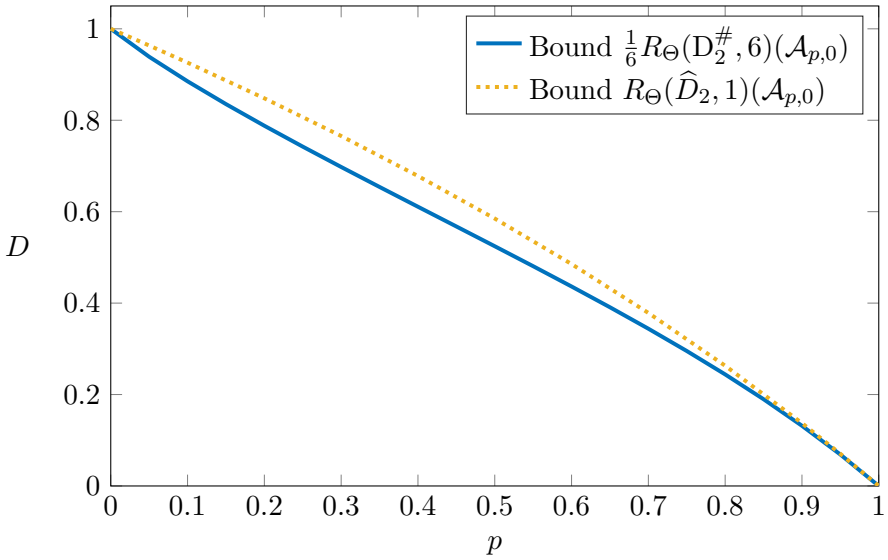

Figure 5.2: Comparison between two strong converse bounds $R_\Theta(\widehat{D}_2, 1)$ and $\frac{1}{6}R_\Theta(\mathrm{D}^\#_2, 6)$ on for two-way/PPT assisted quantum capacity for the qubit amplitude damping channel $\mathcal{A}_{p,0}$ for $p \in [0, 1]$

5.7 Conclusion

Optimal rates for achieving an information processing task are often characterized in terms of regularized information measures. In many cases of quantum tasks, we do not know how to compute such quantities. Here, we exploited the symmetries in $\mathrm{D}^\#$ in order to obtain a hierarchy of semidefinite programming bounds on various regularized quantities. As applications, we gave a general procedure to give efficient bounds on the regularized Umegaki channel divergence as well as the classical capacity and two-way assisted quantum capacity of quantum channels. In particular, we obtained slight improvements for the capacity of the amplitude damping channel. We also proved that for fixed input and output dimensions, the regularized sandwiched Rényi divergence between any two quantum channels can be approximated up to an ϵ accuracy in time that is polynomial in $1/\epsilon$.

Chapter 6

Conclusion

In this thesis, we considered the asymptotic growth behavior of a property for the powers of a fixed object for some types of objects.

The first objects we considered are hypergraphs equipped with the strong product operation and the property of interest is the independence number. The asymptotic growth of the independence number of a hypergraph is known as the Shannon capacity. We introduced the combinatorial degeneration method for finding lower bounds for the Shannon capacity of directed k-uniform hypergraph in Theorem 3.3.3. We then applied this method to improve the lower bound for the corner, square, Lshape, which are special cases of the generalized multidimensional Szemerédi problem in Theorem 3.3.6 and Table 3.1. Moreover, in Corollary 3.4.6, we pointed out how induced matchings in hypergraphs pose a barrier for existing tensor tools (such as slice rank, subrank, analytic rank, geometric rank and G-stable rank) to efficiently obtain an upper bound on the size of independent sets in hypergraphs. This implies a barrier for these tools to effectively establish lower bounds on the communication complexity on the NOF model of the Eval function over any group G.

Tensors are the second considered objects. We equipped them with the tensor product and the property of interest is the symmetric subrank. The symmetric subrank is a notion we introduce motivated by limitations of current tensor methods to bound the Shannon capacity of hypergraphs. In Section 4.2, we presented and proved precise relations and separations between subrank and symmetric subrank. Then, in Theorem 4.3.4, we showed that for symmetric tensors the subrank and the symmetric subrank are asymptotically equal. This proves the asymptotic subrank analogon of a conjecture known as Comon's conjecture in the theory of tensors. This result allows us to prove a strong connection between the general and the symmetric versions of an asymptotic duality theorem of Strassen. Finally, in Section 4.5, we introduced a representation-theoretic method to asymptotically bound the symmetric subrank called the symmetric quantum functional in analogy with the quantum functionals, then studied the relations between these functionals. Nevertheless, the symmetric quantum functional cannot give better bounds than the quantum functionals which itself suffers from the induced matching barrier and cannot be used to make progress on the corner problem. But we hope that future improved asymptotic upper bounds on the symmetric subrank can still overcome the induced matching barrier. In particular, we leave it as an open question to define a good symmetric version

of Strassen's support functionals.

Our last considered objects are quantum channels whose operation and property are the tensor product of channels and the divergence, respectively. Namely, we used the recently introduced $\mathrm{D}^{\#}$ Rényi channel divergence [FF21b] as the property of interest. We exploited the symmetries in $\mathrm{D}^{\#}$ in order to obtain a hierarchy of semidefinite programming bounds on various regularized quantities. As applications, in Section 5.5 and 5.6 we gave a general procedure to give efficient bounds on the regularized Umegaki channel divergence as well as the classical capacity and two-way assisted quantum capacity of quantum channels. In particular, we obtained slight improvements for the capacity of the amplitude damping channel. In Section 5.4, we proved that for fixed input and output dimensions, the regularized sandwiched Rényi divergence between any two quantum channels can be approximated up to an ϵ accuracy in time that is polynomial in $1/\epsilon$.

Appendix A

Deferred proofs

A.1 Representation-theoretic characterization of the moment polytope

In this section we prove Lemma 4.5.10.

We recall some notions and results of geometric invariant theory and representation theory. We refer to [NM84], [Bri87], [Fra02], [Wal14], and [BFG+19] for more information. Let $\mathrm{GL}(d)$ be the group of $d \times d$ invertible matrices over the complex numbers. Let $\mathcal{H}$ be a complex finite-dimensional vector space, with $\dim(\mathcal{H}) = d$. Denote by $M(d)$ the set of complex $d \times d$ matrices, and denote by $\mathrm{Herm}(d)$ the set of $d \times d$ Hermitian matrices. We define the representation π of $\mathrm{GL}(d)$ on $\mathcal{H}^{\otimes k}$ by $\pi(g)f \coloneqq (g \otimes \cdots \otimes g)f$ for all $g \in \mathrm{GL}(d)$ and $f \in \mathcal{H}^{\otimes k}$. Let $\mathrm{GL}(d) \cdot f \coloneqq \{\pi(g)f : g \in \mathrm{GL}(d)\}$ denote the orbit of f under the action of $\mathrm{GL}(d)$. For any nonzero vector $f \in \mathcal{H}^{\otimes k}$, we define the function:

$$\begin{aligned} F_f : \quad & \mathrm{GL}(d) \to \mathbb{R} \\ & g \mapsto \frac{1}{2} \log \|\pi(g)f\|^2. \end{aligned}$$

The following definition defines the gradient of F_f at $g = I$.

Definition A.1.1. The *moment map* is the function $\mu : \mathcal{H}^{\otimes k} \setminus \{0\} \to \mathrm{Herm}(d)$ defined by the property that for all $H \in \mathrm{Herm}(d)$ we have $\mathrm{tr}[\mu(f)H] = \partial_{t=0} F_f(e^{tH})$.

Let $H \in \mathrm{Herm}(d)$. Then $\partial_{t=0} F_f(e^{tH}) = \partial_{t=0} \frac{\langle f, \pi(e^{tH})f \rangle}{\|f\|^2}$. Therefore, we have

$$\begin{aligned} \mathrm{tr}[\mu(f)H] &= \partial_{t=0} \frac{\langle f, \pi(e^{tH})f \rangle}{\|f\|^2} \\ &= \frac{\langle f, (\sum_{j=1}^{k} I^{\otimes j-1} \otimes H \otimes I^{\otimes n-j})f \rangle}{\|f\|^2} \\ &= \sum_{j=1}^{k} \mathrm{tr}\left[\frac{ff^\dagger}{\|f\|^2} (I^{\otimes j-1} \otimes H \otimes I^{\otimes n-j}) \right] \\ &= \sum_{j=1}^{k} \mathrm{tr}[\rho_j(f)H], \end{aligned}$$

where $\rho_j(f)$ denotes the jth reduced density matrix of $\rho(f) = \frac{ff^\dagger}{\|f\|^2}$. Thus, $\mu(f) = \sum_{j=1}^k \rho_j(f)$.

Following [FH91], any rational irreducible representations of $\mathrm{GL}(d)$ can be labeled by highest weight $\lambda \in \mathbb{N}^d$ such that $\lambda_1 \geq \cdots \geq \lambda_d$. For any natural number $n \geq 1$, consider the representation Π of $\mathrm{GL}(d)$ on $(\mathcal{H}^{\otimes k})^{\otimes n}$ by $\Pi(g) \cdot v := (\pi(g) \otimes \cdots \otimes \pi(g))v$ for all $v \in (\mathcal{H}^{\otimes k})^{\otimes n}$. Let V be a finite-dimensional rational representation of $\mathrm{GL}(d)$. For each highest weight λ of $\mathrm{GL}(d)$, we denote by V_λ the λ-isotypical component of V. Let $Z \subseteq V$ be a Zariski closed set. We denote by $\mathbb{C}[Z]_n$ the degree-n part of the *coordinate ring* of Z. Letting $\lambda = (\lambda_1, \ldots, \lambda_d)$ be a highest weight of $\mathrm{GL}(d)$, we define $\lambda^* = (-\lambda_d, \ldots, -\lambda_1)$. For any nonzero vector $f \in \mathcal{H}^{\otimes k}$, the following lemma says that the moment polytope $\Delta(f)$ has another representation theoretic description.

Lemma A.1.2 ([Bri87], [Fra02], [Str05, Theorem 11] or [Zui18, Chapter 6])**.** *Let $f \in \mathcal{H}^{\otimes k}$ be nonzero. Then*

$$\begin{aligned}\Delta(f) &= \overline{\left\{\lambda/n : \exists n \in \mathbb{N}_{\geq 1}, (\mathbb{C}[\overline{\mathrm{GL}(d) \cdot f}]_n)_{\lambda^*} \neq 0\right\}} \\ &= \overline{\{\lambda/n : \exists n \in \mathbb{N}_{\geq 1}, P_\lambda f^{\otimes n} \neq 0\}},\end{aligned}$$

where P_λ is the projector from $(\mathcal{H}^{\otimes k})^{\otimes n}$ onto the λ-isotypical component in the decomposition of $(\mathcal{H}^{\otimes k})^{\otimes n}$ with respect to Π.

Proof of Theorem 4.5.10. By Schur–Weyl duality we have a decomposition of the space $(\mathcal{H}^{\otimes k})^{\otimes n}$ as

$$(\mathcal{H}^{\otimes k})^{\otimes n} \cong \bigoplus_{\lambda \vdash_d kn} \mathbb{S}_\lambda(\mathcal{H}) \otimes [\lambda].$$

For $\lambda \vdash_d kn$, let P_λ be the projector onto the isotypical component of type λ, that is, onto the subspace of $(\mathcal{H}^{\otimes k})^{\otimes n}$ which isomorphic to $\mathbb{S}_\lambda(\mathcal{H}) \otimes [\lambda]$, since all irreducible representations of Π are labeled by the partitions of kn in at most d parts. Therefore,

$$\Delta(f) = \overline{\left\{\frac{\lambda}{n} : \exists n \geq \mathbb{N}_{\geq 1}, \lambda \vdash_d kn, P_\lambda f^{\otimes n} \neq 0\right\}},$$

completing the proof. □

A.2 Sub-multiplicativity of the symmetric quantum functional

In this section we prove that the symmetric quantum functional F is sub-multiplicative. For symmetric tensors this follows from Theorem 4.5.2. (In fact, Theorem 4.5.2 says that the symmetric quantum functional is multiplicative on symmetric tensors.) Here we prove that the symmetric quantum functional is sub-multiplicative on arbitrary tensors (not necessarily symmetric). The argument is an adaptation of the argument in [CVZ18] to the symmetric quantum functional.

Lemma A.2.1. *For all tensors $s \in V^{\otimes k}$ and $t \in W^{\otimes k}$ we have $\Delta(s \otimes t) \subseteq \Delta(s) \otimes_{\mathrm{Kron}} \Delta(t)$ where*

$$\Delta(s) \otimes_{\mathrm{Kron}} \Delta(t) \coloneqq \mathrm{closure}\big\{\bar{\mu} : \bar{\lambda} \in \Delta(s), \bar{\lambda}' \in \Delta(t), P_\mu(P_\lambda \otimes P_{\lambda'}) \neq 0\big\}.$$

Proof. Let $\dim(V) = d$ and $\dim(W) = d'$. If $\bar{\mu} \in \Delta(s \otimes t)$, then for some n, we have $P_\mu(s \otimes t)^{\otimes n} \neq 0$. We have $\sum_{\lambda \vdash_d kn} P_\lambda = \mathrm{Id}_{V^{\otimes kn}}$ and $\sum_{\lambda' \vdash_{d'} kn} P_{\lambda'} = \mathrm{Id}_{W^{\otimes kn}}$. Thus, we can write

$$P_\mu(s \otimes t)^{\otimes n} = P_\mu\Big(\sum_{\lambda,\lambda'} P_\lambda \otimes P_{\lambda'}\Big)(s \otimes t)^{\otimes n}.$$

So there exists λ, λ' such that $P_\mu(P_\lambda \otimes P_{\lambda'})(s \otimes t)^{\otimes n} \neq 0$. But this implies that $P_\lambda s^{\otimes n} \neq 0$, $P_{\lambda'} t^{\otimes n} \neq 0$, and $P_\mu(P_\lambda \otimes P_{\lambda'}) \neq 0$, which completes the proof. □

Proposition A.2.2 (Sub-multiplicativity of the symmetric quantum functional)**.** *For every $s \in V^{\otimes k}$ and $t \in W^{\otimes k}$ we have $F(s \otimes t) \leq F(s)F(t)$.*

Proof. Let $d = \dim(V)$ and $d' = \dim(W)$. Let $E = \log_2 F$. We need to prove $E(s \otimes t) \leq E(s) + E(t)$. By definition

$$E(s \otimes t) = \max_{p \in \Delta(s \otimes t)} H(p) \leq \max_{p \in \Delta(s) \otimes_{\mathrm{Kron}} \Delta(t)} H(p).$$

But if $p \in \Delta(s) \otimes_{\mathrm{Kron}} \Delta(t)$, then there exists μ a partition of kn in at most dd' parts such that $P_\mu(P_\lambda \otimes P_{\lambda'}) \neq 0$ with $\bar{\lambda} \in \Delta(s)$ and $\bar{\lambda}' \in \Delta(t)$ by Lemma A.2.1. It is shown in [CM06, Proposition 3] that if $P_\mu(P_\lambda \otimes P_{\lambda'}) \neq 0$, then $H(\bar{\mu}) \leq H(\bar{\lambda}) + H(\bar{\lambda}')$. This shows that $E(s \otimes t) \leq E(s) + E(t)$. □

A.3 Technical lemmas

Lemma A.3.1 (Fekete's lemma, see [PS72])**.** *Let $x_1, x_2 \cdots \in \mathbb{R}_{\geq 0}$ 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 $\epsilon > 0$. Let $m \in \mathbb{N}$ with $x_m/m < y + \epsilon$. Any $n\mathbb{N}$ 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 $x_n = x_{qm+r} \leq x_m + x_m + \cdots + x_m + x_r = qx_m + r$. 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

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

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

Bibliography

[ACFN15] Anil Ada, Arkadev Chattopadhyay, Omar Fawzi, and Phuong Nguyen. The NOF multiparty communication complexity of composed functions. *Computational Complexity*, 24(3):645–694, 2015. doi:10.1007/s00037-013-0078-4. 10, 25, 32, 33, 38, 45

[Aci01] Antonio Acin. Statistical distinguishability between unitary operations. *Physical review letters*, 87(17):177901, 2001. 12

[Alm19] Josh Alman. Limits on the universal method for matrix multiplication. In *Computational Complexity Conference*, 2019. 15

[Alo98] Noga Alon. The shannon capacity of a union. *Combinatorica*, 18(3):301–310, 1998. 18

[ApS19] MOSEK ApS. *The MOSEK optimization toolbox for MATLAB manual. Version 9.0.*, 2019. URL: http://docs.mosek.com/9.0/toolbox/index.html. 92

[AS20] Noga Alon and Adi Shraibman. Algorithmic number on the forehead protocols yielding dense Ruzsa-Szemerédi graphs and hypergraphs, 2020. arXiv:2001.00387. 28

[ASU13] Noga Alon, Amir Shpilka, and Christopher Umans. On sunflowers and matrix multiplication. *computational complexity*, 22(2):219–243, 2013. doi:10.1007/s00037-013-0060-1. 9, 35

[AW18] Josh Alman and Virginia Vassilevska Williams. Further limitations of the known approaches for matrix multiplication. In *9th Innovations in Theoretical Computer Science Conference (ITCS 2018)*, pages 25:1–25:15, 2018. doi:10.4230/LIPIcs.ITCS.2018.25. 35

[AW21] Josh Alman and Virginia Vassilevska Williams. A refined laser method and faster matrix multiplication. In Dániel Marx, editor, *Proceedings of the 2021 ACM-SIAM Symposium on Discrete Algorithms, SODA 2021, Virtual Conference, January 10 - 13, 2021*, pages 522–539. SIAM, 2021. doi:10.1137/1.9781611976465.32. 8, 16, 44

[Bal68] C. S. Ballantine. Triangularizing matrices by congruence. *Linear Algebra and its Applications*, 1(2):261 – 280, 1968. doi:10.1016/0024-3795(68)90007-4. 58

[BB07] Mark Bauer and Michael A Bennett. On a question of erdos and graham. *ENSEIGNEMENT MATHEMATIQUE*, 53(3/4):259, 2007. 34

[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. 16

[BCS97] Peter Bürgisser, Michael Clausen, and M. Amin Shokrollahi. *Algebraic complexity theory*, volume 315 of *Grundlehren der Mathematischen Wissenschaften*. Springer-Verlag, Berlin, 1997. doi:10.1007/978-3-662-03338-8. 35

[BDPW07] Paul Beame, Matei David, Toniann Pitassi, and Philipp Woelfel. Separating deterministic from nondeterministic NOF multiparty communication complexity. In *International Colloquium on Automata, Languages, and Programming (ICALP 2007)*, pages 134–145, 2007. doi:10.1007/978-3-540-73420-8_14. 25, 32

[BDS97] Charles H Bennett, David P DiVincenzo, and John A Smolin. Capacities of quantum erasure channels. *Physical Review Letters*, 78(16):3217, 1997. 100

[BDSW96] Charles H Bennett, David P DiVincenzo, John A Smolin, and William K Wootters. Mixed-state entanglement and quantum error correction. *Physical Review A*, 54(5):3824, 1996. 100

[BFG+19] Peter Bürgisser, Cole Franks, Ankit Garg, Rafael Mendes de Oliveira, Michael Walter, and Avi Wigderson. Towards a theory of non-commutative optimization: Geodesic 1st and 2nd order methods for moment maps and polytopes. In *Proceedings of the 60th IEEE Annual Symposium on Foundations of Computer Science (FOCS 2019)*, pages 845–861, 2019. doi:10.1109/FOCS.2019.00055. 104

[BGG06] Richard Beigel, William Gasarch, and James Glenn. The multiparty communication complexity of Exact-T: Improved bounds and new problems. In *International Symposium on Mathematical Foundations of Computer Science (MFCS 2006)*, pages 146–156, 2006. doi:10.1007/11821069_13. 25, 32, 33

[BGKL04] László Babai, Anna Gál, Peter G. Kimmel, and Satyanarayana V. Lokam. Communication complexity of simultaneous messages. *SIAM J. Comput.*, 33(1):137–166, 2004. doi:10.1137/S0097539700375944. 25, 32

[BK12] Michael Bateman and Nets Hawk Katz. New bounds on cap sets. *J. Amer. Math. Soc.*, 25(2):585–613, 2012. doi:10.1090/S0894-0347-2011-00725-X. 33

[BL96] Vitaly Bergelson and Alexander Leibman. Polynomial extensions of van der waerden's and szemerédi's theorems. *Journal of the American Mathematical Society*, 9(3):725–753, 1996. 29, 30

[BM85] E. Bidamon and H. Meyniel. On the Shannon capacity of a directed graph. *European J. Combin.*, 6(4):289–290, 1985. doi:10.1016/S0195-6698(85)80042-1. 42, 43

[Boh03] Tom Bohman. A limit theorem for the shannon capacities of odd cycles i. *Proceedings of the American Mathematical Society*, 131(11):3559–3569, 2003. 18

[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. 69, 70, 104, 105

[Bri19] Jop Briët. Subspaces of tensors with high analytic rank, 2019. arXiv:1908.04169. 29

[BS06] Genrich Belitskii and Vladimir Sergeichuk. Congruence of multilinear forms. *Linear Algebra and its Applications*, 418(2):751–762, 2006. doi:10.1016/j.laa.2006.03.003. 59

[BSST99] Charles H. Bennett, Peter W. Shor, John A. Smolin, and Ashish V. Thapliyal. Entanglement-assisted classical capacity of noisy quantum channels. *Phys. Rev. Lett.*, 83:3081–3084, Oct 1999. URL: https://link.aps.org/doi/10.1103/PhysRevLett.83.3081, doi:10.1103/PhysRevLett.83.3081. 99

[BT94] Richard Beigel and Jun Tarui. On ACC. *Computational Complexity*, 4(4):350–366, 1994. doi:10.1007/BF01263423. 25, 32

[Bür90] Peter Bürgisser. *Degenerationsordnung und Trägerfunktional bilinearer Abbildungen*. PhD thesis, Universität Konstanz, 1990. http://nbn-resolving.de/urn:nbn:de:bsz:352-opus-20311. 53, 65, 67

[BW18] Mario Berta and Mark M Wilde. Amortization does not enhance the max-rains information of a quantum channel. *New Journal of Physics*, 20(5):053044, 2018. 100

[CFL83] Ashok K. Chandra, Merrick L. Furst, and Richard J. Lipton. Multi-party protocols. In *Proceedings of the 15th Annual ACM Symposium on Theory of Computing (STOC 1983)*, pages 94–99, 1983. doi:10.1145/800061.808737. 9, 25, 32, 33

[CFTZ21] Matthias Christandl, Omar Fawzi, Hoang Ta, and Jeroen Zuiddam. Symmetric subrank of tensors and applications. *arXiv preprint arXiv:2104.01130*, 2021. 5, 52

[CFTZ22] Matthias Christandl, Omar Fawzi, Hoang Ta, and Jeroen Zuiddam. Larger Corner-Free Sets from Combinatorial Degenerations. In Mark Braverman, editor, *13th Innovations in Theoretical Computer Science Conference (ITCS 2022)*, volume 215 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 48:1–48:20, Dagstuhl, Germany, 2022. Schloss Dagstuhl –

Leibniz-Zentrum für Informatik. URL: https://drops.dagstuhl.de/opus/volltexte/2022/15644, doi:10.4230/LIPIcs.ITCS.2022.48. 5, 24

[CGLM08] Pierre Comon, Gene Golub, Lek-Heng Lim, and Bernard Mourrain. Symmetric tensors and symmetric tensor rank. *SIAM Journal on Matrix Analysis and Applications*, 30(3):1254–1279, 2008. doi:10.1137/060661569. 11, 52, 54, 56, 58, 62

[CKSU05] H. Cohn, R. Kleinberg, B. Szegedy, and C. Umans. Group-theoretic algorithms for matrix multiplication. In *Proceedings of the 46th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2005)*, pages 379–388, 2005. doi:10.1109/SFCS.2005.39. 9

[CLP17] Ernie Croot, Vsevolod F. Lev, and Péter Pál Pach. Progression-free sets in $\mathbb{Z}_4^n$ are exponentially small. *Annals of Mathematics*, pages 331–337, 2017. doi:10.4007/annals.2017.185.1.7. 9

[CLZ20] Matthias Christandl, Vladimir Lysikov, and Jeroen Zuiddam. Weighted slice rank and a minimax correspondence to strassen's spectra, 2020. arXiv:2012.14412. 67

[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. 106

[CMT18] Brian Cook, Ákos Magyar, and Tatchai Titichetrakun. A multidimensional szemerédi theorem in the primes via combinatorics. *Annals of Combinatorics*, 22(4):711–768, 2018. 30

[CMW16] Tom Cooney, Milán Mosonyi, and Mark M Wilde. Strong converse exponents for a quantum channel discrimination problem and quantum-feedback-assisted communication. *Communications in Mathematical Physics*, 344(3):797–829, 2016. 12

[CPR00] Andrew M Childs, John Preskill, and Joseph Renes. Quantum information and precision measurement. *Journal of modern optics*, 47(2-3):155–176, 2000. 12

[CS14] Arkadev Chattopadhyay and Michael E. Saks. The power of super-logarithmic number of players. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2014)*, volume 28, pages 596–603, 2014. doi:10.4230/LIPIcs.APPROX-RANDOM.2014.596. 25, 32

[CVZ18] Matthias Christandl, Péter Vrana, and Jeroen Zuiddam. Universal points in the asymptotic spectrum of tensors. In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing (STOC 2018)*, pages 289–296, 2018. arXiv:1709.07851, doi:10.1145/3188745.3188766. 11, 12, 15, 16, 28, 35, 46, 53, 54, 65, 67, 68, 69, 70, 71, 105

[CVZ19a] Matthias Christandl, Péter Vrana, and Jeroen Zuiddam. Asymptotic tensor rank of graph tensors: beyond matrix multiplication. *Comput. Complex.*, 28(1):57–111, 2019. arXiv:1609.07476, doi:10.1007/s00037-018-0172-8. 47

[CVZ19b] Matthias Christandl, Péter Vrana, and Jeroen Zuiddam. Barriers for fast matrix multiplication from irreversibility. In *Computational Complexity Conference*, 2019. 15

[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, 1987. 44, 47

[Der20] Harm Derksen. The G-stable rank for tensors, 2020. arXiv:2002.08435. 27, 29

[DLS20] Dong Dong, Xiaochun Li, and Will Sawin. Improved estimates for polynomial roth type theorems in finite fields. *Journal d'Analyse Mathématique*, 141(2):689–705, 2020. 30

[Ede04] Yves Edel. Extensions of generalized product caps. *Designs, Codes and Cryptography*, 31(1):5–14, 2004. doi:10.1023/A:1027365901231. 33, 37

[EG17] Jordan S. Ellenberg and Dion Gijswijt. On large subsets of $\mathbb{F}_q^n$ with no three-term arithmetic progression. *Ann. of Math.*, 185(1):339–343, 2017. doi:10.4007/annals.2017.185.1.8. 9, 33

[Erd73] Paul Erdös. Problems and results on combinatorial number theory. In *A survey of combinatorial theory*, pages 117–138. Elsevier, 1973. 34

[FF21a] Kun Fang and Hamza Fawzi. Geometric rényi divergence and its applications in quantum channel capacities. *Communications in Mathematical Physics*, pages 1–63, 2021. 74, 75, 76, 81, 95, 99, 100

[FF21b] Hamza Fawzi and Omar Fawzi. Defining quantum divergences via convex optimization. *Quantum*, 5:387, 2021. 12, 13, 74, 75, 76, 77, 81, 100, 103

[FFRS20] Kun Fang, Omar Fawzi, Renato Renner, and David Sutter. Chain rule for the quantum relative entropy. *Physical Review Letters*, 124(10):100501, 2020. 12, 13, 73, 74, 75, 81, 92, 95

[FGW22] Kun Fang, Gilad Gour, and Xin Wang. Towards the ultimate limits of quantum channel discrimination. *arXiv preprint arXiv:2110.14842*, 2022. 12, 13

[FH91] William Fulton and Joe Harris. *Representation theory*, volume 129 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1991. doi:10.1007/978-1-4612-0979-9. 105

[FH13] William Fulton and Joe Harris. *Representation theory: a first course*, volume 129. Springer Science & Business Media, 2013. 20

[FK79] Hillel Furstenberg and Yitzhak Katznelson. An ergodic szemer'edi theorem for commuting transformations. In *J. Analyse Math.* Citeseer, 1979. 9, 26, 29

[FK00] Emanuela Fachini and János Körner. Colour number, capacity, and perfectness of directed graphs. *Graphs Combin.*, 16(4):389–398, 2000. doi:10.1007/PL00007226. 49

[Fra02] Matthias Franz. Moment polytopes of projective G-varieties and tensor products of symmetric group representations. *J. Lie Theory*, 12(2):539–549, 2002. 70, 104, 105

[FS17] Hamza Fawzi and James Saunderson. Lieb's concavity theorem, matrix geometric means, and semidefinite optimization. *Linear Algebra and its Applications*, pages 240–263, 2017. 77

[FSP18] Hamza Fawzi, James Saunderson, and Pablo A. Parrilo. Semidefinite approximations of the matrix logarithm. *Foundations of Computational Mathematics*, 2018. Package cvxquad at https://github.com/hfawzi/cvxquad. 92

[FST21] Omar Fawzi, Ala Shayeghi, and Hoang Ta. A hierarchy of efficient bounds on quantum capacities exploiting symmetry. In *2021 IEEE International Symposium on Information Theory (ISIT)*, pages 272–277. IEEE, 2021. 5, 73

[FZ15] Jacob Fox and Yufei Zhao. A short proof of the multidimensional szemerédi theorem in the primes. *American Journal of Mathematics*, pages 1139–1145, 2015. 30

[GB14] Michael Grant and Stephen Boyd. CVX: Matlab software for disciplined convex programming, version 2.1. http://cvxr.com/cvx, March 2014. 92

[Gij05] Dion Gijswijt. Matrix algebras and semidefinite programming techniques for codes. *PhD thesis, University of Amsterdam, Netherlands*, July 2005. 78

[Gij09] Dion Gijswijt. Block diagonalization for algebra's associated with block codes. *arXiv preprint arXiv:0910.4515*, 2009. 86, 89

[GKV92] L. Gargano, J. Körner, and U. Vaccaro. Qualitative independence and sperner problems for directed graphs. *Journal of Combinatorial Theory, Series A*, 61(2):173–192, 1992. doi:10.1016/0097-3165(92)90016-N. 24

[GKV93] L. Gargano, J. Körner, and U. Vaccaro. Sperner capacities. *Graphs and Combinatorics*, 9(1):31–46, 1993. doi:10.1007/BF01195325. 24

[GLS12] Martin Grötschel, László Lovász, and Alexander Schrijver. *Geometric algorithms and combinatorial optimization*, volume 2. Springer Science & Business Media, 2012. 22

[Gow07] W Timothy Gowers. Hypergraph regularity and the multidimensional szemerédi theorem. *Annals of Mathematics*, pages 897–946, 2007. 29

[Gro94] Vince Grolmusz. The BNS lower bound for multi-party protocols in nearly optimal. *Inf. Comput.*, 112(1):51–54, 1994. doi:10.1006/inco.1994.1051. 25, 32

[GT08] Ben Green and Terence Tao. The primes contain arbitrarily long arithmetic progressions. *Annals of mathematics*, pages 481–547, 2008. 29

[GW09] Roe Goodman and Nolan R. Wallach. *Symmetry, representations, and invariants*, volume 255. Springer, 2009. 64

[GW11] W. T. Gowers and J. Wolf. Linear forms and higher-degree uniformity for functions on $\mathbb{F}_p^n$. *Geom. Funct. Anal.*, 21(1):36–69, 2011. doi:10.1007/s00039-010-0106-3. 27, 29

[Hae79] Willem H. Haemers. On some problems of lovász concerning the shannon capacity of a graph. *IEEE Trans. Inf. Theory*, 25(2):231–232, 1979. doi:10.1109/TIT.1979.1056027. 18, 27

[Has09] Matthew B Hastings. Superadditivity of communication capacity using entangled inputs. *Nature Physics*, 5(4):255–257, 2009. 74, 93

[HLY21] Rui Han, Michael T Lacey, and Fan Yang. A polynomial roth theorem for corners in finite fields. *Mathematika*, 67(4):885–896, 2021. 30

[HM17] Fumio Hiai and Milán Mosonyi. Different quantum f-divergences and the reversibility of quantum operations. *Reviews in Mathematical Physics*, 29(07):1750023, 2017. 76

[Hol98] Alexander S. Holevo. The capacity of the quantum channel with general signal states. *IEEE Transactions on Information Theory*, 44(1):269–273, 1998. doi:10.1109/18.651037. 93

[Hua14] Yichen Huang. Computing quantum discord is np-complete. *New journal of physics*, 16(3):033027, 2014. 100

[HW01] Alexander S Holevo and Reinhard F Werner. Evaluating capacities of bosonic gaussian channels. *Physical Review A*, 63(3):032312, 2001. 100

[IK99] Anthony Iarrobino and Vassil Kanev. *Power sums, Gorenstein algebras, and determinantal loci*, volume 1721 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1999. doi:10.1007/BFb0093426. 60

[KA80] Fumio Kubo and Tsuyoshi Andô. Means of positive linear operators. *Mathematische Annalen*, 246:205–224, 1980. 76

[Kar72] Richard M. Karp. *Reducibility among Combinatorial Problems*, pages 85–103. Springer US, Boston, MA, 1972. doi:10.1007/978-1-4684-2001-2_9. 7, 18

[Kay12] Neeraj Kayal. Affine projections of polynomials: extended abstract. In *Proceedings of the 44th Symposium on Theory of Computing Conference, (STOC 2012)*, pages 643–662, 2012. doi:10.1145/2213977.2214036. 62

[KKN95] Mauricio Karchmer, Eyal Kushilevitz, and Noam Nisan. Fractional covers and communication complexity. *SIAM Journal on Discrete Mathematics*, 8(1):76–92, 1995. 9

[KM90] János Körner and Katalin Marton. On the capacity of uniform hypergraphs. *IEEE Trans. Inf. Theory*, 36(1):153–156, 1990. doi:10.1109/18.50381. 28

[KMZ20] Swastik Kopparty, Guy Moshkovitz, and Jeroen Zuiddam. Geometric Rank of Tensors and Subrank of Matrix Multiplication. In *Proceedings of the 35th Computational Complexity Conference (CCC 2020)*, pages 35:1–35:21, 2020. arXiv:2002.09472, doi:10.4230/LIPIcs.CCC.2020.35. 27, 29

[KPS05] János Körner, Concetta Pilotto, and Gábor Simonyi. Local chromatic number and sperner capacity. *J. Comb. Theory, Ser. B*, 95(1):101–117, 2005. doi:10.1016/j.jctb.2005.03.006. 49, 50

[KPS07] Etienne Klerk, Dmitrii Pasechnik, and Alexander Schrijver. Reduction of symmetric semidefinite programs using the regular $*$-representation. *Math. Program.*, 109:613–624, March 2007. doi:10.1007/s10107-006-0039-7. 79, 85

[KSS16] Robert Kleinberg, William Sawin, and David Speyer. The growth rate of tri-colored sum-free sets. *Discrete Anal.*, 2016. doi:10.19086/da.3734. 29

[KSV02] Alexei Yu. Kitaev, Alexander H. Shen, and Mikhail N. Vyalyi. *Classical and Quantum Computation*. American Mathematical Society, USA, 2002. 73

[Kuc19] Borys Kuca. Further bounds in the polynomial szemer\'{e} di theorem over finite fields. *arXiv preprint arXiv:1907.08446*, 2019. 28, 30

[Kuc21] Borys Kuca. Multidimensional polynomial szemer\'{e} di theorem in finite fields for polynomials of distinct degrees. *arXiv preprint arXiv:2103.12606*, 2021. 28, 30

[LG14] François Le Gall. Powers of tensors and fast matrix multiplication. In *Proceedings of the 39th international symposium on symbolic and algebraic computation*, pages 296–303, 2014. 8, 16

[LKDW18] Felix Leditzky, Eneet Kaur, Nilanjana Datta, and Mark M. Wilde. Approaches for approximate additivity of the holevo information of quantum channels. *Phys. Rev. A*, 97:012332, Jan 2018. URL: https://link.aps.org/doi/10.1103/PhysRevA.97.012332, doi:10.1103/PhysRevA.97.012332. 12, 77

[LM07] Michael Lacey and William McClain. On an argument of Shkredov on two-dimensional corners. *Online Journal of Analytic Combinatorics*, 2007. URL: https://hosted.math.rochester.edu/ojac/vol2/Lacey_McClain_2007.pdf, arXiv:math/0510491. 9, 25, 31, 32

[LM11] Jean-Bernard Lasserre and Anjos F. Miguel. *Handbook on Semidefinite, Conic and Polynomial Optimization*. International Series in Operations Research & Management Science. Springer, November 2011. URL: https://hal.archives-ouvertes.fr/hal-00697757. 77, 78, 85

[Lov79] L. Lovasz. On the Shannon capacity of a graph. *IEEE Transactions on Information Theory*, 25(1):1–7, 1979. doi:10.1109/TIT.1979.1055985. 18, 27

[Lov19] Shachar Lovett. The analytic rank of tensors and its applications. *Discrete Anal.*, 2019. arXiv:1806.09179. 27, 29, 52

[LPS17] Bart Litjens, Sven Polak, and Alexander Schrijver. Semidefinite bounds for nonbinary codes based on quadruples. *Designs, Codes and Cryptography*, 84(1):87–100, 2017. 81, 86, 87

[LPS18] Nati Linial, Toni Pitassi, and Adi Shraibman. On the communication complexity of high-dimensional permutations. *In 10th Innovations in Theoretical Computer Science Conference (ITCS 2019)*, 124, page 54:1–54:20, 2018. arXiv:1706.02207, doi:10.4230/LIPIcs.ITCS.2019.54. 9, 10, 25, 26, 28, 32, 38, 44, 45

[LS21] Nati Linial and Adi Shraibman. Larger corner-free sets from better NOF Exactly-n protocols, 2021. arXiv:2102.00421. 25, 28, 32

[LT10] J. M. Landsberg and Zach Teitler. On the ranks and border ranks of symmetric tensors. *Found. Comput. Math.*, 10(3):339–366, 2010. doi:10.1007/s10208-009-9055-3. 64

[Mat15] Keiji Matsumoto. A new quantum version of f-divergence. In *Nagoya Winter Workshop: Reality and Measurement in Algebraic Quantum Theory*, pages 229–273. Springer, 2015. 76

[MDS+13] Martin Müller-Lennert, Frédéric Dupuis, Oleg Szehr, Serge Fehr, and Marco Tomamichel. On quantum Rényi entropies: A new generalization and some properties. *Journal of Mathematical Physics*, 54(12):122203–122203, December 2013. arXiv:1306.3142, doi:10.1063/1.4838856. 13, 74, 76

[Mes95] Roy Meshulam. On subsets of finite abelian groups with no 3-term arithmetic progressions. *Journal of Combinatorial Theory, Series A*, 71(1):168 – 172, 1995. doi:10.1016/0097-3165(95)90024-1. 33

[Nas20] Eric Naslund. The partition rank of a tensor and k-right corners in $\mathbb{F}_q^n$. *Journal of Combinatorial Theory, Series A*, 174:105190, 2020. doi:10.1016/j.jcta.2019.105190. 29

[NC11] Michael A. Nielsen and Isaac L. Chuang. *Quantum computation and quantum information*. Cambridge University Press, Cambridge, 2011. 71

[NM84] Linda Ness and David Mumford. A stratification of the null cone via the moment map. *American Journal of Mathematics*, 106(6):1281–1329, 1984. URL: http://www.jstor.org/stable/2374395. 69, 104

[NN94] Yurii Nesterov and Arkadii Nemirovskii. *Interior-point polynomial algorithms in convex programming*. SIAM, 1994. 22

[NRS06] Brendan Nagle, Vojtěch Rödl, and Mathias Schacht. The counting lemma for regular k-uniform hypergraphs. *Random Structures & Algorithms*, 28(2):113–179, 2006. 29

[NS17] Eric Naslund and Will Sawin. Upper bounds for sunflower-free sets. *Forum of Mathematics, Sigma*, 5, 2017. doi:10.1017/fms.2017.12. 29

[OPW97] Masanori Ohya, Dénes Petz, and Noboru Watanabe. On capacities of quantum channels. *Probability and Mathematical Statistics*, 17:179–196, 1997. 73, 94

[Pel18] Sarah Peluse. Three-term polynomial progressions in subsets of finite fields. *Israel Journal of Mathematics*, 228(1):379–405, 2018. 28, 30

[Pel19] Sarah Peluse. On the polynomial szemerédi theorem in finite fields. *Duke Mathematical Journal*, 168(5):749–774, 2019. 28, 30

[Pel22] Sarah Peluse. Bounds for subsets of $\mathbb{F}_p^n \times \mathbb{F}_p^n$ without l-shaped configurations, 2022. URL: https://www.ias.edu/math/events/computer-sciencediscrete-mathematics-seminar-ii-495. 34

[Pol19] Sven Carel Polak. *New methods in coding theory : Error-correcting codes and the Shannon capacity*. PhD thesis, University of Amsterdam, 2019. 80, 91

[PR98] Dénes Petz and Mary B. Ruskai. Contraction of generalized relative entropy under stochastic mappings on matrices. *Infinite Dimensional Analysis, Quantum Probability and Related Topics*, 01:83–89, 1998. 76

[Pre15] Sean Prendiville. Matrix progressions in multidimensional sets of integers. *Mathematika*, 61(1):14–48, 2015. 34

[Pro07] Claudio Procesi. Lie groups. an approach through invariants and representations. *Bull. Amer. Math. Soc*, 2007. 80

[PS72] George Pólya and Gabor Szegö. *Problems and Theorems in Analysis: Series, integral calculus, theory of functions*. Springer, 1972. 106

[PS19] Sven C Polak and Alexander Schrijver. New lower bound on the shannon capacity of c7 from circular graphs. *Information Processing Letters*, 143:37–40, 2019. 18

[PV10] Yury Polyanskiy and Sergio Verdú. Arimoto channel coding converse and Rényi divergence. *48th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pages 1327–1333, September 2010. 76

[Rai99] Eric M Rains. Bound on distillable entanglement. *Physical Review A*, 60(1):179, 1999. 100

[Rai01] Eric M Rains. A semidefinite program for distillable entanglement. *IEEE Transactions on Information Theory*, 47(7):2921–2933, 2001. 100

[Sac05] Massimiliano F Sacchi. Optimal discrimination of quantum operations. *Physical Review A*, 71(6):062340, 2005. 12

[Sag13] Guillaume Sagnol. On the semidefinite representation of real functions applied to symmetric matrices. *Linear Algebra and its Applications*, 439(10):2829 – 2843, 2013. URL: http://www.sciencedirect.com/science/article/pii/S002437951300520X, doi:https://doi.org/10.1016/j.laa.2013.08.021. 77

[Ser77] Jean-Pierre Serre. *Linear Representations of Finite Groups*. Cambridge University Press, 1977. 20

[Sha56] Claude E. Shannon. The zero error capacity of a noisy channel. *IRE Trans. Inf. Theory*, 2(3):8–19, 1956. doi:10.1109/TIT.1956.1056798. 8, 18

[Shi18] Yaroslav Shitov. A counterexample to comon's conjecture. *SIAM J. Appl. Algebra Geom.*, 2(3):428–443, 2018. doi:10.1137/17M1131970. 53, 54, 58, 62

[Shi22] Yaroslav Shitov. The ubrank of a complex symmetric tensor can exceed its symmetric subrank. *Personal communication*, 2022. 53, 58

[Shk06a] I. D. Shkredov. On a generalization of Szemerédi's theorem. *Proc. London Math. Soc.*, 93(3):723–760, 2006. doi:10.1017/S0024611506015991. 9, 25, 31

[Shk06b] I. D. Shkredov. On a problem of Gowers. *Izvestiya: Mathematics*, 70(2):385–425, 2006. doi:10.1070/im2006v070n02abeh002316. 9, 25, 31

[Shk15] Ilya D Shkredov. On some problems of euclidean ramsey theory. *Analysis Mathematica*, 41(4):299–310, 2015. 34

[Shr18] Adi Shraibman. A note on multiparty communication complexity and the Hales–Jewett theorem. *Information Processing Letters*, 139:44–48, 2018. doi:10.1016/j.ipl.2018.07.002. 25, 28, 32

[SS21] Ilya Shkredov and Jozsef Solymosi. Tilted corners in integer grids. *arXiv preprint arXiv:2103.01645*, 2021. 34

[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 1986)*, pages 49–54, 1986. doi:10.1109/SFCS.1986.52. 11, 18, 53, 65, 67

[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. 11, 15, 16, 18, 35, 52

[Str88] Volker Strassen. The asymptotic spectrum of tensors. *J. Reine Angew. Math*, 384:102–152, 1988. doi:10.1515/crll.1988.384.102. 8, 11, 18, 19, 53, 65, 66, 67

[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. 11, 18, 27, 28, 46, 47, 53, 65, 67

[Str05] Volker Strassen. Komplexität und Geometrie bilinearer Abbildungen. *Jahresber. Deutsch. Math.-Verein.*, 107(1):3–31, 2005. 67, 70, 105

[SW97] Benjamin Schumacher and Michael D. Westmoreland. Sending classical information via noisy quantum channels. *Phys. Rev. A*, 56:131–138, Jul 1997. URL: https://link.aps.org/doi/10.1103/PhysRevA.56.131, doi:10.1103/PhysRevA.56.131. 93

[SW13] Naresh Sharma and Naqueeb Ahmad Warsi. Fundamental bound on the reliability of quantum information transmission. *Phys. Rev. Lett.*, 110:080501, Feb 2013. URL: https://link.aps.org/doi/10.1103/PhysRevLett.110.080501, doi:10.1103/PhysRevLett.110.080501. 76

[Sze75] Endre Szemerédi. On sets of integers containing no k elements in arithmetic progression. *Acta Arith*, 27(299-345):5, 1975. 29

[Tao16] Terence Tao. A symmetric formulation of the Croot-Lev-Pach-Ellenberg-Gijswijt capset bound. *Tao's blog post*, 2016. URL: https://terrytao.wordpress.com. 16, 29

[TGW14] Masahiro Takeoka, Saikat Guha, and Mark M Wilde. The squashed entanglement of a quantum channel. *IEEE Transactions on Information Theory*, 60(8):4987–4998, 2014. 100

[Tob91] Verena Tobler. *Spezialisierung und Degeneration von Tensoren*. PhD thesis, Universität Konstanz, 1991. http://nbn-resolving.de/urn:nbn:de:bsz:352-opus-20324. 53, 65, 67

[Tod01] Michael J Todd. Semidefinite optimization. *Acta Numerica*, 10:515–560, 2001. 22

[Tom15] Marco Tomamichel. *Quantum information processing with finite resources: mathematical foundations*, volume 5. Springer, 2015. https://arxiv.org/pdf/1504.00233.pdf. 76

[TS16] Terence Tao and Will Sawin. Notes on the "slice rank" of tensors. *Tao's blog post*, 2016. URL: https://terrytao.wordpress.com/2016/08/24/notes-on-the-slice-rank-of-tensors/. 16, 27, 28, 52

[TZ15] Terence Tao and Tamar Ziegler. A multi-dimensional szemerédi theorem for the primes via a correspondence principle. *Israel Journal of mathematics*, 207(1):203–228, 2015. 30

[Vio19] Emanuele Viola. Guest column: Non-abelian combinatorics and communication complexity. *SIGACT News*, 50(3):52–74, 2019. doi:10.1145/3364626.3364637. 28

[Vra21] Péter Vrana. A generalization of strassen's theorem on preordered semirings. *Order*, pages 1–20, 2021. 18

[Wal14] Michael Walter. *Multipartite Quantum States and their Marginals*. PhD thesis, Zurich, ETH, 2014. arXiv:1410.6820, doi:10.3929/ethz-a-010250985. 70, 104

[WBHK20] Mark M Wilde, Mario Berta, Christoph Hirche, and Eneet Kaur. Amortized channel divergence for asymptotic quantum channel discrimination. *Letters in Mathematical Physics*, 110(8):2277–2336, 2020. 12, 73

[WFD17] Xin Wang, Kun Fang, and Runyao Duan. Semidefinite programming converse bounds for quantum communication. *IEEE Transactions on Information Theory*, 65(4):2583–2592, 2017. doi:10.1109/TIT.2018.2874031. 99

[WFT19] Xin Wang, Kun Fang, and Marco Tomamichel. On converse bounds for classical communication over quantum channels. *IEEE Transactions on Information Theory*, 65(7):4609–4619, 2019. doi:10.1109/TIT.2019.2898656. 75, 94, 95

[Wil13] Mark M. Wilde. *Quantum Information Theory*. Cambridge University Press, 2013. doi:10.1017/CBO9781139525343. 73

[WW19] Xin Wang and Mark M Wilde. Resource theory of asymmetric distinguishability for quantum channels. *Physical Review Research*, 1(3):033169, 2019. 12, 73

[WWY14] Mark Wilde, Andreas Winter, and Dong Yang. Strong converse for the classical capacity of entanglement-breaking and hadamard channels via a sandwiched rényi relative entropy. *Communications in Mathematical Physics*, 331, October 2014. doi:10.1007/s00220-014-2122-x. 13, 74, 76

[WZ21] Avi Wigderson and Jeroen Zuiddam. Asymptotic spectra: Theory, applications and extensions, 2021. URL: https://staff.fnwi.uva.nl/j.zuiddam/papers/convexity.pdf. 18, 53

[Yao79] Andrew Chi-Chih Yao. Some complexity questions related to distributive computing. In *Proceedings of the Eleventh Annual ACM Symposium on Theory of Computing*, STOC '79, pages 209–213. ACM, 1979. doi:10.1145/800135.804414. 25, 32

[Zha19] Yufei Zhao. Graph theory and additive combinatorics. Lecture Notes, 2019. URL: https://yufeizhao.com/gtac/. 34

[Zui18] Jeroen Zuiddam. *Algebraic complexity, asymptotic spectra and entanglement polytopes*. PhD thesis, University of Amsterdam, 2018. 18, 19, 52, 53, 65, 66, 70, 105

[Zui19] Jeroen Zuiddam. The asymptotic spectrum of graphs and the Shannon capacity. *Combinatorica*, 39:1173–1184, 2019. URL: https://doi.org/10.1007/s00493-019-3992-5. 18, 20