Quasirandomness in quantum information theory

Farrokh Labib

Quasirandomness in quantum information theory

ILLC Dissertation Series DS-2022-03

INSTITUTE FOR LOGIC, LANGUAGE AND COMPUTATION

For further information about ILLC-publications, please contact

Institute for Logic, Language and Computation
Universiteit van Amsterdam
Science Park 107
1098 XG Amsterdam
phone: +31-20-525 6051
e-mail: illc@uva.nl
homepage: http://www.illc.uva.nl/

The work in this dissertation was supported by the Gravitation-grant NETWORKS-024.002.003 from the Dutch Research Council (NWO).

Cover design by Farrokh Labib.
Printed and bound by Ipskamp Printing.

Quasirandomness in quantum information theory

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. dr. ir. K.I.J. Maex
ten overstaan van een door het College voor Promoties ingestelde commissie,
in het openbaar te verdedigen in de Agnietenkapel
op woensdag 26 januari 2022, te 10.00 uur

door

Farrokh Sieyar Labib

geboren te Kabul

Promotiecommisie

Promotor:	prof. dr. H.M. Buhrman	Universiteit van Amsterdam
Co-promotor:	dr. J. Briët	Centrum Wiskunde & Informatica
Overige leden:	prof. dr. R.M. de Wolf	Universiteit van Amsterdam
	prof. dr. C.J.M. Schoutens	Universiteit van Amsterdam
	prof. dr. D.C. Gijswijt	TU Delft
	dr. M. Ozols	Universiteit van Amsterdam
	dr. M. Walter	Universiteit van Amsterdam

Faculteit der Natuurwetenschappen, Wiskunde en Informatica

This dissertation is based on the following articles. For all these articles, the co-autorship is shared equally.

1) T. Bannink, J. Briët, H. Buhrman, F. Labib, and T. Lee. Bounding quantum-classical separations for classes of nonlocal games. In *36th International Symposium on Theoretical Aspects of Computer Science (STACS 2019)*, 126, pages 12:1–12:11. March 2019. doi: 10.4230/LIPIcs.STACS.2019.12

2) T. Bannink, J. Briët, F. Labib, and H. Maassen. Quasirandom quantum channels. *Quantum*, 4:298, 2020. doi: 10.22331/q-2020-07-16-298. Best paper award TQC 2020.

3) F. Labib. Stabilizer rank and higher-order Fourier analysis. *arXiv preprint arXiv:2107.10551*, 2021

4) J. Briët and F. Labib. High-entropy dual functions over finite fields and locally decodable codes. *Forum of Mathematics, Sigma*, 9:e19, 2021. doi: 10.1017/fms.2021.1

The author has additionally co-authored in the following paper which is not included in this dissertation.

5) T. Giurgica-Tiron, I. Kerenidis, F. Labib, A. Prakash, and W. Zeng. Low depth algorithms for quantum amplitude estimation. *arXiv:2012.03348*, 2020

Contents

Acknowledgments

First of all, I want to thank my promotor Harry Buhrman and my supervisor Jop Briët for having me as their PhD student. Thank you Jop for all the amazing research discussions we have had over the years, introducing me to beautiful mathematics and teaching me how to become a researcher of my own. I am also grateful for the detailed reading of previous versions of this dissertation.

Next I want to thank all my coauthors for all the successful collaborations over the past four years: Tom Bannink, Jop Briët, Harry Buhrman, Chris Cade, Tudor Giurgica-Tiron, Iordanis Kerenidis, Troy Lee, Hans Maassen, Ido Niesen, Anupam Prakash, and William Zeng.

I also want to thank Dion Gijswijt, Maris Ozols, Kareljan Schoutens, Michael Walter, and Ronald de Wolf for agreeing to be part of my PhD committee and taking the time to read my thesis. Special thanks to Ronald de Wolf for the very detailed comments.

I want to thank Rajiv Krishnakumar and William Zeng for having me in their research group at Goldman Sachs during the summer of 2020.

Next I want to thank all the amazing colleagues I have had in the past four years at CWI: Jeroen Zuiddam, Alvaro Piedrafita, Arjan Cornelissen, Subhasree Patro and Harold Nieuwboer for sharing the "peanut butter room" with me, Tom Bannink for creating the foosball system and playing lots of chess with me, Chris Cade and Marten Folkertsma for the amazing bouldering sessions, and Srinivasan Arunachalam, Jan Czajkowski, Koen Groenland, Freek Witteveen, Joris Kattemölle, Yfke Dulek, András Gilyén, Yinan Li, Isabella Pozzi, Mathé Zeegers, Sander Gribling, Ruben Brokkelkamp, Joran van Apeldoorn, Florian Speelman, Sebastian de Bone, Ido Niesen, Chris Majenz, Jonas Helsen, Peter van der Gulik, Jana Sotakova, Philip Verduyn Lunel, Rene Allerstorfer, Jordi Weggemans, Sebastian Zur for being such great colleagues.

Last but not least, I want to thank my family and friends. But most importantly my parents for fleeing Afghanistan to ensure a better future for their children.

Chapter 1

Introduction

In this dissertation, we study *quasirandomness* in several contexts, mostly in quantum information theory. An object is quasirandom if it shares properties with a *random* object. What these properties are, depends on the context. Consider, for example, uniformly random 3-regular graphs. They have the property that they are likely highly connected, while at the same time the number of edges is quite small (the graph is "sparse") [AS04a]. Highly connected refers to the property that, for example, a random walk on the graph mixes very rapidly: after a small number of steps, the position of the walker is close to uniformly random. So when an explicit 3-regular graph has this property as well, we say that it is quasirandom. Such graphs are also called *expanders* and they have several other nice properties and have found many applications in mathematics and computer science [HLW06].

Other interesting objects are linear maps from matrices to matrices, complex-valued functions on a finite abelian group, etc. These objects appear in quantum information theory in the form of quantum channels or the amplitudes of quantum states for example. Analogous to the graph case, random quantum channels have the property that they quickly "mix" any quantum state by applying it a small number of times. That is, the resulting quantum state is close to the "maximally mixed state" [HL09]. So explicit quantum channels with this property may be called quasirandom. For quantum states we consider the notion of rank, where you want to express a quantum state in terms of a minimal number of simpler states called *stabilizer states*, this is the *stabilizer rank*. In this case, random quantum states have a "high" stabilizer rank.

In most cases, it is then a challenge to show that an *explicit* object that we believe is quasirandom, shares a typical property with a truly random object. Taking again the example of expanders, while it is not hard to show that they exist (random 3-regular graphs are likely to be expanders through the probabilistic method), it is considerably more difficult to construct them explicitly. Similarly, we know that random quantum states have a high stabilizer rank, but it is still

unknown whether the *magic state*, an explicit quantum state crucial for universal quantum computing, has a high stabilizer rank [BBC+19].

The results from this thesis rely heavily on tools from *higher-order Fourier analysis*. Higher-order Fourier analysis is a still nascent area of mathematics that grew out of Gowers's Fourier-analytic proof of Szemerédi's theorem on *arithmetic progressions* [Gow01]. An arithmetic progression of length k is a sequence of integers $x, x+d, x+2d, \ldots, x+(k-1)d$ for some integers x, d. Here x is called the starting point and d the common difference of the progression. Szemeredi's theorem states that any "dense" subset in the interval $\{1, 2, \ldots, N\}$ contains arbitrarily long arithmetic progressions, provided that N is large enough. This is a fundamental result in additive combinatorics that has a number of different proofs, one of which is Fourier-analytic and which led to the development of higher-order Fourier analysis. The applications found in this thesis motivate the further study of tools in higher-order Fourier analysis in the context of quantum information theory, where they have thus far not been used much. Higher-order Fourier analysis has already found many applications in classical theoretical computer science, such as in property testing, coding theory, and complexity theory [HHL19]. The results in this thesis contribute to the development of higher-order Fourier analysis in quantum information-theoretic settings.

We now give a high-level overview of each topic that we study in this dissertation.

Nonlocal games. The framework of *nonlocal games* studies a fundamental property of nature, that of *nonlocality* in quantum mechanics. Entanglement between spatially separated quantum systems allows for correlations that would not be possible "classically", that is, in a *hidden variable theory*. This statement was made precise by John Bell in [Bel64] using so-called "Bell tests". Such a test, which is a setup of physical systems, can demonstrate the violation of "Bell inequalities". In the framework of nonlocal games, the CHSH game [CHSH69] provides us with a Bell inequality as follows.

The game consists of two players, usually named Alice and Bob, and a referee. The referee samples two uniformly random bits $x, y \in \{0, 1\}$ and sends x to Alice and y to Bob. Without communicating, Alice and Bob send answers $a, b \in \{0, 1\}$ to the referee respectively. They "win" if $a + b = xy \mod 2$. Alice and Bob are allowed to devise a strategy together before the game starts. Such a strategy tells the players what bit to answer given the question. Once the referee sends the questions, they are not allowed to communicate. They are also allowed to use shared randomness (hidden variables), i.e. coin tosses that they both can see, and decide what to answer depending on the outcome of the toss and their input. This will however not help them gain an advantage over *deterministic strategies*, i.e. strategies without randomness. Such strategies are also called *classical strategies*. It is not hard to show that the best Alice and Bob can do using classical strategies

is answer correctly 3 out of 4 times, see Section 1.2. In a hidden variable model of nature, classical strategies are the only possible strategies that can be used to play the game. Hence, this game provides us with a Bell inequality, meaning that the probability of winning the CHSH game is at most 3/4.

It is however possible to play better (win with probability higher than 3/4) when the players are allowed to share a quantum state. In this case, the players answer according to a *measurement* performed on their part of the shared entangled state. Such strategies are called quantum strategies. Using a quantum strategy, it can be shown that the probability that Alice and Bob win, is approximately 0.85, see Section 1.2. In particular, "the Bell inequality is violated" which is not possible in a classical world. Experimentally showing this implies that nature can not be described by a hidden variable theory. This was first done in [AGR82] and more recently [HBD$^+$15] (in a "loophole-free" way).

It appears to be the case that shared entanglement can provide useful correlations for the players in playing nonlocal games. The following question then arises naturally: how much advantage can there be in using quantum strategies in nonlocal games over classical strategies? In this dissertation, we will look at multiplayer *XOR* games, a subclass of nonlocal games. These games have the property that the answers of the players are bits and that the winning condition of the game only depends on the XOR of the answers. The CHSH game just discussed is an example. We will see in Section 1.2 that such games are given by a *game tensor* which is a map from the questions to $\{\pm 1\}$. XOR games have the property that the winning probability is always at least 1/2, because the players can always output a random bit. Therefore, we look at the *bias* of the game, which measures how much better than the random strategy the game can be played. For two-player XOR games, it is known that the quantum bias is at most a constant times the classical bias [Tsi87], which implies that quantum strategies do not give an arbitrary big advantage over classical. This follows from a deep theorem in Banach space theory, Grothendieck's inequality [Gro53], together with Tsirelson's Theorem [Tsi87].

In the case of three or more players, less is known about the ratio of quantum and classical bias. There is a sequence of three-player XOR games for which this ratio goes to infinity [PGWP$^+$08]. The games for which this happens were not made explicit, merely its existence was shown. Later in [BV13] this was quantitatively improved, but the games were still not explicit, as the techniques used are based on the probabilistic method. These games also have the additional property that the quantum bias goes to zero. It is an open problem to find an explicit family of three-player XOR games for which the ratio of quantum and classical bias diverges. This is another example of the phenomenon that through the probabilistic method existence of objects (XOR games in this case) can be shown with certain properties, but explicitly constructing them turns out to be hard. In Chapter 2 we study the slightly refined problem where we require that the quantum bias does not go to zero. We rule out certain natural subclasses of

XOR games for which this might, a priori, be possible. Roughly speaking, we show that the quantum bias of such games is bounded from above by certain tensor norms that arise in the combinatorial and Fourier-analytic proofs of Szemerédi's theorem, of the game tensor. This implies that if the quantum bias is large, the norm of the game tensor norm is also large. This in turn implies the existence of "structure", the opposite of what we would expect from a random tensor. It turns out that this structure can be turned into a classical strategy for which the classical bias is large as well, meaning that the ratio of quantum and classical bias is bounded.

Quasirandom graphs. In seminal work [CGW89], Chung, Graham, and Wilson introduced the notion of quasirandom graphs. They showed that a number of properties (seven to be precise) that are typical for random graphs are equivalent for dense graphs. In particular, if an explicit family of dense graphs satisfies one of the properties, it must simultaneously satisfy all the other properties. This means that this family of graphs behaves as though it is random, in a number of different ways, which is why we call it quasirandom. Two of these properties that we will focus on are (spectral) *expansion* and *uniformity*. A family of d-regular graphs is an expander if the second to largest eigenvalue in absolute value of the adjacency matrix is bounded away from d. One of the (many) properties of such graphs is that random walks on these graphs converge rapidly to the uniform (limit) distribution. This can for example be used to reduce the amount of randomness that a probabilistic algorithm uses, as randomness can be quite an expensive resource [HLW06]. Uniformity on the other hand is a combinatorial property of the graph. Roughly speaking, a graph is uniform if, for arbitrary pairs of subsets of the vertex set, the edge density between these two subsets is approximately the same as the overall edge density of the graph. It is straightforward to show that expander graphs are also uniform, showing that expansion is a stronger property of a graph. Chung, Graham, and Wilson showed that the converse also holds for dense graphs. That is, if a dense graph is uniform, it also has to be an expander. In [KS06] it was shown that this is not true in general, they found counterexamples to this converse in sparse graphs. However, Conlon and Zhao [CZ17] showed that such a converse holds for *vertex-transitive* graphs, surprisingly using Grothendieck's inequality.

In Chapter 3 we generalize the theory of quasirandom graphs to quantum information theory. In quantum information theory, the notion of a *quantum channel* is central. It is the most general transformation physically realizable that a quantum state can undergo. In classical information theory, on the other hand, transition matrices, or Markov chains, are the most general transformations on probability distributions, i.e. classical information. Classical channels can therefore be identified with a weighted graph and as such, a natural generalization of graphs are quantum channels, or *superoperators* in general.

We will study the relationship between expansion for quantum channels, a notion that has already been extensively studied [Has07, BST10], and a natural generalization of uniformity for quantum channels. We are able to prove quantum analogues of the results of Chung, Graham, and Wilson on quasirandomness of graphs. We show that expansion and uniformity for "randomizing" quantum channels [Aub09] are equivalent. We also show that Conlon and Zhao's result can be generalized to say that "irreducibly covariant" quantum channels admit a converse (from uniformity to expansion). Irreducible covariance is an important notion for quantum channels relevant to "additivity conjectures" in quantum information theory [Hol05]. Here we use the noncommutative Grothendieck inequality [Haa85] for proving this converse.

Stabilizer rank. According to the Gottesman-Knill Theorem [Got98, NC02] we can efficiently simulate any quantum circuit consisting of *stabilizer operations*, which are Clifford gates and Pauli measurements, on a classical computer. Such circuits can be promoted to universal quantum computation by adding a non-Clifford gate to our gate set or if we have access to a "magic state". It is widely believed that universal quantum computers cannot be efficiently simulated on classical computers: state-of-the-art simulators using modern-day supercomputers are only able to simulate a few dozen qubits [CZH$^+$18, HS17, PGN$^+$17, SSAG16]. The quantum states that we obtain from the canonical all-zero state after applying Clifford gates are the *stabilizer states* [Got97]. Then, lower bounds on the minimal number of terms needed to express the n-qubit magic state as a superposition of stabilizer states gives a lower bound on the cost of certain simulation algorithms of stabilizer operations applied to the n-qubit magic state. This minimal number of terms is also called the *stabilizer rank* of, in this case, the magic state.

Pick a random n-qubit state (from the Haar measure) and with a very high probability its stabilizer rank is exponential in n, this follows from a dimension counting argument. But showing that an explicit n-qubit state has an exponentially large stabilizer rank turns out to be hard. Here we have a candidate explicit n-qubit state, the n-qubit magic state, for which it is expected that its stabilizer rank is exponential in n [BBC$^+$19]. In [BSS16] it was shown that the stabilizer rank of the n-qubit magic state has to be at least $\Omega(\sqrt{n})$ and recently this was improved to $\Omega(n)$ [PSV21]. A huge gap persists between the hoped-for lower bound and the best known lower bound, and the techniques used in these results will not be able to close this gap. In Chapter 4 we take an approach that is different from the previous two results.

It turns out that the amplitudes of stabilizer states and the n-qubit magic states are *polynomial phase functions* (to be precise *nonclassical* polynomial phase functions defined on affine subspaces) [DDM03, HDDM05, HV12]. Such objects are of central importance in higher-order Fourier analysis, a generalization of

Fourier analysis where instead of studying the correlation of functions with characters, the correlation with polynomial phase functions are studied. In Chapter 4 we use this observation together with tools from higher-order Fourier analysis [HHL19] to show that the stabilizer rank of the n-qubit magic state is $\Omega(n)$, giving an alternative Fourier-analytic proof of the result in [PSV21].

Decomposition theorems. The final topic in this dissertation concerns a result in higher-order Fourier analysis, instead of an application of it, relevant to a certain refinement of Szemerédi's theorem referred to as *Szemerédi's theorem with random differences*. Before explaining this, recall that Szemerédi's theorem states that any "dense" subset of $\{1, 2, \dots, N\}$, with N large enough, contains arbitrarily long arithmetic progressions. In this theorem, there is no restriction as to what the common difference of the progression can be and one would like to know if the theorem still holds when the common difference is from certain special sets. There are some known results in this direction, for example for $k \geq 2$ it is known that any dense subset in $\{1, 2, \dots, N\}$, with N large enough, contains a k-term arithmetic progression $x, x+d, \dots, x+(k-1)d$ with the common difference d in the squares $\{1^2, 2^2, 3^2, \dots\}$ [BL96]. Other sets of common differences known with this property are the shifted primes $\{p-1 \colon p \text{ prime}\}$ and $\{p+1 \colon p \text{ prime}\}$ [WZ12]. Frantzikinakis et al. [FLW16] asked if these sets are really special or that a random subset $D \subset \{1, 2, \dots, N\}$ would satisfy this property as well. Here, the random subset D is defined by adding each $d \in \{1, 2, \dots, N\}$ to D with probability $\rho > 0$. Roughly speaking, we want to know what the smallest probability ρ is such that with the random subset D, Szemerédi's theorem with random differences holds. An avenue for solving this problem is through *decomposition theorems* for certain functions called *dual functions*.

In higher-order Fourier analysis, decomposition theorems play an important role in understanding counts of linear configurations in subsets of the interval $\{1, 2, \dots, N\}$ where N is a large integer [Gow10]. For example, the linear configuration might be arithmetic progressions. One usually decomposes the indicator function of the subset in terms of polynomial phase functions plus an error term and a quasirandom component that is small in the "Gowers uniformity norm". The polynomial part is then easier to analyze, while the error term and quasirandom part do not affect the count of the linear configuration that much. In this way, one obtains a Fourier-analytic proof of Szemerédi's theorem [Gow01, Tao05].

The types of decomposition theorems we will look at in Chapter 5 are for more structured functions known as dual functions. Instead of decomposing an arbitrary function on a finite abelian group, we wish to decompose dual functions in terms of polynomial phase functions plus an error term. Dual functions can, for example, count progressions of length three in subsets of $\{1, 2, \dots, N\}$ with a given common difference. Let $A \subset \{1, 2, \dots, N\}$ and $\mathbb{1}_A$ the indicator function of A, that is, $\mathbb{1}_A(x) = 1$ if $x \in A$ and 0 otherwise. The dual function, in this case,

is given by

$$\phi_A \colon d \mapsto \frac{1}{N} \sum_{x=1}^{N} \mathbb{1}_A(x)\mathbb{1}_A(x+d)\mathbb{1}_A(x+2d).$$

In other words, $\phi_A(d)$ counts the number of progressions of length three in A with common difference d. The powerful Gowers inverse theorem together with the Hahn-Banach theorem allows us to decompose these functions such that the error term is small on "average" [Gow10]. However, if we can decompose dual functions in such a way that the error term is small *everywhere*, then this would imply certain optimality results on Szemerédi's theorem with random differences, more on this implication in Section 1.5. However, we show in Chapter 5 that such a decomposition theorem is not true in general in vectors spaces over finite fields. Our proof uses breakthrough constructions of error-correcting codes due to Yekhanin [Yek08].

1.1 Preliminaries

First we will list the basic notation and terminology. Then we will go through some background information for the upcoming chapters.

Notation. For a finite set S, we write $\mathbb{E}_{x\in S}$ for $\frac{1}{|S|}\sum_{x\in S}$. For integer $n \in \mathbb{N}$, define $[n] := \{1, 2, \dots, n\}$. Let S be a set and $T \subset S$ a subset. We write $\overline{T}$ for the complement of T. We define $\mathbb{1}_T \colon S \to \{0, 1\}$ to be the indicator function of T, that is $\mathbb{1}_T(s) = 1$ if and only if $s \in T$. For a complex number $z \in \mathbb{C}$ we write $\Re(z)$ for its real part and $\Im(z)$ for its imaginary part. Let $\mathcal{H}$ be a finite-dimensional Hilbert space. We denote by $\mathcal{L}(\mathcal{H})$ the set of linear maps $A \colon \mathcal{H} \to \mathcal{H}$. If $\mathcal{K}$ is another finite-dimensional Hilbert space, their tensor product, denoted by $\mathcal{H} \otimes \mathcal{K}$, is again a Hilbert space and is the set of all $v \otimes w$ for $(v, w) \in \mathcal{H} \times \mathcal{K}$ such that for $\alpha \in \mathbb{C}$

- $(\alpha v) \otimes w = \alpha v \otimes w$,
- $v \otimes (\alpha w) = \alpha v \otimes w$,
- $(v + v') \otimes w = v \otimes w + v' \otimes w$,
- $v \otimes (w + w') = v \otimes w + v \otimes w'$.

Let p be a prime. We denote by $\mathbb{F}_p$ the field of p elements and $\mathbb{T} = \mathbb{R}/\mathbb{Z}$. Define $\iota \colon \mathbb{F}_p \to \mathbb{T}$ to be the map given by

$$\iota \colon x \mapsto |x|/p \mod 1,$$

where $|\cdot|\colon \mathbb{F}_p \to \{0, 1, \ldots, p-1\}$ is the natural map. We also define this map on $\mathbb{F}_p^n$ by $|\cdot|\colon \mathbb{F}_p^n \to \mathbb{Z}_{\geq 0} : x \mapsto |x_1| + \cdots + |x_n|$, this is an abuse of notation, but it should be clear from context what the domain of the map $|\cdot|$ is. The exponential map $e\colon \mathbb{T} \to \mathbb{C}$ is defined to be $e(t) := e^{2\pi i t}$.

Fourier analysis. We recall now the basics of Fourier analysis over prime finite fields. In Fourier analysis, the protagonists are the *characters* of the group we are looking at. These characters have many nice properties, of which the orthogonality might be the most important one.

1.1.1. Definition (Characters). Let G be a finite abelian group. Homomorphisms $\chi\colon G \to \mathbb{C}^*$ are called characters. The set of characters is denoted by $\widehat{G}$.

The set of characters $\widehat{G}$ is a finite abelian group with pointwise multiplication as group operation.

1.1.2. Example. Let $G = \mathbb{Z}_N$ be the cyclic group of order N. One can show that the map $\chi_a\colon \mathbb{Z}_N \to \mathbb{C}\colon x \mapsto e(ax/N)$ for any $a \in \mathbb{Z}_N$ is a character. The map $a \mapsto \chi_a$ gives an isomorphism $G \to \widehat{G}$. Similarly, for prime finite fields they take the following explicit form. For $z \in \mathbb{F}_p^n$, the map $\chi_z(x) := e(\langle z, x\rangle/p)$ is a character. It is not hard to check that $z \mapsto \chi_z$ again gives an isomorphism between $\mathbb{F}_p^n$ and $\widehat{\mathbb{F}_p^n}$.

For two functions $f, g\colon \mathbb{F}_p^n \to \mathbb{C}$, their inner product is defined by

$$\langle f, g\rangle := \mathop{\mathbb{E}}_{x\in\mathbb{F}_p^n} f(x)\overline{g(x)} = p^{-n}\sum_{x\in\mathbb{F}_p^n} f(x)\overline{g(x)}.$$

Using this inner product, the characters form an orthonormal basis for the space of functions $\{f\colon \mathbb{F}_p^n \to \mathbb{C}\}$.

1.1.3. Proposition (Orthogonality of characters). *Let $z, z' \in \mathbb{F}_p^n$. Then*

$$\langle \chi_z, \chi_{z'}\rangle = \begin{cases} 1 & \text{if } z = z' \\ 0 & \text{otherwise.} \end{cases}$$

As the characters form an orthonormal basis for the space of functions, any function $f\colon \mathbb{F}_p^n \to \mathbb{C}$ can be written as a linear combination of characters. This decomposition is used so often that it has a name: the *Fourier transform*. Denote by $\widehat{f}$ the Fourier transform of f, i.e.

$$\widehat{f}(z) = \mathop{\mathbb{E}}_{x\in\mathbb{F}_p^n} f(x)\chi_z(-x), \quad z \in \mathbb{F}_p^n.$$

The inner product for the Fourier transforms is defined by

$$\langle \widehat{f}, \widehat{g} \rangle = \sum_{z \in \mathbb{F}_p^n} \widehat{f}(z)\overline{\widehat{g}(z)}.$$

Denote by $*$ the convolution operator, i.e.

$$f * g(x) = \mathop{\mathbb{E}}_{y+z=x} f(y)g(z).$$

Norms. Norms play an important role in this thesis. In a normed vector space, the norm gives a notion of how large an object in that space is and also defines a notion of distance, turning it naturally into a metric space. They can also define a notion of quasirandomness: an object is quasirandom if it is small in a certain norm. The spaces in which we consider norms are usually $\mathbb{C}^n$, the space of $n \times n$ matrices or spaces of functions on finite sets. Here we give the definitions of the norms that we use throughout. Let $M_n(\mathbb{C})$ be the set of $n \times n$ matrices with complex entries. The trace of $A \in M_n(\mathbb{C})$ is defined to be $\mathsf{Tr}(A) = \sum_{i \in [n]} A_{ii}$.

- For $p \in [1, \infty)$, $x \in \mathbb{C}^n$ the L_p-norm and ℓ_p-norm are defined as

$$\|x\|_{L_p} = \Big(\mathop{\mathbb{E}}_{i \in [n]} |x_i|^p \Big)^{1/p} \quad \text{and} \quad \|x\|_{\ell_p} = \Big(\sum_{i \in [n]} |x_i|^p \Big)^{1/p}$$

  and $\|x\|_{L_\infty} = \|x\|_{\ell_\infty} = \max_i |x_i|$.

- For $A \in M_n(\mathbb{C})$, its operator norm, denoted $\|\cdot\|_{\mathrm{op}}$, is defined by

$$\|A\|_{\mathrm{op}} = \max_{x \in \mathbb{C}^n \colon \|x\|_{\ell_2} = 1} \|Ax\|_{\ell_2},$$

  or equivalently, its largest singular value.

- The Schatten-p norms, denoted by $\|\cdot\|_{S_p}$, for $p \in (0, \infty)$ and $A \in M_n(\mathbb{C})$ is defined to be

$$\|A\|_{S_p} = \Big(\frac{1}{n} \mathsf{Tr}\left[(A^* A)^{p/2} \right] \Big)^{1/p},$$

  and for $p = \infty$ define $\|A\|_{S_\infty} = \|A\|_{\mathrm{op}}$.

- Let G be a finite abelian group and $f \colon G \to \mathbb{C}$ a complex-valued function on G. Define the multiplicative derivative Δ_h for any $h \in G$ for such functions as

$$\Delta_h f(x) = f(x+h)\overline{f(x)}.$$

We define for any $s \geq 1$ the Gowers norm $\|\cdot\|_{U^s(G)}$

$$\|f\|_{U^s(G)} = (\mathop{\mathbb{E}}_{h_1,\ldots,h_s,x\in G} \Delta_{h_1} \cdots \Delta_{h_s} f(x))^{1/2^s}. \tag{1.1}$$

For $s = 1$ we get the absolute value of the mean of the function

$$\|f\|_{U^1(G)} = (\mathop{\mathbb{E}}_{h,x\in G} \Delta_h f(x))^{1/2} = |\mathop{\mathbb{E}}_{x\in G} f(x)|,$$

so technically it is not a norm, but for $s > 1$ it is indeed a norm [TV06]. By the recursion

$$\|f\|^{2^{s+1}}_{U^{s+1}(G)} = \mathop{\mathbb{E}}_{h\in G} \|\Delta_h f\|^{2^s}_{U^s(G)}$$

one sees that the expectation in Equation 1.1 is a non-negative real.

Quantum mechanics. We refer to [NC02, dW19] for the basics in quantum computing and quantum information theory.

We use braket notation for quantum states. A pure quantum state $|\psi\rangle$ is a unit vector in a finite-dimensional complex Hilbert space $\mathcal{H}$, which is isomorphic to $\mathbb{C}^n$ for some $n \in \mathbb{N}$. We say that the dimension of $|\psi\rangle$ is n. Write $\langle\psi|$ for the dual of $|\psi\rangle$ and $\langle\psi|\phi\rangle$ for the inner product between $|\psi\rangle$ and $|\phi\rangle$. We also write $|\psi\rangle\langle\psi|$ for the pure quantum state in density matrix form. A general quantum state $\rho \in M_n(\mathbb{C})$ is a probabilistic mixture of pure quantum states, i.e. $\rho = \sum_{i=1}^m p_i|\psi_i\rangle\langle\psi_i|$ where $p_i \geq 0$ and $\sum_{i=1}^m p_i = 1$. If it is not a pure state, we also say that it is a mixed state. Alternatively, any $n \times n$ matrix ρ is a quantum state if it is positive semi-definite and satisfies $\mathsf{Tr}(\rho) = 1$.

If $\mathcal{H}_1$ and $\mathcal{H}_2$ describe two quantum systems, then the composite quantum system is described by the tensor product $\mathcal{H}_1 \otimes \mathcal{H}_2$. If $|\psi\rangle \in \mathcal{H}_1$ and $|\phi\rangle \in \mathcal{H}_2$ are pure states, then $|\psi\rangle \otimes |\phi\rangle \in \mathcal{H}_1 \otimes \mathcal{H}_2$ is a product state. Note that not all states in $\mathcal{H}_1 \otimes \mathcal{H}_2$ are product states.

A quantum channel $\Phi\colon M_n(\mathbb{C}) \to M_{n'}(\mathbb{C})$ is a linear map that is completely positive and trace-preserving (CPTP). Trace-preserving refers to the condition that $\mathsf{Tr}(\Phi(X)) = \mathsf{Tr}(X)$ for any $X \in M_n(\mathbb{C})$ and complete positivity requires that for any $m \in \mathbb{N}$ the map $\Phi \otimes \mathrm{Id}\colon M_n(\mathbb{C}) \otimes M_m(\mathbb{C}) \to M_{n'}(\mathbb{C}) \otimes M_m(\mathbb{C})$ is a positive map, where Id is the identity map sending $Y \in M_m(\mathbb{C})$ to Y. A general linear map $\Phi\colon M_n(\mathbb{C}) \to M_{n'}(\mathbb{C})$ is also called a *superoperator*. Note that the action of a superoperator is completely determined by its action on quantum states, since one can form a basis for $M_n(\mathbb{C})$ out of density matrices.

Let $|\psi\rangle \in \mathbb{C}^n$ be a pure quantum state given by $|\psi\rangle = \sum_{i\in[n]} \alpha_i|i\rangle$, where $(|i\rangle)_{i\in[n]}$ is a basis for $\mathbb{C}^n$. When we *measure* $|\psi\rangle$ in the basis $(|i\rangle)_{i\in[n]}$, the state *collapses* to $|i\rangle$ with probability $|\alpha_i|^2$.

The most general kind of measurement one can perform on a mixed state $\rho \in M_n(\mathbb{C})$ is a *positive operator-valued measure* (POVM). It is given by a set of positive semi-definite matrices $F_1, \ldots, F_m \in M_n(\mathbb{C})$ such that $\sum_{i=1}^m F_i = \mathrm{I}_n$, where I_n is the $n \times n$ identity matrix. The POVM element F_i is associated with measurement outcome i and obtaining this outcome happens with probability $\mathsf{Tr}(\rho F_i)$. A *projective measurement* is a POVM where the operators F_i are projectors, for example when the F_i are projectors on the i-th basis vector.

Observables are Hermitian matrices $H \in M_n(\mathbb{C})$. A special case is when the eigenvalues are ± 1, relevant for XOR games, in which case we write the set of all such matrices as $\mathrm{Obs}^{\pm}(\mathbb{C}^n)$.

Suppose $\mathcal{H}_1$ and $\mathcal{H}_2$ describe two quantum systems and $\rho \in \mathcal{L}(\mathcal{H}_1 \otimes \mathcal{H}_2)$ is a (mixed) quantum state in the composite system. Let $F_1, \ldots, F_m \in \mathcal{L}(\mathcal{H}_1)$ be a POVM for, say, the first quantum system. To perform this measurement on the first part of ρ, we apply the POVM given by the operators $F_1 \otimes \mathrm{I}, \ldots, F_m \otimes \mathrm{I}$ where $\mathrm{I} \in \mathcal{L}(\mathcal{H}_1)$ is the identity linear map.

1.2 Nonlocal games

This section serves as motivation and background information for Chapter 2. Here we follow the paper [CHTW04].

In Chapter 2, we consider *nonlocal games* with an arbitrary number of players. But for now we stick with the case of two players, where we call the two players Alice and Bob (as always). There is also a *referee* who samples questions from a known set according to a known probability distribution. The referee sends each of the players one question and Alice and Bob are not allowed to communicate once they receive it. In particular, they don't know each others questions. The goal of Alice and Bob is to separately send answers back to the referee such that they satisfy a known *predicate*, i.e. they win if their combined answer satisfies some condition. The set of questions, the probability distribution according to which the referee samples questions, the answer set and the predicate are all known before the game starts. This means that Alice and Bob can come together and devise a *strategy*.

Definitions. Write X and Y for the set of questions for Alice and Bob respectively and let π be a probability distribution on $X \times Y$. Let A and B be the set of answers that Alice and Bob can answer from and $V \colon A \times B \times X \times Y \to \{0, 1\}$ be a predicate. The pair (V, π) defines a nonlocal game $G = G(V, \pi)$ as follows. The referee picks a pair of questions $(x, y) \in X \times Y$ according to π and sends x to Alice and y to Bob. After receiving the questions, Alice and Bob must return an answer $a \in A$ and $b \in B$ without communicating. They are allowed to agree on a strategy before they receive the questions. Alice and Bob *win* the game if

the predicate V evaluates to 1 on the instance (a, b, x, y) and *lose* if it evaluates to 0.

The strategy that Alice and Bob agree on before the game starts, can be either *classical* or *quantum*. The *classical value* of the game is the maximum probability with which the players can win the game using a classical strategy. A classical strategy is simply given by two maps $a\colon X \to A$ and $b\colon Y \to B$, here the map a dictates Alice what to answer on a given question, similarly the map b determines what Bob should answer given a question. Such a strategy is *deterministic*, meaning that it doesn't use any source of randomness (shared or private). Then, the classical value of the game G, denoted by $\omega(G)$, is given by the following expression

$$\omega(G) = \max_{a,b} \sum_{(x,y)\in X\times Y} \pi(x,y)V(a(x), b(y)|x, y),$$

here the maximum is taken over maps $a\colon X \to A$ and $b\colon Y \to B$. Using more compact notation, this is also equal to

$$\omega(G) = \max_{a,b} \mathop{\mathbb{E}}_{(x,y)\in X\times Y} V(a(x), b(y)|x, y),$$

where the expectation is taken over the probability distribution π. Note that using shared or private randomness cannot increase the classical value, since the above maximum can always be achieved by a deterministic strategy.

A quantum strategy for Alice and Bob consists of the following.

- A bi-partite quantum state $|\psi\rangle \in \mathcal{A} \otimes \mathcal{B}$ where $\mathcal{A}$ and $\mathcal{B}$ are isomorphic copies of $\mathbb{C}^n$ for some n. Alice has the $\mathcal{A}$ part of $|\psi\rangle$ and Bob the $\mathcal{B}$ part.
- Two collections $\{P_x^a\colon x \in X, a \in A\}$ and $\{Q_y^b\colon y \in Y, b \in B\}$ of $n \times n$ matrices that form a POVM for every fixed $x \in X$ and $y \in Y$.

These ingredients turn into a quantum strategy for the game G as follows. When Alice gets question x, she performs the measurement described by the collection $\{P_x^a\colon a \in A\}$ on her part of the quantum state $|\psi\rangle$. Similarly if Bob gets question y, he performs the measurement given by $\{Q_y^b\colon b \in B\}$ on his part of $|\psi\rangle$. The probability that they obtain the pair of answers (a, b) upon measurement is $\langle\psi|P_x^a \otimes Q_y^b|\psi\rangle$. Then, the probability that they win the game G using this strategy is

$$\mathop{\mathbb{E}}_{(x,y)\in X\times Y} \sum_{(a,b)\in A\times B} \langle\psi|P_x^a \otimes Q_y^b|\psi\rangle V(a, b|x, y).$$

The *quantum value* of the game G, which we denote by $\omega^*(G)$, is the supremum of the above expression over all quantum strategies. It is not clear if one can achieve the supremum using finite-dimensional strategies. Here finite-dimensionality refers to the dimension of the quantum state $|\psi\rangle$. It is entirely

possible that there exists a sequence of strategies with increasing dimension that approaches the quantum value, but never reaches it in any finite dimension. Indeed, it has been shown quite recently in [Slo19] that there exists a nonlocal game which can be played perfectly (winning probability equals one) using a limit of finite-dimensional quantum strategies but which cannot be played perfectly using any finite-dimensional strategy.

1.2.1. Example. The *CHSH game* is a famous two-player game, based on the *CHSH inequality* [CHSH69], for which a quantum strategy does better than any classical strategy. The set of questions and answers are all equal to $\{0,1\}$. The distribution on the question set is uniform and the game tensor is given by

$$V(a,b|x,y) = \begin{cases} 1 & \text{if } a \oplus b = x \wedge y \\ 0 & \text{otherwise.} \end{cases}$$

The classical value of this game is 3/4. This can be seen by considering the following table.

x	y	Alice $\oplus$ Bob	$x \wedge y$
0	0	$a_0 \oplus b_0$	0
0	1	$a_0 \oplus b_1$	0
1	0	$a_1 \oplus b_0$	0
1	1	$a_1 \oplus b_1$	1

Here, a_x is the answer of Alice to question x and b_y the answer of Bob to question y. Assuming that there is a perfect strategy, i.e. $a_x \oplus b_y = x \wedge y$ for all $x, y \in \{0,1\}$, the sum of the rows of the third and fourth column should be equal. The sum of the rows of the third column is 0 (mod 2), but for the fourth column it is 1 (mod 2). This is a contradiction, so there cannot be a perfect classical strategy for this game. The best the players can do is win with probability 3/4. We will show in a bit that the quantum winning probability is equal to $\cos^2(\pi/8) \approx 0.85$.

XOR games. Of particular interest in this thesis are multiplayer XOR games. XOR games are a subclass of nonlocal games where the allowed answers are bits and the predicate only depends on the XOR of the anwers of the players. A t-player XOR game $G = (f, \pi)$ is defined by a function $f : X_1 \times \cdots \times X_t \to \{0,1\}$ and a probability distribution π over $X_1 \times \cdots \times X_t$ where $X_1, X_2, \ldots, X_t$ are question sets. An input $(x_1, \ldots, x_t) \in X_1 \times \cdots \times X_t$ is chosen by a referee according to π, who then gives x_i to player i. Without communicating, player i then outputs a bit $a_i \in \{0,1\}$ with the collective goal of the players being that $a_1 \oplus \cdots \oplus a_t = f(x_1, \ldots, x_t)$.

In XOR games, the notion of *bias* is more interesting than the winning probability: any XOR game can be won with probability 1/2 and this is achieved when

the players output a random bit. The bias of a strategy measures the amount with which we can play better than the "random strategy".

More precisely, let G be an XOR game and S a classical strategy. The bias of this strategy, which we denote by $\beta(G, S)$, is then defined as

$$\begin{aligned}\beta(G,S) &:= \Pr(\text{Win using strategy } S) - \Pr(\text{Lose using strategy } S)\\ &= 2\Pr(\text{Win using strategy } S) - 1.\end{aligned}$$

The classical bias of a game G, denoted $\beta(G)$, is then the maximum over all possible classical strategies S of $\beta(G, S)$. The quantum bias $\beta^*(G)$ is defined similarly, where we replace classical strategies with quantum strategies. An explicit expression for the bias can be obtained by letting the strategies and the function f take values in $\{\pm 1\}$ instead of $\{0, 1\}$. More precisely, we apply the transformation $a \mapsto (-1)^a$ for $a \in \{0, 1\}$. The classical bias is then given by

$$\beta(G) = \max_{a_i\colon X_i \to \{0,1\}} \mathop{\mathbb{E}}_{x \in X_1 \times \cdots \times X_t} (-1)^{\sum_{i=1}^t a_i(x_i) + f(x_1,\ldots,x_t)},$$

where the expectation is taken over the distribution π, this will always be the case in this setting unless otherwise stated. This follows from the following computation. First define

$$\mathcal{X} = \{x \in X_1 \times \cdots \times X_t \colon \sum_{i=1}^t a_i(x_i) = f(x_1, \ldots, x_t) \mod 2\},$$

then

$$\begin{aligned}&\mathop{\mathbb{E}}_{x \in X_1 \times \cdots \times X_t} (-1)^{\sum_{i=1}^t a_i(x_i) + f(x_1,\ldots,x_t)} = \mathop{\mathbb{E}}_{x \in X_1 \times \cdots \times X_t} \mathbb{1}_{\mathcal{X}}(x) - \mathbb{1}_{\overline{\mathcal{X}}}(x)\\ &= \Pr(\text{Win}) - \Pr(\text{Lose}).\end{aligned}$$

It is in fact more convenient to immediately assume that the strategies a_i are maps $X_i \to \{\pm 1\}$ and defining $T(x_1, \ldots, x_t) := (-1)^{f(x_1,\ldots,x_t)}$, so that

$$\beta(G) = \max_{a_i\colon X_i \to \{\pm 1\}} \mathop{\mathbb{E}}_{x \in X_1 \times \cdots \times X_t} T(x_1, \ldots, x_t) \prod_{i=1}^t a_i(x_i). \tag{1.2}$$

This is the form that is most convenient in Chapter 2. The map T is also called the *game tensor*.

To obtain a similar expression for the quantum bias, let $G = (f, \pi)$ be a two-player XOR game for simplicity. Consider the quantum strategy given by a shared entangled state $|\psi\rangle$ and projective measurements $\{P_x^0, P_x^1\}$ and $\{Q_y^0, Q_y^1\}$

for Alice and Bob. We can assume that their measurements are projective by the result in [CHTW04]. Then, the bias is given by

$$\begin{aligned}
&\Pr(\text{Win}) - \Pr(\text{Lose}) = \mathop{\mathbb{E}}_{x,y} \sum_{a,b\in\{0,1\}} (-1)^{f(x,y)+a+b} \langle\psi| P_x^a \otimes Q_y^b |\psi\rangle \\
&= \mathop{\mathbb{E}}_{x,y} (-1)^{f(x,y)} (\langle\psi| P_x^0 \otimes Q_y^0 |\psi\rangle - \langle\psi| P_x^0 \otimes Q_y^1 |\psi\rangle - \langle\psi| P_x^1 \otimes Q_y^0 |\psi\rangle + \langle\psi| P_x^1 \otimes Q_y^1 |\psi\rangle) \\
&= \mathop{\mathbb{E}}_{x,y} (-1)^{f(x,y)} \langle\psi| (P_x^0 - P_x^1) \otimes (Q_y^0 - Q_y^1) |\psi\rangle.
\end{aligned}$$

If we define $A_x := P_x^0 - P_x^1$ and $B_y := Q_y^0 - Q_y^1$, we see that the quantum bias of this strategy is given by

$$\mathop{\mathbb{E}}_{x,y} T(x,y) \langle\psi| A_x \otimes B_y |\psi\rangle,$$

where $T(x,y) := (-1)^{f(x,y)}$. The matrices A_x and B_y are ± 1-valued observables (Hermitian with eigenvalues ± 1). To get an expression for the quantum bias of the game G, we take supremum over all such observables and shared quantum state $|\psi\rangle$. Since the 2-norm of $|\psi\rangle$ is 1 and the matrices A_x, B_y are Hermitian, the supremum over such states can be neatly replaced by the operator norm, so

$$\beta^*(G) = \sup_{A,B} \| \mathop{\mathbb{E}}_{x,y} T(x,y) A_x \otimes B_y \|_{\text{op}}. \tag{1.3}$$

Here the supremum is taken over maps $A\colon X \to \text{Obs}^{\pm}(\mathbb{C}^N)$ and $B\colon Y \to \text{Obs}^{\pm}(\mathbb{C}^N)$ and integer $N \in \mathbb{N}$, where $\text{Obs}^{\pm}(\mathbb{C}^N)$ is the set of ± 1-valued observables in dimension N. Here N is the dimension of the part of the shared quantum state of Alice and Bob.

In a similar fashion, we can obtain an explicit expression for the quantum bias for any number of players. Let $G = (T, \pi)$ be a t-player XOR game given by a function $T\colon X_1 \times \cdots \times X_t \to \{\pm 1\}$ and π a probability distribution on $X_1 \times \cdots \times X_t$. The quantum bias can be conveniently written as

$$\beta^*(G) = \sup_{A_i} \| \mathop{\mathbb{E}}_{x \in X_1 \times \cdots \times X_t} T(x_1, \ldots, x_t) \otimes_{i=1}^t A_i(x_i) \|_{\text{op}}, \tag{1.4}$$

where the supremum is over maps $A_i\colon X_i \to \text{Obs}^{\pm}(\mathbb{C}^N)$ and $N \in \mathbb{N}$. Here $\text{Obs}^{\pm}(\mathbb{C}^N)$ is the set of ± 1-valued observables in dimension N.

1.2.2. Example. Continuing the example of the CHSH game, we will now show that there is a quantum strategy that wins this game with probablity $\cos^2(\pi/8) \approx 0.85$. Let the shared quantum state be $|\psi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ and the game tensor is given by $T\colon \{0,1\}^2 \to \{\pm 1\}\colon (x,y) \mapsto (-1)^{xy}$. Let $X = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ and $Y = \left(\begin{smallmatrix} 0 & -i \\ i & 0 \end{smallmatrix}\right)$. These matrices are ± 1-valued observables. Consider the following strategy: Let A_x and B_y be Alice and Bob's observables be given by $A_0 = X, A_1 = Y$ and

$B_0 = (X - Y)/\sqrt{2}, B_1 = (X + Y)/\sqrt{2}$ respectively. By definition of X, Y and $|\psi\rangle$ we have

$$\langle\psi|X \otimes X|\psi\rangle = 1, \quad \langle\psi|X \otimes Y|\psi\rangle = 0$$
$$\langle\psi|Y \otimes X|\psi\rangle = 0, \quad \langle\psi|Y \otimes Y|\psi\rangle = -1.$$

It follows that $\langle\psi|A_x \otimes B_y|\psi\rangle = (-1)^{xy}/\sqrt{2}$ which implies that the bias of this strategy is

$$\mathop{\mathbb{E}}_{x,y}(-1)^{xy}\langle\psi|A_x \otimes B_y|\psi\rangle = \frac{1}{\sqrt{2}}.$$

So the winning probability of this strategy is $\frac{1}{2} + \frac{1}{2\sqrt{2}} = \cos^2(\pi/8)$. In particular, this is strictly better than the classical winning probability 3/4. The fact that the quantum value of this game is also $\cos^2(\pi/8)$ follows from Tsirelson's inequality [TC80].

1.2.3. Example. *Line games* will be discussed in Chapter 2, but here we will look at a small example that can be obtained from a modification of the famous three-player Magic Square Game [Mer90, Mer93] which was analyzed in [IKP$^+$08]. Line games can be described by a simple geometric structure: the game is played over the plane $\mathbb{F}_3^2$ where the referee picks a line and sends the three consecutive points to each player. The predicate only depends on the direction of the line. In the Magic Square Game in [IKP$^+$08], the referee restricts to horizontal and vertical lines.

Grothendieck's inequality and two-player XOR games. There is a surprising connection between Grothendieck's inequality [Gro53], a fundamental result from Banach space theory, and the quantum/classical bias of two-player XOR games.

1.2.4. Definition. The Grothendieck constant K_G is the smallest real number such that for all $n \in \mathbb{N}$ the following holds. Let $M \in \mathbb{R}^{n \times n}$ be a real matrix such that for all $a, b \in [-1, 1]^n$ the inequality

$$\left|\sum_{s,t=1}^{n} M_{s,t} a_s b_t\right| \leq 1$$

holds. Then for all unit vectors $u_1, \ldots, u_n, v_1, \ldots, v_n \in \mathbb{R}^N$ (for any $N \in \mathbb{N}$)

$$\left|\sum_{s,t=1}^{n} M_{s,t}\langle u_s, v_t\rangle\right| \leq K_G.$$

Grothendieck showed that K_G does not depend on n. The exact value of K_G is unknown to this day. We do have the following bounds

$$1.6769... \leq K_G < \frac{\pi}{2\log(1+\sqrt{2})} \approx 1.7822.$$

The lower bound is from [Dav84] and independently [Ree91]. The upper bound is due to Krivine [Kri77] who conjectured that Grothendieck's constant is equal to the upper bound. But in [BMMN13], Braverman et al. show that Grothendieck's constant is strictly smaller than Krivine's bound, by some absolute constant.

The following theorem shows that for two-player XOR games, the quantum bias can be at most a constant factor larger than the classical bias and that this constant is given by K_G. It implies that quantum strategies in the two-player setting can not have an arbitrary large advantage over classical strategies.

1.2.5. Theorem ([Tsi87]). *Let $G = (T, \pi)$ be a two-player XOR game. Then*

$$\beta^*(G) \leq K_G \beta(G).$$

Proof:
Assume, without loss of generality, that the question sets X, Y have both size n. Define the matrix $M \in \mathbb{R}^{n\times n}$ to be

$$M_{x,y} = \frac{1}{\beta(G)} \pi(x,y) T(x,y).$$

It follows that for all $a, b \in \{\pm 1\}^n$ we have

$$\left| \sum_{x,y=1}^{n} M_{x,y} a_x b_y \right| \leq 1.$$

By convexity, this holds true for all $a, b \in [-1, 1]^n$. By Grothendieck's inequality, this implies that for all unit vectors $u_1, \ldots, u_n, v_1, \ldots, v_n \in \mathbb{R}^N$ (for any $N \in \mathbb{N}$)

$$\left| \mathop{\mathbb{E}}_{x,y} T(x,y) \langle u_x, v_y \rangle \right| \leq K_G \beta(G).$$

Tsirelson's correspondence [Tsi87] implies that maximizing over all real unit vectors $u_x, v_y \in \mathbb{R}^N$ for any $N \in \mathbb{N}$ yields the quantum bias on the left hand side. The result follows. □

Multiplayer XOR games. For three-player XOR games, such a result does not hold. In [PGWP+08] it was shown that there is a family of three-player XOR games $(G_i)_{i\in\mathbb{N}}$ such that $\beta^*(G_i)/\beta(G_i) \to \infty$ as $i \to \infty$. Later, this was

quantitatively improved in [BV13]. Both showed only the existence of such games, explicitly constructing such games is still an open problem. The games in these papers have the additional property that both $\beta^*(G_i)$ and $\beta(G_i)$ go to zero. In Chapter 2 we try to answer the question whether there is such a family, but with the extra condition that $\beta^*(G_i) \geq c$ for some $c > 0$. Such games have implications in communication complexity, see Chapter 2 for more details.

1.3 Quasirandom graphs

This section serves as motivation and background information for Chapter 3.

In a seminal paper [CGW89] Chung, Graham and Wilson showed that a set of seemingly different properties of graphs are equivalent for dense graphs. Two of these properties are *expansion* and *uniformity*. We will give the definitions of these properties and show how they are equivalent for certain classes of graphs. In Chapter 3 we will generalize these notions and apply them to quantum channels.

Expansion, uniformity and quasirandomness. We follow the survey the on expander graphs [HLW06] for the basic definitions. All graphs in this section are simple. Let $G = (V, E)$ be an undirected d-regular graph with vertex set V and edge set E on $|V| = n$ vertices. By d-regular we mean that each vertex has exactly d neighbours. For subsets $S, T \subset V$, denote by $E(S, T)$ the set of edges between S and T. Also, denote by $\partial(S) := E(S, \overline{S})$ the set of outgoing edges from S to its complement $\overline{S}$. The following is a combinatorial definition of expansion. The *edge expansion*, denoted $h(G)$ of the graph G is defined to be

$$h(G) = \min_{S \subset V\colon |S| \leq n/2} \frac{|\partial(S)|}{|S|}. \tag{1.5}$$

A large edge expansion $h(G) \geq \varepsilon > 0$ means that for any $S \subset V$ there are many edges between S and $\overline{S}$, at least an ε fraction of the size of S, i.e. the graph is highly connected. It is possible to define expansion in terms of how many vertices a set S is connected with in the complement $\overline{S}$. This is called *vertex expansion*, but we will not consider this in this dissertation.

1.3.1. Definition. A sequence of graphs $(G_i)_{i \in \mathbb{N}}$ is a *family of expander graphs* if there is an $\varepsilon > 0$ such that for all $i \in \mathbb{N}$ we have $h(G_i) \geq \varepsilon$.

Alternatively, one can define expansion in terms of the eigenvalues of the *adjacency matrix*. The (normalized) adjacency matrix of a d-regular graph G, denoted by $A = A(G)$, has its rows and columns indexed by the vertices $v \in V$ and the (u, v) entry is $1/d$ if there is an edge between u and v and 0 otherwise. The adjacency matrix A is real and symmetric, so it has n real eigenvalues which we denote (and order) by $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$. We have the following properties.

- $\lambda_1 = 1$ with eigenvector $e := (1, 1, \ldots, 1)$.
- The graph is connected if and only if $\lambda_2 < \lambda_1$.
- The graph is bipartite if and only if $\lambda_n = -\lambda_1$.

The first property follows from a small computation, while the second follows from the Perron-Frobenius theorem. For the third property, if G is a bipartite graph with left vertex set L and right vertex set R, then the vector with 1 on coordinates of L and -1 on coordinates of R is an eigenvector of A with eigenvalue $-\lambda_1 = -1$. In the other direction, if $x \in \mathbb{R}^V$ is a vector such that $Ax = -x$, we see that

$$0 = \langle x, x\rangle + \langle x, Ax\rangle = \frac{1}{d}\sum_{(v,w)\in E}(x_v + x_w)^2.$$

This implies that $x_v \neq 0$ for all $v \in V$, since assuming the existence of v such that $x_v = 0$ would imply that $x = 0$. We then define the left vertex set to be $L := \{v \in V : x_v > 0\}$ and $R := \{v \in V : x_v < 0\}$. It can be seen that there are no edges inside L or R, which means that G is bipartite.

We also refer to the eigenvalues of A as the *spectrum* of G. An important quantity of the spectrum is the second largest eigenvalue in absolute value, i.e. $\lambda(G) := \max\{|\lambda_2|, |\lambda_n|\}$. We say that G is an (n, d, λ)-graph if $\lambda(G) \leq \lambda$.

The following theorem relates the *spectral gap*, defined to be $d - \lambda_2$, with the edge expansion of G.

1.3.2. Theorem ([Che70], [Dod84]). *Let G be a d-regular graph on n vertices with spectrum $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$. Then*

$$\frac{d - \lambda_2}{2} \leq h(G) \leq \sqrt{d(d - \lambda_2)}.$$

In particular, an (n, d, λ)-graph has edge expansion at least $d\lambda/2$. A sequence of graphs $(G_i)_{i\in\mathbb{N}}$ is a family of *spectral expander graphs* if there is a $\lambda > 0$ such that $\lambda(G_i) \leq \lambda$ for each $i \in \mathbb{N}$. Such a family is also a family of expanders using the combinatorial definition of expansion.

Uniformity is a property that captures how random-like the edge distribution is between any two subsets of vertices. More precisely:

1.3.3. Definition. Let $G = (V, E)$ be a d-regular graph on n vertices and $\varepsilon > 0$. We say that G is ε-uniform if for all $S, T \subset V$:

$$\left|\,|E(S, T)| - \frac{d}{n}|S||T|\,\right| \leq \varepsilon dn.$$

Denote by $\varepsilon(G)$ the smallest such ε.

Having small uniformity parameter $\varepsilon(G)$ implies that for any pair of subsets S, T, the number of edges is roughly the same as one would expect from a random graph with edge density d/n. In this sense, uniform graphs, graphs that have small uniformity parameter $\varepsilon(G)$, are quasirandom. The following result relates expansion with uniformity, observed by several researchers but appeared first in print in [AC88], saying that a graph with small expansion parameter $\lambda(G)$ is quasirandom in the sense that it is a uniform graph.

1.3.4. Lemma (Expander mixing lemma). *Let $G = (V, E)$ be a d-regular graph on n vertices and set $\lambda = \lambda(G)$. Then, for all $S, T \subset V$*

$$\left| |E(S,T)| - \frac{d}{n}|S||T| \right| \leq \lambda d \sqrt{|S||T|}.$$

In particular, G is λ-uniform.

Proof:
For a subset $S \subset V$, write $\mathbb{1}_S \in \mathbb{R}^V$ for the indicator function of S. Let A be the adjacency matrix of G. Then for $S, T \subset V$

$$\langle \mathbb{1}_S, A\mathbb{1}_T \rangle = |E(S,T)|/d.$$

Expand the vectors $\mathbb{1}_S$ and $\mathbb{1}_T$ in the orthonormal basis of eigenvectors of A, i.e. if $v_1, \ldots, v_n$ are the eigenvectors corresponding to the eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$,

$$\mathbb{1}_S = \sum_{i=1}^{n} \alpha_i v_i \quad \text{and} \quad \mathbb{1}_T = \sum_{i=1}^{n} \beta_i v_i.$$

Then

$$\begin{aligned} |E(S,T)| = d\langle \mathbb{1}_S, A\mathbb{1}_T \rangle &= d \sum_{i=1}^{n} \lambda_i \alpha_i \beta_i \\ &= \frac{d}{n}|S||T| + d \sum_{i=2}^{n} \lambda_i \alpha_i \beta_i. \end{aligned}$$

We used the orthogonality of the vectors v_i and in the last equality the fact that $\alpha_1 = \langle \mathbb{1}_S, \frac{e}{\sqrt{n}} \rangle = \frac{|S|}{\sqrt{n}}$ and $\beta_1 = \langle \mathbb{1}_T, \frac{e}{\sqrt{n}} \rangle = \frac{|T|}{\sqrt{n}}$. Recall that $\lambda = \max\{|\lambda_2|, |\lambda_n|\}$. So

$$\begin{aligned} \left| |E(S,T)| - \frac{d}{n}|S||T| \right| = d \left| \sum_{i=2}^{n} \lambda_i \alpha_i \beta_i \right| &\leq \lambda d \sum_{i=2}^{n} |\alpha_i \beta_i| \\ &\leq \lambda d \sqrt{\sum_{i=2}^{n} |\alpha_i|^2 \sum_{i=2}^{n} |\beta_i|^2} \\ &\leq \lambda d \|\alpha\|_2 \|\beta\|_2 = \lambda d \|\mathbb{1}_S\|_2 \|\mathbb{1}_T\|_2 \\ &= \lambda d \sqrt{|S||T|}. \end{aligned}$$

□

The Expander mixing lemma tells us that an (n, d, λ)-graph is λ-uniform. The question then arises whether a λ-uniform graph is also an (n, d, λ')-graph for, possibly, a different constant λ' depending on λ. In [CGW89] it was shown that such a converse holds for *dense* graphs, i.e. graphs $G = (V, E)$ on n vertices such that $|E| = \Omega(n^2)$. For convenience, we prove this particular result here for d-regular graphs.

1.3.5. Proposition ([CGW89]). *Let $G = (V, E)$ be a d-regular graph on n vertices and let $\delta = d/n$. Then $\lambda(G) \leq (2\varepsilon(G)/\delta^3)^{1/4}$.*

Proof:
We prove this in two steps. In [CGW89] it is shown that if the number of 4-cycles in G is $(1 + o(1))d^4$, then $|\lambda_2| = o(1)$. We use the fact that $\mathsf{Tr}(A^4)$ counts the number of 4-cycles in G divided by d^4, where A is the normalized adjacency matrix of G. Let $\varepsilon > 0$ and assume

$$\mathsf{Tr}(A^4) = 1 + 2\varepsilon/\delta^3.$$

Note that δ is constant if G is a dense graph. We have that

$$\mathsf{Tr}(A^4) = \sum_{i=1}^{n} |\lambda_i|^4 = 1 + \sum_{i=2}^{n} |\lambda_i|^4.$$

By assumption

$$\sum_{i=2}^{n} |\lambda_i|^4 \leq 2\varepsilon/\delta^3,$$

and it follows that $|\lambda_2| \leq (2\varepsilon/\delta^3)^{1/4}$. The next step in the proof is to show that uniformity of the graph G implies the 'right' 4-cycle count. The proof of this is found at Tim Gowers's blog [Gow21]. Let $\varepsilon = \varepsilon(G)$ be the uniformity parameter of the graph G. We will now show that

$$\mathsf{Tr}(A^4) \leq 1 + 2\varepsilon/\delta^3.$$

In this case, it is more convenient to work with the unnormalized adjacency matrix, which we denote by A'. Let $v^{(l)}$ and $w^{(k)}$ be the l-th row and k-th column of A' respectively. We have that

$$\langle v^{(l)}, A'w^{(k)} \rangle \leq \langle v^{(l)}, \delta J w^{(k)} \rangle + \varepsilon dn = \delta \sum_{ij} A'_{li} A'_{jk} + \varepsilon dn,$$

where J is the all-ones matrix. This follows from the definition of the uniformity parameter ε. We use this as follows,

$$\begin{aligned} d^4 \,\mathsf{Tr}(A^4) = \mathsf{Tr}(A'^4) &= \sum_{ijkl} A'_{ij}A'_{jk}A'_{kl}A'_{li} = \sum_{kl} A'_{kl}\langle v^{(l)}, A'w^{(k)}\rangle \\ &\leq \sum_{kl} A'_{kl}\left(\delta \sum_{ij} A'_{li}A'_{jk} + \varepsilon dn\right) \\ &= \delta \sum_{ij} \langle v^{(i)}, A'w^{(j)}\rangle + \varepsilon dn^3 \\ &= \delta^2 \sum_{ijkl} A'_{jk}A'_{li} + 2\varepsilon dn^3 \\ &= d^4 + 2\varepsilon dn^3. \end{aligned}$$

From this we see that indeed $\mathsf{Tr}(A^4) \leq 1 + 2\varepsilon/\delta^3$. This completes the proof. □

This shows that we have a converse for dense d-regular graphs, i.e. if $\delta = \Omega(1)$, we have that $|\lambda_2| = O(\varepsilon(G)^{1/4})$.

But in [KS06] it was shown that the converse does not hold in general, thereby answering the question posed in [CG02]. In particular, they constructed a sequence of sparse graphs that is $o(1)$-uniform, but is $\Omega(1)$-expanding. Then, in [KRS16], it was shown that such a converse does in fact hold for Cayley graphs over finite abelian groups (even sparse ones) which in turn was generalized to work for *vertex-transitive* graphs in [CZ17]. A graph is vertex-transitive if the automorphism group of the graph acts transitively on the set of vertices. They also showed that vertex-transitivity is in fact a necessary condition by giving an example of a sparse graph that is not vertex-transitive such that the converse to the Expander mixing lemma does not hold.

Generalization to quantum channels. In Chapter 3 we will look at the relationship between the expansion parameter [Has07, BST10] and uniformity parameter for quantum channels, or superoperators in general. We will see that there is an analogue of the Expander mixing lemma in the quantum setting and a converse for randomizing quantum channels [Aub09]. Irreducible covariant quantum channels will play the role of vertex-transitive graphs in proving a converse to the Expander mixing lemma.

1.4 Higher-order Fourier analysis

This section serves as background information for Chapter 2 where we use higher-order Fourier analysis to analyze certain classes of XOR games, in Chapter 4 in the context of the stabilizer rank of the magic state, and Chapter 5 where we discuss

decompositions of dual functions. For more detailed information on higher-order Fourier analysis, we refer to [HHL19].

Higher-order Fourier analysis grew out of the Fourier-analytic proof of Szemerédi's theorem by Gowers [Gow01].

1.4.1. Theorem (Szemerédi's theorem). *Let $k \geq 2$ a positive integer and $\delta > 0$. There exists a positive integer N such that any subset $A \subset [N]$ of size at least δN contains a non-trivial arithmetic progression of length k.*

In this section, we will be mainly concerned with arithmetic progressions in subsets of vector spaces over finite fields, the main reason being (non-trivial) technicalities that arise in the setting of $[N]$.

Meshulam's theorem. Let $f, g \colon \mathbb{F}_p^n \to \mathbb{C}$ be functions. Recall the definition of the Fourier transform in Section 1.1. We have the following basic properties.

- $\langle f, g \rangle = \langle \widehat{f}, \widehat{g} \rangle$ (Plancherel),
- $\|f\|_2 = \|\widehat{f}\|_2$ (Parseval, follows directly from Plancherel),
- $\widehat{f * g}(z) = \widehat{f}(z)\widehat{g}(z)$ (convolution identity).

These results are already powerful enough to prove Roth's theorem [Rot53] on three-term arithmetic progression in subset of $[N]$. We will however discuss the analogous result where we replace $[N]$ with vector spaces over a finite field [Mes95] as the proof becomes much more clean.

1.4.2. Theorem (Meshulam's theorem). *Let p be an odd prime and $\delta > 0$. There is n_0 such that if $n \geq n_0$, any subset $A \subset \mathbb{F}_p^n$ of density $|A|/p^n \geq \delta$ contains a three-term arithmetic progression.*

Proof:
We will prove the statement using the "density increment strategy". Let $A \subset \mathbb{F}_p^n$. Note that x, y, z are in arithmetic progression if and only if $x + z = 2y$. Define $\mathbb{1}_{A_2}(x) := \mathbb{1}_A(x/2)$. Then, the (normalized) number of three-term arithmetic progressions is

$$\Lambda_3(A) := \mathop{\mathbb{E}}_{x,d} \mathbb{1}_A(x)\mathbb{1}_A(x+d)\mathbb{1}_A(x+2d) = \mathop{\mathbb{E}}_{x+z=2y} \mathbb{1}_A(x)\mathbb{1}_A(y)\mathbb{1}_A(z).$$

Write $\alpha := |A|/p^n$ for the density of A in $\mathbb{F}_p^n$. Then

$$\begin{aligned}\Lambda_3(A) &= \mathop{\mathbb{E}}_{x+z=y} \mathbb{1}_A(x)\mathbb{1}_A(y/2)\mathbb{1}_A(z) = \mathop{\mathbb{E}}_{y}(\mathbb{1}_A * \mathbb{1}_A)(y)\mathbb{1}_A(y/2)\\ &= \langle \mathbb{1}_A * \mathbb{1}_A, \mathbb{1}_{A_2}\rangle\\ &= \langle \widehat{\mathbb{1}_A}^2, \widehat{\mathbb{1}_{A_2}}\rangle\\ &= \sum_{z\in\mathbb{F}_p^n} \widehat{\mathbb{1}_A}(z)^2\overline{\widehat{\mathbb{1}_{A_2}}(z)}\\ &= \sum_{z\in\mathbb{F}_p^n} \widehat{\mathbb{1}_A}(z)^2\widehat{\mathbb{1}_A}(-2z)\\ &= \alpha^3 + \sum_{z\neq 0} \widehat{\mathbb{1}_A}(z)^2\widehat{\mathbb{1}_A}(-2z).\end{aligned}$$

From the second to the third line we used Plancherel. Then we used the convolution identity and in the last equality, we used that $\alpha = \widehat{\mathbb{1}_A}(0)$. What have we achieved here? The first expression for $\Lambda_3(A)$ was a sum over x, y, z that satisfies a linear equation, namely $x+z = 2y$. The last expression is just a sum without any constraints. This allows us to use another very useful tool: the Cauchy-Schwarz inequality.

$$\begin{aligned}|\sum_{z\neq 0} \widehat{\mathbb{1}_A}(z)^2\widehat{\mathbb{1}_A}(-2z)| &\leq \max_{z\neq 0}|\widehat{\mathbb{1}_A}(z)||\sum_{z\neq 0} \widehat{\mathbb{1}_A}(z)\widehat{\mathbb{1}_A}(-2z)|\\ &\leq \max_{z\neq 0}|\widehat{\mathbb{1}_A}(z)||\sum_{z\neq 0} \widehat{\mathbb{1}_A}^2(z)|^{1/2}|\sum_{z\neq 0} \widehat{\mathbb{1}_A}^2(-2z)|^{1/2}\\ &= \max_{z\neq 0}|\widehat{\mathbb{1}_A}(z)|\|\widehat{\mathbb{1}_A}\|_2^2\\ &= \alpha \max_{z\neq 0}|\widehat{\mathbb{1}_A}(z)|.\end{aligned}$$

This implies the following lower bound on the normalized count of three-term progressions in A,

$$\Lambda_3(A) \geq \alpha^3 - \alpha \max_{z\neq 0}|\widehat{\mathbb{1}_A}(z)|.$$

From this it is seen that if the Fourier coefficients of A are all small, for example that $\max_{z\neq 0}|\widehat{A}(z)| \leq \alpha^2/2$, then $\Lambda_3(A) \geq \alpha^3/2$. In this case we are done: the number of three-term progressions in A is at least $\alpha^3 n^2/2$. So assuming that A does not contain any three-term progression, this implies that A has a large non-zero Fourier coefficient. In this case, we define the function $f(x) := \mathbb{1}_A(x) - \alpha$ which has the property that for $z \neq 0$ we have $\widehat{f}(z) = \widehat{\mathbb{1}_A}(z)$. So there is a $z \neq 0$ such that

$$|\widehat{f}(z)| = |\mathop{\mathbb{E}}_{x} f(x)\chi_z(-x)| \geq \alpha^2/2.$$

Now, let $H = \{x \colon \langle z, x\rangle = 0\}$ be the kernel of χ_z. We can write the expectation as an expectation over hyperplanes $H + a$ on which χ_z is constant.

$$\alpha^2/2 \le |\mathop{\mathbb{E}}_x f(x)\chi_z| = |\mathop{\mathbb{E}}_a \mathop{\mathbb{E}}_{x\in H+a} f(x)\chi_z| \le \mathop{\mathbb{E}}_a |\mathop{\mathbb{E}}_{x\in H+a} f(x)|.$$

Since $\mathbb{E}_{x\in\mathbb{F}_p^n} f(x) = 0$, we can add this to the right hand side to get

$$\mathop{\mathbb{E}}_a |\mathop{\mathbb{E}}_{x\in H+a} f(x)| + \mathop{\mathbb{E}}_{x\in H+a} f(x) \ge \alpha^2/2.$$

By the averaging principle, there exists $a \in \mathbb{F}_p^n$ such that

$$|\mathop{\mathbb{E}}_{x\in H+a} f(x)| + \mathop{\mathbb{E}}_{x\in H+a} f(x) \ge \alpha^2/2.$$

Recal that $f(x) = \mathbb{1}_A(x) - \alpha$. This last inequality implies that

$$\mathop{\mathbb{E}}_{x\in H+a} \mathbb{1}_A(x) \ge \alpha + \alpha^2/4,$$

i.e. on the hyperplane $H+a$, the set A has a higher density. This is the sought for density increment. We can continue this argument a constant number of times, at some point the density will exceed 1 which is of course not possible. □

Quadratic Fourier analysis. To solve the problem of progressions of length four, Gowers [Gow98] developed quadratic Fourier analysis. In this case, we need that the prime $p \ge 5$. The main reason why Fourier analysis was useful for progressions of length three, was the identity

$$\begin{aligned}\Lambda_3(f,g,h) &:= \mathop{\mathbb{E}}_{x+z=2y} f(x)g(y)h(z)\\ &= \mathop{\mathbb{E}}_{x,y,z} f(x)g(y)h(z)\sum_{r\in\mathbb{F}_p^n}\chi_r(2y-x-z)\\ &= \sum_r \widehat{f}(r)\widehat{g}(r)\widehat{h}(-2r),\end{aligned}$$

which we proved for $f = g = h = \mathbb{1}_A$. This allowed us to use some basic inequalities that gave information about the Fourier coefficients of A and its relation with the number of three-term progressions. More explicitly, after Hölder's inequality and Cauchy-Schwarz inequality we have

$$|\Lambda_3(f,g,h)| \le \max_r |\widehat{f}(r)|\|g\|_2\|h\|_2.$$

Such an inequality does not work in the case of progressions of length four. Consider the following expression.

$$\Lambda_4(f_1,f_2,f_3,f_4) := \mathop{\mathbb{E}}_{x,d} f_1(x)f_2(x+d)f_3(x+2d)f_4(x+3d). \tag{1.6}$$

This expression can be quite large while at the same time the Fourier coefficients of the functions f_1, f_2, f_3, f_4 are tiny. For an example, consider the functions

$$f_1(x) = \omega^{\langle x,x\rangle}, f_2(x) = \omega^{-3\langle x,x\rangle}, f_3(x) = \omega^{3\langle x,x\rangle}, f_4(x) = \omega^{-\langle x,x\rangle},$$

where ω is a p-th root of unity. These functions have Fourier coefficients whose magnitude is $p^{-n/2}$, but $\Lambda_4(f_1, f_2, f_3, f_4) = 1$.

It seems that we need to find an inequality for Equation (1.6) that is consistent with the above example. Such an inequality must necessarily detect whether the functions correlate with *quadratic phase functions*: functions of the form $\omega^{q(x)}$ where q is a quadratic polynomial over $\mathbb{F}_p^n$. If our main tool in bounding the expression (1.6) is the Cauchy-Schwarz inequality (which was also the case in the three-term progression case), then there is only one possibility. For this, we need the Gowers U^3-norm which we defined in Definition 1.1.

The following result tells that we can bound (1.6) in terms of the U^3-norm of one of the functions. First, define for functions $f_i \colon \mathbb{F}_p^n \to \mathbb{C}$ for $i \in [s]$

$$\Lambda_s(f_1, \ldots, f_s) := \mathop{\mathbb{E}}_{x,d} f_1(x) f_2(x+d) \cdots f_s(x+(s-1)d)$$

1.4.3. Theorem (Generalized von Neumann inequality [TV06]). *Let $s \geq 2$ and $f_i \colon \mathbb{F}_p^n \to \mathbb{C}$ for $i \in [s]$ be functions that take values of modulus at most 1. Then*

$$|\Lambda_s(f_1, \ldots, f_s)| \leq \min_{i \in [s]} \|f_i\|_{U^s}. \tag{1.7}$$

We now give a sketch of how the proof of progressions of length four goes. For subset $A \subset \mathbb{F}_p^n$ of density α, define f such that $\mathbb{1}_A = f + \alpha$. If we expand

$$\Lambda_4(\mathbb{1}_A) := \Lambda_4(\mathbb{1}_A, \mathbb{1}_A, \mathbb{1}_A, \mathbb{1}_A) = \alpha^4 + \text{ 15 other terms},$$

we get that, besides the α^4 term, each term is of the form $\Lambda_4(f_1, f_2, f_3, f_4)$ with at least one of the f_i is equal to f. So using the Generalized von Neumann inequality on these 15 terms, we see that

$$\Lambda_4(\mathbb{1}_A) \geq \alpha^4 - 15\|f\|_{U^3}.$$

We now follow the same strategy as in the three-term progression case. Assuming $\|f\|_{U^3}$ is small, say smaller than $\alpha^4/2$, then the number of four-term progressions in A will be at least $\alpha^4 n^2/2$. Otherwise $\|f\|_{U^3}$ is large, say $\|f\|_{U^3} > \alpha^4/2$. In that case, we would like to say that f correlates with a character, so that we can apply the same density increment strategy. Unfortunately, it is not clear why this should be the case. But if we allow a broader class of functions with which f can correlate, then this is possible again. In this case, they are the quadratic phase functions. This is made precise in the following theorem.

1.4.4. THEOREM (Gowers inverse theorem for U^s [TZ12]). *Let $p \geq s$ be a prime and $f\colon \mathbb{F}_p^n \to \mathbb{D}$ a function such that $\|f\|_{U^s} > c$ for some $c > 0$. Then, there exists a polynomial $q \in \mathbb{F}_p[x_1, \ldots, x_n]$ of degree at most $s-1$ such that*

$$|\langle f, \omega^q \rangle| \geq \delta,$$

for some $\delta = \delta(s, p, c)$ independent of n.

Functions of the form ω^q for ω a p-th root of unity and q a polynomial of degree d are called *polynomial phase functions of degree d.*

1.4.5. REMARK. This theorem is actually true even if $p < s$, but one has to consider a larger class of polynomials, namely *nonclassical polynomials*. We will go in more detail as to what these objects are in a bit, since we will need it in analyzing the stabilizer rank of magic states in Chapter 4.

We apply the above for $s = 3$. So assume that $\|f\|_{U^3} > \alpha^4/2$ where $f = A - \alpha$. By the Gowers inverse theorem, we see that f correlates with some quadratic phase function ω^q. The idea is now to partition $\mathbb{F}_p^n$ into affine subspaces of large dimension on which q is constant to get the density increment, analogous to the three-term progression case.

Nonclassical polynomials. We will now define what nonclassical polynomials are. To this end, define the additive derivative operation Δ_h for any $h \in \mathbb{F}_2^n$ on functions $P\colon \mathbb{F}_2^n \to \mathbb{T}$ to be

$$\Delta_h P(x) := P(x+h) - P(x), \tag{1.8}$$

which we will also call the *derivative* in direction h.

Now let $P \in \mathbb{F}_p[x_1, \ldots, x_n]$ be a "classical" polynomial of degree $d \geq 1$. One can show that $\Delta_h P(x) = P(x+h) - P(x)$ is again a polynomial but of degree at most $d-1$. So after taking $d+1$ derivatives, the resulting polynomial will be the zero polynomial. Using this observation, we can define a broader class of polynomials as functions that take values in $\mathbb{T}$ and satisfy a condition on its derivatives.

1.4.6. DEFINITION. For an integer $d \geq 1$, a map $P\colon \mathbb{F}_p^n \to \mathbb{T}$ is a nonclassical polynomial of degree at most d if for all $h_1, \ldots, h_{d+1} \in \mathbb{F}_p^n$ we have

$$\Delta_{h_{d+1}} \cdots \Delta_{h_1} P(x) = 0. \tag{1.9}$$

The degree of P is the smallest such d.

Using the map $\iota\colon \mathbb{F}_p \to \mathbb{T}\colon x \to |x|/p$, one can view a polynomial $P \in \mathbb{F}_p[x_1, \ldots, x_n]$ as a map $\iota(P)\colon \mathbb{F}_p^n \to \mathbb{T}$. Nonclassical polynomials that arise in this way are called classical polynomials and they are a subset of the nonclassical polynomials. Note that they take values in $\frac{1}{p}\mathbb{Z}/\mathbb{Z}$. This confusing terminology has unfortunately become standard in the literature. The following example shows that this containment of classical polynomials in the set of nonclassical polynomials is indeed proper.

1.4.7. Example. Consider the map $P\colon \mathbb{F}_2^n \to \mathbb{T}$ be given by $x \mapsto |x|/4$. This is a nonclassical polynomial of degree two. To show this, we take derivatives:

$$|x+h|/4 - |x|/4 = |x|/4 + |h|/4 - |x \circ h|/2 - |x|/4 = |h|/4 - |x \circ h|/2 \mod 1,$$

where $\circ$ is entry-wise product of vectors. Here we used the property that for $a, b \in \mathbb{F}_2$,

$$|a+b| = |a| + |b| - 2|a||b|. \tag{1.10}$$

Taking one more derivative

$$\Delta_{h'}\Delta_h P(x) = -|x \circ h + h' \circ h|/2 + |x \circ h|/2 = -|h' \circ h|/2 \mod 1.$$

Indeed, $P(x) = |x|/4$ is a nonclassical polynomial of degree two. Note that it is *not* a classical polynomial since it takes values in $\frac{1}{4}\mathbb{Z}/\mathbb{Z}$.
In general, the polynomial $P(x) = |x|/2^k$ is a nonclassical polynomial of degree k. We will see later (Section 4.3.1) that for $k = 3$, this polynomial corresponds to the n-qubit magic state.

The above is a local definition of nonclassical polynomials. A global definition of classical polynomials of degree at most d is that they take the form $\sum_{i_1+\cdots+i_n \leq d} c_{i_1,\ldots,i_n} x_1^{i_1} \cdots x_n^{i_n}$. Similarly, nonclassical polynomials have the following global description.

1.4.8. Proposition ([TZ12]). *A map $P\colon \mathbb{F}_p^n \to \mathbb{T}$ is a nonclassical polynomial of degree at most d if and only if it has a representation of the form*

$$P(x_1, \ldots, x_n) = \alpha + \sum_{\substack{0 \leq i_1, \ldots i_n \leq 1; j \geq 0: \\ 0 < i_1 + \cdots + i_n \leq d - j(p-1)}} \frac{c_{i_1,\ldots,i_n,j} |x_1|^{i_1} \cdots |x_n|^{i_n}}{p^{j+1}}, \tag{1.11}$$

for some unique coefficients $c_{i_1,\ldots,i_n,j} \in \{0, 1, \ldots, p-1\}$ and $\alpha \in \mathbb{T}$. The maximal j in this decomposition is called the depth *of P.*

Note that the depth of a nonclassical polynomial is always at most $\lceil d/(p-1) \rceil - 1$ since at least one of the indices $i_1, \ldots, i_n$ must be positive. Using this proposition, one quickly sees that the polynomial $x \mapsto |x|/4$ for $x \in \mathbb{F}_2^n$ from the example above has degree two and depth 1. The polynomial $x \mapsto |x|/2^k$ has degree k and depth $k-1$.

1.5 Dual functions and decomposition theorems

In Chapter 5 we will be looking at *decomposition theorems* for a class of functions called *dual functions*. They are important in understanding certain refinements of Szemerédi's theorem, which we will define in a moment.

1.5.1. Definition. Let $k \geq 2$, $\mathbf{i} = (i_1, \dots, i_k) \in \mathbb{Z}^k_{\geq 0}$ and G a finite abelian group. The associated set of order-k dual functions is given by

$$\Delta_{\mathbf{i}} = \{\phi \colon y \mapsto \mathop{\mathbb{E}}_{x \in G} f_1(x + i_1 y) \cdots f_k(x + i_k y) | f_i \colon G \to \mathbb{D}\}.$$

For example, let $G = \mathbb{F}_p^n$, $\mathbf{i} = (0, 1, 2) \in \mathbb{Z}^3_{\geq 0}$ and $A \subset \mathbb{F}_p^n$. Then, the dual function

$$\phi(y) = \mathop{\mathbb{E}}_{x \in \mathbb{F}_p^n} \mathbb{1}_A(x) \mathbb{1}_A(x + y) \mathbb{1}_A(x + 2y)$$

gives the normalized count of three term progressions with common difference equal to y. By a decomposition theorem, we mean to write a certain (complicated) function as a linear combination of simpler functions plus an error term in a certain norm. As an example, consider dual functions of order two. For functions $f, g \colon \mathbb{F}_p^n \to \mathbb{D}$ and $\mathbf{i} = (i, j) \in \mathbb{Z}^2_{\geq 0}$, the associated dual function is

$$\phi(y) = \mathop{\mathbb{E}}_{x \in \mathbb{F}_p^n} f(x + iy) g(x + jy).$$

It follows after Fourier inversion that

$$\phi(y) = \sum_{\chi \in \widehat{\mathbb{F}_p^n}} \alpha_\chi \chi((j - i)y),$$

where the coefficients α_χ satisfy $\|\alpha\|_{\ell_1} = \sum_{\chi \in \widehat{\mathbb{F}_p^n}} |\alpha_\chi| \leq 1$.

The following proposition, provided to us by Shao [Sha20], tells us that we can write dual functions of any order in terms of polynomial phase functions of degree one lower than the order of the dual function. In contrast with the order two case, we have to allow an error term. It follows from an application of the Hahn-Banach theorem, the Generalized Von Neumann inequality and the Gowers inverse theorem.

1.5.2. Proposition. *Let $p \geq k + 1$ be a prime and let $G = \mathbb{F}_p^n$. Then, for any $\varepsilon > 0$ and $\mathbf{i} \in \mathbb{Z}^k_{\geq 0}$, there is an $M = M(\varepsilon, k, p) > 0$ such that any dual function $\phi \in \Delta_{\mathbf{i}}$ can be decomposed as*

$$\phi = \sum_{i=1}^{r} \alpha_i \psi_i + \tau, \tag{1.12}$$

where $\alpha_1, \dots, \alpha_r \in \mathbb{C}$ satisfy $|\alpha_1| + \cdots + |\alpha_r| \leq M$, $\psi_1, \dots, \psi_r$ are polynomial phases of degree at most $k - 1$ and $\|\tau\|_{L_1} \leq \varepsilon$.

Proof:
Define Φ to be the convex hull of $\Delta_{\mathbf{i}}$. We can assume without loss of generality that ϕ is real-valued: if ϕ is complex-valued, then $\Re(\phi) = (\phi + \overline{\phi})/2 \in \Phi$ and $\Im(\phi) = i(\overline{\phi} - \phi)/2 \in \Phi$. So once we show that the real and imaginary part of ϕ admits an L_1-decomposition, then ϕ itself must admit such a decomposition.

We identify real-valued functions on $\mathbb{F}_p^n$ as elements in $\mathbb{R}^N$ with $N = p^n$. Let ω be a p-th root of unity and let $K \subset \mathbb{R}^N$ be the convex hull generated by the functions of the form $\Re(\omega^{P(x)})$ and $\Im(\omega^{P(x)})$ where P is a polynomial of degree at most $k-1$. Let $K_\varepsilon \subset \mathbb{R}^N$ be the set of functions whose L_1-norm is at most ε. We wish to show that for a real-valued $\phi \in \Phi$ we have

$$\phi \in M \cdot K + K_\varepsilon,$$

for some large enough real constant M. Here $M \cdot K = \{cf \colon |c| \leq M, f \in K\}$. Assume now that this is not possible, then by a corollary of the Hahn-Banach theorem [Gow10, Corollary 3.3] there is an $f \in \mathbb{R}^N$ such that

- $\langle \phi, f \rangle > 1$,
- $\langle g, f \rangle \leq 1$ for all $g \in M \cdot K \cup K_\varepsilon$.

By definition of K_ε, the second property implies that $\|\varepsilon f\|_\infty \leq 1$. By (a more general version of) the Generalized Von Neumann inequality [TV06, Lemma 11.4], we have that

$$\langle \phi, \varepsilon f \rangle \leq \|\varepsilon f\|_{U^k}.$$

So by the first property above, we see that $\|\varepsilon f\|_{U^k} > \varepsilon$. Hence, by the Gowers inverse theorem 1.4.4, there exists a polynomial P of degree at most $k-1$ such that

$$|\langle \varepsilon f, \omega^P \rangle| \geq c_\varepsilon,$$

for some constant $c_\varepsilon > 0$. Then, for either $h = \pm\Re(\omega^P)$ or $h = \pm\Im(\omega^P)$ we have

$$\langle \varepsilon f, h \rangle \geq c_\varepsilon/2.$$

This contradicts the second property above whenever $M > 2\varepsilon/c_\varepsilon$. The result follows since $\Re(\omega^P) = \frac{1}{2}(\omega^P + \omega^{-P})$ and $\Im(\omega^P) = \frac{1}{2i}(\omega^{-P} - \omega^{-P})$. □

In the decomposition theorem that we just proved, the error term was small *on average*. Even though such decomposition theorems prove to be useful in higher-order Fourier analysis (see [Gow10]), a natural finite-field analog of a conjecture by Frantzikinakis [Fra16] (see also [Alt20]) asks whether such a decomposition is possible if we require $\|\tau\|_{L_\infty} \leq \varepsilon$.

1.5.3. Conjecture. *Proposition 1.5.2 holds with the error term τ satisfying $\|\tau\|_{L_\infty} \leq \varepsilon$.*

Truth of this conjecture has implications for *Szemerédi's theorem with random differences*, which we will now explain.

Does Szemerédi's theorem 1.4.1 still hold when we restrict the common difference to be in certain subsets of $[N]$? More precisely, let $A \subset [N]$ of density at least $\varepsilon > 0$ and $k \geq 3$. Are there subsets $D \subset [N]$ such that if N is large enough that A contains a proper k-term arithmetic progression $x, x+d, \ldots, x+(k-1)d$ with $d \in D$? In this setting, there are some known results. For example, the squares $\{1^2, 2^2, 3^2, \ldots\}$ [BL96] and the shifted primes $\{p-1 \colon p \text{ prime}\}$ and $\{p+1 \colon p \text{ prime}\}$ [WZ12] satisfy this. Frantzikinakis [FLW12] then asks if these sparse sets are special or that a random subset of $[N]$ also satisfies this property. We will be only focussing on the finite field setting, for which the problem statement is as follows.

1.5.4. Definition. Let $\rho, \varepsilon > 0$ and $k \geq 2$. Let $D \subset \mathbb{F}_p^n$ be a Bernoulli-ρ random subset, i.e. each element of $\mathbb{F}_p^n$ is picked with probability ρ to be in D independently of the others. Let E be the event that for any subset $A \subset \mathbb{F}_p^n$ of size at least εp^n contains a proper k-term arithmetic progression with common difference in D. We say that Szemerédi's theorem with random differences holds if $\Pr(E) \geq 1/2$.

1.5.5. problem. *In the setting of Definition 1.5.4, what is the smallest $\rho > 0$ such that Szemerédi's theorem with random differences hold?*

We write ρ_k for this smallest such ρ and also refer to it as the critical density. A connection with Locally Decodable Codes (LDC's) [BG18] shows that this probability ρ_k is upperbounded by $p^{-(1-o(1))n/\lceil k/2 \rceil}$. Truth of Conjecture 1.5.3 implies much stronger bounds for the probability ρ_k, namely [Alt20]

$$\rho_k \leq O(p^{-n} n^{k-1}).$$

But in Chapter 5 we show that Conjecture 1.5.3 is not true.

In Definition 1.5.4 we can replace the group $\mathbb{F}_p^n$ with the cyclic group $\mathbb{Z}_N$ for $N \in \mathbb{N}$ and ask the same question as in Problem 1.5.5. In this case, the best known upper bound on the critical density is $N^{-(1-o(1))/\lceil k/2 \rceil}$ [BG18].

In the setting of the cyclic groups, there is an analogous conjecture to 1.5.3, where instead of decomposing a dual function in terms of polynomial phase functions, we allow for a broader class of function called "nilsequences". Analogous to the finite-field setting, if this conjecture is true, it would imply far better upper bounds on the critical density in Szemerédi's theorem with random differences, namely $O_k(N^{-1}\log(N))$ [Alt20]. However, this conjecture is again false as

shown in [BG20]. This means that the strategy of proving good upper bounds on the critical density, in the finite-field and cyclic setting, through Conjecture 1.5.3 seems to be failing and new ideas are needed to tackle this problem of Szemerédi's theorem with random differences.

Chapter 2

Bounding quantum-classical separations of nonlocal games

This chapter is based on the paper [BBB+19] which is joint work with Tom Bannink, Harry Buhrman, Jop Briët and Troy Lee. We use the notation introduced in Section 1.2.

2.1 Introduction

The study of multiplayer games has been extremely fruitful in theoretical computer science across diverse areas including the study of complexity classes [BOGKW88], hardness of approximation [Kho02], and communication complexity [SZ08, LS09, BBLV13]. They are also a great framework in which to study Bell inequalities [Bel64] and analyze the nonlocal properties of entanglement. A particularly simple kind of multiplayer game is an XOR game. Recall that an XOR game $G = (f, \pi)$ between t players is defined by a function $f : X_1 \times X_2 \times \cdots \times X_t \to \{0,1\}$ and a probability distribution π over $X_1 \times \cdots \times X_t$. An input $(x_1, \ldots, x_t) \in X_1 \times \cdots \times X_t$ is chosen by a referee according to π, who then gives x_i to player i. Without communicating, player i then outputs a bit $a_i \in \{0,1\}$ with the collective goal of the players being that $a_1 + \cdots + a_t = f(x_1, \ldots, x_t) \mod 2$. In an XOR game without entanglement, the players' strategies are deterministic. In an XOR game with entanglement, players are allowed to share a quantum state and make measurements on this state to inform their outputs.

Our motivating question in this chapter is the following:

2.1.1. Question. *Is there a family of t-player XOR games $(G_n)_{n\in\mathbb{N}}$ such that $\beta^*(G_n) = 1$ and $\beta(G_n) \to 0$ as $n \to \infty$?*

This question has important implications for multi-party communication complexity. For $t \geq 2$, in t-party communication complexity, t players have to compute some function $f : X_1 \times \cdots X_t \to \{0,1\}$. Player i receives $x_i \in X_i$

and together they have to compute $f(x_1, \ldots, x_t)$ while communicating as little as possible. They are allowed to use shared randomness and local computation is free. Usually, the value $f(x_1, \ldots, x_t)$ will depend on all variables x_i, so that communication is necessary. Let $R(f)$ denote the t-party randomized communication complexity of f, that is the minimal number of bits of communication that is necessary to compute f using shared randomness and local computation. Also, let $R^*(f)$ denote the t-party randomized communication complexity of f where the parties are allowed to share entanglement. A positive answer to Question 2.1.1 gives a family of functions $(f_n)_{n\in\mathbb{N}}$ with $R^*(f_n) = O(1)$ and $R(f_n) = \omega(1)$, i.e. an unbounded separation between these two communication models.

In the reverse direction, a family of functions $(f_n)_{n\in\mathbb{N}}$ with $R^*(f_n) = O(1)$ and $R(f_n) = \omega(1)$ gives a family of games $G_n = (f_n, \pi_n)$ with $\beta^*(G_n) \geq c$ for some constant c and $\beta(G_n) \to 0$ as $n \to \infty$. Thus there is a very close connection between Question 2.1.1 and the existence of an unbounded separation between randomized communication complexity with and without entanglement.

As we discussed in Section 1.2, it is known that for two-player XOR games the answer to Question 2.1.1 is negative. Linial and Shraibman [LS09] and Shi and Zhu [SZ08] showed that the bias of an XOR game (f, π) can be used to lower bound the communication complexity of f, both in the randomized setting and the setting with entanglement. Together with Grothendieck's inequality they used this to show that $R(f) = O(2^{2R^*(f)})$ for any partial two-party function f. Thus in the two-party case an unbounded communication separation is not possible between the randomized model with and without entanglement. Raz has given an example of a partial function f with $R(f) = 2^{\Omega(R^*(f))}$ [Raz99], thus the upper bound of Linial-Shraibman and Shi-Zhu is essentially optimal.

In the case of three or more parties, Question 2.1.1 and the corresponding question of an unbounded separation between the entangled and non-entangled communication complexity models remain open. We already saw in Section 1.2 that there is no analogue of Grothendieck's inequality in the three-player setting. In particular, Peréz-García et al. [PGWP$^+$08] showed that there exists an infinite family of three-player XOR games $(G_n)_{n\in\mathbb{N}}$ with the property that the ratio of the entangled and classical biases of G_n goes to infinity with n. This result was later quantitatively improved by Briët and Vidick [BV13]. Both results rely crucially on non-constructive (probabilistic) methods, and in both separating examples the entangled bias $\beta^*(G_n)$ also goes to zero with increasing n. These works leave open the question, posed explicitly in [BV13], of whether there is such a family of games in which the entangled bias does not vanish with n, but instead stays above a fixed positive threshold while the classical bias decays to zero. Crucially, having a separation in XOR bias where $\beta^*(G_n)$ remains constant is what is needed to also obtain an unbounded separation between randomized communication complexity with and without entanglement.

Our contribution to answering Question 2.1.1 One approach to Question 2.1.1 is to look at different classes of games and identify which ones could possibly lead to a positive answer.

Peréz-García et al. [PGWP$^+$08] show that in any XOR game where the entangled strategy uses a GHZ state, there is a bounded gap between the classical and entangled bias: namely, the bias with a GHZ state in a t-player XOR game G is at most $K_G(2\sqrt{2})^{t-1}\beta(G)$. This bound is essentially tight as there are examples of t-player XOR games achieving a ratio between the GHZ state bias and classical bias of $\frac{\pi}{2}t$ [Zuk93]. Briët et al. [BBLV13] later extended the Grothendieck- type inequality of Peréz-García et al. to a larger class of entangled states called Schmidt states. Thus any game that can be played perfectly with a strategy where the players share a Schmidt state cannot give a positive answer to Question 2.1.1.

Watts et al. [WHKN18] recently investigated Question 2.1.1 and found that a t-player XOR game G that is symmetric, i.e. invariant under the renaming of players, and where $\beta^*(G) = 1$ always has a perfect entangled strategy where the players share a GHZ state. Thus symmetric games also cannot give a positive answer to Question 2.1.1. Subsequently, Watts et al. [WH20] showed that in the 3-player setting, the symmetry condition on the game can be dropped.

We rule out other types of games that could positively answer Question 2.1.1 as well. A t-player *free* XOR game $G = (f, \pi)$ is a game where π is a product distribution. For such games we show that $\beta(G) \geq \beta^*(G)^{2^t}$, and thus they cannot be used for a positive answer to Question 2.1.1.

Another class of XOR games we consider are *line games*, where the questions asked to the players are related by a geometric property. An example of a line game is a slight modification of the Magic Square game [IKP$^+$08]. We show that line games cannot give a positive answer to Question 2.1.1 either.

In the next subsections, we discuss our results in more detail.

2.1.1 Free XOR games

In this subsection we identify two types of games, namely free games and line games, for which either the ratio of the entangled and classical biases is small, or the entangled bias itself is small. Thus these games will not be able to give a positive answer to Question 2.1.1. Free games are a general and natural class of games in which the players' questions are independently distributed. Line games appear to be less studied (see below for their definition), but turn out to be relevant in the context of parallel repetition (also see below). The main idea behind these results is that a large entangled bias implies that the games are in a sense far from random. This is quantified by the magnitude of certain norms of the game tensors. The particular norms of interest here are related to norms used in Gowers' celebrated hypergraph- and Fourier-analytic proofs of Szemerédi's Theorem. A crucial fact of these norms is that they are large if and only if there is "correlation with structure", the opposite of what one would

expect from randomness. We show that this structure can be turned into good classical strategies, thus establishing a relationship between the entangled and classical biases.

2.1.2. THEOREM (Polynomial bias relation for free XOR games). *For integer $t \geq 2$ and any free t-player XOR game with entangled bias β, the classical bias is at least β^{2^t}.*

In free XOR games all questions have a non-zero probability of being asked, so this result may be considered as an analogue of a well-known result on quantum query algorithms for total functions: in [BBC$^+$01] it is shown that the bounded-error quantum and classical query complexities of total functions are polynomially related.

2.1.2 Line games

Line games are not free, but have a simple geometric structure. For a finite field $\mathbb{F}$ of characteristic at least t and positive integer n, a t-player line game is given by a map $\tau : \mathbb{F}^n \to \{0, 1\}$. In the game, the referee independently samples two uniformly random points $x, y \in \mathbb{F}^n$ and sends the point $x + (i-1)y$ to the ith player. The players win the game if and only if the XOR of their answers equals $\tau(y)$. In other words, the players' questions correspond to consecutive points (or an arithmetic progression) on a random affine line through $\mathbb{F}^n$ and the winning criterion depends only on the direction of the line. Refer to this as a line game *over* $\mathbb{F}^n$.

2.1.3. THEOREM. *For any $\varepsilon \in (0, 1]$, integer $t \geq 2$ and finite field $\mathbb{F}$ of characteristic at least t, there exists a $\delta(\varepsilon, t, \mathbb{F}) \in (0, 1]$ such that the following holds. For any positive integer n and any t-player line game over $\mathbb{F}^n$ with entangled bias ε, the classical bias is at least $\delta(\varepsilon, t, \mathbb{F})$.*

Note that in the above result, the value of the classical bias is independent of the dimension n of the vector space determining the players' question sets.

Parallel repetition. While it is not relevant to Question 2.1.1, the proof techniques used for Theorem 2.1.3 allow us to prove a parallel repetition theorem for a class of games that include line games. The k-fold parallel repetition for $k \geq 1$ of a t-player XOR game $G = (f, \pi)$, which we denote by $G^k = (f^k, \pi^k)$, is a game where t-tuples of questions $(q_1^i, q_2^i, \ldots, q_t^i)$ for $i = 1, \ldots, k$ are sampled from the distribution π and sent to the players all at once. The players have to play the XOR game G on each t-tuple of questions. To win the game, they have to answer correctly on each t-tuple of questions.

It is known that the value of free games and so-called anchored games decays exponentially under parallel repetition. Dinur et al. [DHVY16] identified

a general criterion of multi-player games to behave like this, encompassing free and anchored games. They showed that it is sufficient for a certain graph that can be obtained from a game to be expanding, a well-known quasirandom property that gives a measure of graph connectivity. Line games do not belong to this class, as their graphs are not even connected. However, we show that if a map $\tau : \mathbb{F}^n \to \{0,1\}^n$ is quasirandom in a different sense, then a line game defined by τ has exponential decaying value under parallel repetition. More generally, we show that this is the case for a family of XOR games over an arbitrary finite abelian group Γ. These games are given by a positive integer m, a family of affine linear maps $\psi_0, \dots, \psi_t : \Gamma^m \to \Gamma$ such that no two are multiples of each other, and a "game map" $\rho : \Gamma \to \{0,1\}$. In the game, the referee samples a uniform random element x from Γ^m and sends the group element $\psi_i(x)$ to the ith player. The winning criterion is given by $\rho(\psi_0(x))$. The relevant notion of quasiranomness is quantified by the Gowers t-uniformity norm of the map $(-1)^\rho : x \mapsto (-1)^{\rho(x)}$.

2.1.4. Lemma. *Let m, t be positive integers and let Γ be a finite abelian group. Let $\psi_0, \dots, \psi_t : \Gamma^m \to \Gamma$ be affine linear maps such that no two are multiples of each other and let $\rho : \Gamma \to \{0,1\}$. Let G be the t-player XOR game given by the system $\{\psi_0, \dots, \psi_t, \rho\}$. Then, for every positive integer k,*

$$\omega(G^k) \leq \Big(\frac{1 + \|(-1)^\rho\|_{U^t}}{2}\Big)^k.$$

2.2 Techniques

This section provides an overview of the proof techniques that we use. We give sketches of the main ideas which are worked out in full detail in later sections.

2.2.1 Norming hypergraphs and quasirandomness

Our main tool for proving Theorem 2.1.2 is a relation between the entangled and classical biases and a norm on the set of game tensors. For t-tensors, this norm is given in terms of a certain t-partite hypergraph H. Recall that such a hypergraph consists of t finite and pairwise disjoint vertex sets $V_1, \dots, V_t$ and a collection of t-tuples $E(H) \subseteq V_1 \times \cdots \times V_t$, referred to as the edge set of H. For a t-tensor $T \in \mathbb{R}^{n_1 \times \cdots \times n_t}$, the norm has the following form:

$$\|T\|_H = \Big(\mathbb{E}_{\phi_i : V_i \to [n_i]} \Big[\prod_{(v_1,\dots,v_t) \in E(H)} T\big(\phi_1(v_1), \dots, \phi_t(v_t)\big) \Big]\Big)^{\frac{1}{|E(H)|}}, \tag{2.1}$$

where the expectation taken with respect to the uniform distribution over all t-tuples of mappings ϕ_i from V_i to $[n_i]$. Expressions such as (2.1) play an important role in the context of graph homomorphisms [BCL$^+$06]. If T is the adjacency

matrix of a bipartite graph with left and right node sets $[n_1]$ and $[n_2]$ respectively, then each product in (2.1) is 1 if and only if the maps ϕ_1 and ϕ_2 preserve edges. In other words, the expression in (2.1) counts the number of graph homomorphisms from the graph to itself.

Criteria for H under which (2.1) defines a norm or a semi-norm were determined by Hatami [Hat10, Hat09] and Conlon and Lee [CL17]. Famous examples of graph norms include the Schatten-p norms for even $p \geq 4$ (in which case H is a p-cycle) and a well-known family of hypergraph norms are the Gowers octahedral norms. The latter were introduced for the purpose of quantifying a notion of quasirandomness of hypergraphs as an important part of Gowers' graph-theoretic proof of the multi-dimensional version of Szemerédi's theorem on arithmetic progressions [Gow07]. Having large Gowers norm turns out to imply *correlation with structure*, as opposed to quasirandomness. This is true also for the norm relevant for our setting. In particular, it turns out that the structure with which a game tensor correlates can be turned into a classical strategy for the game. As such, a large norm of the game tensor implies a large classical bias of the game itself. At the same time, we show that the entangled bias is bounded from above by the norm of the game tensor, provided the game is free. Putting these observations together gives the proof of Theorem 2.1.2, which we give in Section 2.3.

The particular hypergraph norm relevant in our setting was introduced in [CHPS12] and can be obtained recursively as follows. Starting with a t-partite hypergraph H with vertex set $V_1 \cup \cdots \cup V_t$, write $\mathrm{db}_i(H)$ for $i \in [t]$ for the t-partite hypergraph obtained by making two vertex-disjoint copies of H and gluing them together so that the vertices in the two copies of V_i are identified. We obtain our hypergraph by starting with a single edge $e = (v_1, \ldots, v_t)$ (and vertex sets of size 1), and applying this operation to all parts, forming the hypergraph $\mathrm{db}_1(\mathrm{db}_2(\ldots \mathrm{db}_t(e)))$ with vertex sets of size 2^{t-1} and 2^t edges. The fact that this hypergraph defines a norm via (2.1) was proved in [CL17].

2.2.2 Line games and Gowers uniformity norms

The proof of Theorem 2.1.3 is based on two fundamental results from additive combinatorics: the generalized von Neumann inequality and the Gowers Inverse Theorem. The former easily shows that the classical bias of a line game is bounded from above by the Gowers t-uniformity norm of the game map. We show that in fact the same upper bound holds for the entangled bias as well. A large entangled bias thus implies a large uniformity norm for the game map. Analogous to the above-mentioned octahedral norms for tensors, uniformity norms were introduced to quantify a notion of quasirandomness for bounded maps over abelian groups as an important step in Gowers' other proof of Szemerédi's Theorem, based on higher-order Fourier analysis. The highly non-trivial Gowers Inverse Theorem of Tao and Ziegler [TZ12] establishes that high uniformity norm again implies correlation with structure. Although structure in this context means something

quite different than for tensors, we show that it still implies a lower bound on the classical bias. The above observations together prove Theorem 2.1.3.

2.3 Free XOR games

In this section we will go over the details and techniques needed to prove Theorem 2.1.2. The main lemma that we will prove is a relation between the hypergraph norm (with respect to a certain hypergraph) of the game tensor of a free XOR game and its quantum bias. Our main tool is a Cauchy-Schwarz type of inequality for operators.

2.3.1. Proposition. *Let $A_i, B_i \in M_n(\mathbb{C})$ for $i = 1, \dots, k$. Then*

$$\|\sum_{i\in[k]} A_i B_i\| \leq \|\sum_{i\in[k]} A_i A_i^*\|^{1/2} \, \|\sum_{i\in[k]} B_i^* B_i\|^{1/2},$$

where all the norms are operator norms.

Proof:
Define E_{ij} to be the matrix that has 1 in the (i, j) position of the matrix and 0 everywhere else. Write

$$A = \begin{bmatrix} A_1 & A_2 & \cdots & A_k \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} B_1 & 0 & \cdots & 0 \\ B_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ B_k & 0 & \cdots & 0 \end{bmatrix}.$$

Then we use the properties that for $C \in M_n(\mathbb{C})$

$$\|C \otimes E_{ij}\| = \|C\| \quad \text{and} \quad \|C\|^2 = \|C^*C\| = \|CC^*\|,$$

to conclude

$$\begin{aligned} \|\sum_{i\in[k]} A_i B_i\| &= \|\sum_{i\in[k]} A_i B_i \otimes E_{11}\| = \|AB\| \leq \|A\|\|B\| \\ &= \|AA^*\|^{1/2}\|B^*B\|^{1/2} = \|\sum_{i\in[k]} A_i A_i^*\|^{1/2}\|\sum_{i\in[k]} B_i^* B_i\|^{1/2}. \end{aligned}$$

□

A t-player free XOR game G is given by finite non-empty sets $X_1, \dots, X_t$, a product distribution π over $X_1 \times \cdots \times X_t$ and a game tensor

$$T\colon X_1 \times \cdots \times X_t \to \{\pm 1\}. \tag{2.2}$$

Recall from Section 1.2 that the classical bias of the free XOR game G, which we denote by $\beta(G)$ is given by

$$\beta(G) = \max_{a_i \colon X_i \to \{\pm 1\}} | \mathop{\mathbb{E}}_{(x_1,\ldots,x_t)\sim\pi} T(x_1,\ldots,x_t) \prod_{i=1}^{t} a_i(x_i)|,$$

and that the quantum bias, which we denote by $\beta^*(G)$, is given by the expression

$$\beta^*(G) = \sup_{N\in\mathbb{N}, A_i \colon X_i \to \mathrm{Obs}^{\pm}(\mathbb{C}^N)} \| \mathop{\mathbb{E}}_{(x_1,\ldots,x_t)\sim\pi} T(x_1,\ldots,x_t) \otimes_{i=1}^{t} A_i(x_i)\|_{\mathrm{op}}. \tag{2.3}$$

For $n \in \mathbb{N}$ define

$$\beta^*(G,n) := \max_{N\leq n, A_i \colon X_i \to \mathrm{Obs}^{\pm}(\mathbb{C}^N)} \| \mathop{\mathbb{E}}_{(x_1,\ldots,x_t)\sim\pi} T(x_1,\ldots,x_t) \prod_{i=1}^{t} A_i(x_i)\|_{\mathrm{op}}. \tag{2.4}$$

The supremum is taken over $\pm$-observable valued functions A_i such that $[A_i, A_j] = 0$ for $i \neq j$, this corresponds to a quantum strategy of the players in the commuting operator model. Since the commuting operator model and the tensor product model for quantum strategies coincide in finite dimension [Tsi06], we have that $\lim_{n\to\infty} \beta^*(G,n) = \beta^*(G)$. The expectation is taken over the given distribution, which we will suppress in the notation from here onwards.

Before we go into the details of the proof of Theorem 2.1.2 for any number of players, we first sketch the core idea of the proof for *two* players, for which we do not yet need to resort to hypergraphs. For a two-player game G with game tensor T, the commuting-operator strategies A, B yield a bias of

$$\eta = \| \mathop{\mathbb{E}}_{(x,y)\in X\times Y} T(x,y)A(x)B(y)\|.$$

where the norm is the operator norm. Using Proposition 2.3.1 we peel off the operator $B(y)$

$$\begin{aligned}
\eta &= \Big\| \mathop{\mathbb{E}}_{y\in Y} \Big(\mathop{\mathbb{E}}_{x\in X} T(x,y)A(x)\Big)B(y)\Big\|. && \text{(independent questions)}\\
&\leq \Big\| \mathop{\mathbb{E}}_{y\in Y} \Big(\mathop{\mathbb{E}}_{x\in X} T(x,y)A(x)\Big)\Big(\mathop{\mathbb{E}}_{x'\in X} T(x',y)A(x')\Big)^*\Big\|^{1/2} \quad \Big\| \mathop{\mathbb{E}}_{y\in Y} B(y)B(y)^*\Big\|^{1/2}\\
&\leq \Big\| \mathop{\mathbb{E}}_{y\in Y} \mathop{\mathbb{E}}_{x,x'\in X} T(x,y)T(x',y)A(x)A(x')^*\Big\|^{1/2}. && \text{(using } \|B(y)\| \leq 1)
\end{aligned}$$

Now we apply the inequality again on the sum over (x,x') to get rid of the A operator.

$$\begin{aligned}
\eta &\leq \Big\| \mathop{\mathbb{E}}_{x,x'\in X} \Big(\mathop{\mathbb{E}}_{y\in Y} T(x,y)T(x',y)\Big)A(x)A(x')^*\Big\|^{1/2}\\
&\leq \Big| \mathop{\mathbb{E}}_{x,x'} \big(\mathop{\mathbb{E}}_{y} T(x,y)T(x',y)\big)\big(\mathop{\mathbb{E}}_{y'} T(x,y')T(x',y')\big)\Big|^{\frac{1}{4}} \Big\| \mathop{\mathbb{E}}_{x,x'} \big(A(x)A(x')^*\big)\big(A(x)A(x')^*\big)^*\Big\|^{\frac{1}{4}}\\
&\leq \Big| \mathop{\mathbb{E}}_{x,x'\in X} \mathop{\mathbb{E}}_{y,y'\in Y} T(x,y)T(x',y)T(x,y')T(x',y')\Big|^{1/4}. && \text{(using } \|A(x)\| \leq 1)
\end{aligned}$$

We proceed by rewriting the last expression

$$\begin{aligned}\eta^4 &\leq \Big| \mathop{\mathbb{E}}_{(x',y')\in X\times Y} T(x',y') \mathop{\mathbb{E}}_{(x,y)\in X\times Y} T(x,y)T(x',y)T(x,y') \Big| \\ &\leq \mathop{\mathbb{E}}_{(x',y')\in X\times Y} \Big| \mathop{\mathbb{E}}_{(x,y)\in X\times Y} T(x,y)T(x',y)T(x,y') \Big|. \qquad \text{(triangle inequality)}\end{aligned}$$

By the averaging principle there must be choices of x', y' such that

$$\eta^4 \leq \Big| \mathop{\mathbb{E}}_{(x,y)\in X\times Y} T(x,y)T(x',y)T(x,y') \Big|,$$

which is the expression for the bias of the classical strategies $a(x) = T(x,y')$ and $b(y) = T(x',y)$, proving Theorem 2.1.2 for $t = 2$ players. For $t \geq 3$ we can apply the same idea, peeling off the operators one by one, but the final expression is more involved. We will now develop the techniques to deal with this. In particular, we need the notion of hypergraph norms. For our purposes, we only consider *t-partite* hypergraphs.

2.3.2. Definition. For $t \geq 2$, let $V_1, \ldots, V_t$ be finite non-empty sets and $V := V_1 \times \cdots \times V_t$. Given a subset $E \subset V$, we say that the pair $H = (V_1 \cup \cdots \cup V_t, E)$ is a t-partite hypergraph with vertex set $V_1 \cup \cdots \cup V_t$ and edge set E. The set V is the set of vertices and E the set of hyperedges.

Here t-partite refers to the property that the set of vertices can be partitioned into t parts such that each hyperedge contains exactly one vertex from each part.

2.3.3. Definition. Let $t \geq 2$ and $X_1, \ldots, X_t$ be finite non-empty sets and suppose a product distribution on $X := X_1 \times \cdots \times X_t$ is given. Let $T\colon X \to \mathbb{R}$ be a function and $H = (V_1 \cup \cdots \cup V_t, E)$ be a t-partite hypergraph. We define a non-negative function $\|\cdot\|_H$ on the function T by

$$\|T\|_H := \left| \mathop{\mathbb{E}}_{\phi_i\colon V_i \to X_i} \prod_{(v_1,\ldots,v_t)\in E} T(\phi_1(v_1), \ldots, \phi_t(v_t)) \right|^{\frac{1}{|E|}}, \tag{2.5}$$

where the expectation is taken with respect to the following distribution: a particular map $\phi_i\colon V_i \to X_i$ occurs with probability $\prod_{v\in V_i} p_i(\phi_i(v))$ where p_i is the probability distribution on X_i.

The particular hypergraph which arises naturally when we study the quantum bias of free XOR games is constructed as follows. Starting with a t-partite hypergraph H, write $\mathrm{db}_i(H)$ for the t-partite hypergraph obtained by making two vertex-disjoint copies of H and gluing them together so that the vertices in the two copies of V_i are identified. To construct our hypergraph, we start with the hypergraph given by a single edge $e = (v_1, \ldots, v_t)$ and vertex sets of size 1 and

apply the doubling operation to all parts, i.e. $\mathrm{db}_1(\mathrm{db}_2(\ldots \mathrm{db}_t(e)))$. We denote this hypergraph by $H(t)$. A more useful way to define $H(t)$ is as follows. We will do this first for $t = 2$ and explain how to do it for any t afterwards. We use 2-bit strings to label vertices. We start with the hypergraph with a single edge $(x_{00}, y_{00}) \in V_1 \times V_2$. As we will start using the doubling operator, we make copies of the vertex sets. We can use a table to visualize it.

	V_1	V_2
starting position	x_{00}	y_{00}
db_2	x_{01}	y_{00}
db_1	x_{00}	y_{10}
	x_{01}	y_{10}

The table may be read as follows; the rows are the edges of the hypergraph and columns are the vertex sets. In this example we have that $V_1 = \{x_{00}, x_{01}\}$ and $V_2 = \{y_{00}, y_{10}\}$ and the edge set consists of $\{(x_{00}, y_{00}), (x_{01}, y_{00}), (x_{00}, y_{10}), (x_{01}, y_{10})\}$. The algorithm for constructing the table is as follows: we start with the starting position row, which corresponds to the (hyper)graph with a single edge (x_{00}, y_{00}), and as we apply the doubling operator db_2, we add a new row (which corresponds to making a vertex-disjoint copy) where we increase the second bit in the subscript of x but leave y alone (so we have a new copy of V_1 but not of V_2). After this first step we have a graph with vertex sets $V_1 = \{x_{00}, x_{01}\}$ and $V_2 = \{y_{00}\}$ and edge set $\{(x_{00}, y_{00}), (x_{01}, y_{00})\}$. Next we apply db_1 and we get a new copy of V_2, but leave V_1 alone.

For arbitrary $t \geq 2$, let v^i_ω be a formal variable with $i \in [t]$ and ω a t-bit string. We define for $j \in [t]$ an operation Δ_j on the formal variable by

$$\Delta_j(v^i_\omega) := v^i_{\omega_1,\ldots,\omega_j+1,\ldots,\omega_t} \text{ for } j \neq i$$
$$\Delta_i(v^i_\omega) := v^i_\omega.$$

where we add modulo 2. The table then looks like

	V_1	V_2	$\ldots$	V_t
starting position	$v^1_{0^t}$	$v^2_{0^t}$	$\ldots$	$v^t_{0^t}$
db_t	$\Delta_t(v^1_{0^t})$	$\Delta_t(v^2_{0^t})$	$\ldots$	$\Delta_t(v^t_{0^t})$
db_{t-1}	$\Delta_{t-1}(v^1_{0^t})$	$\Delta_{t-1}(v^2_{0^t})$	$\ldots$	$\Delta_{t-1}(v^t_{0^t})$
	$\Delta_{t-1}(\Delta_t(v^1_{0^t}))$	$\Delta_{t-1}(\Delta_t(v^2_{0^t}))$	$\ldots$	$\Delta_{t-1}(\Delta_t(v^t_{0^t}))$
$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$

At step k, the algorithm takes all the rows of the previous steps together and applies Δ_{t-k+1} on each of the formal variables in the rows. We also write $\mathrm{db}_i(e)$ for the row where we apply Δ_i on each variable of the row e. We see in this way that, for example, the edge set of $H(t)$ has cardinality 2^t and the number of vertices in each V_i is 2^{t-1}. In the following proposition we list some properties of $H(t)$ which we prove using this description. We will be using the terms row and edge interchangeably as they mean the same in this context.

2.3.4. Proposition. *The hypergraph $H(t)$ has the following properties: (1) it is t-partite, (2) it is 2-regular and (3) for all vertices v the following holds: let e, e' be the unique edges such that $v \in e$, $v \in e'$ and $e \neq e'$. For $w \in e \setminus \{v\}$, denote by e, e'' the unique edges such that $w \in e$, $w \in e''$ and $e \neq e''$. Then $e' \cap e'' = \emptyset$.*

Proof:
(1) follows directly from the algorithm described above using the table. We can prove (2) as follows. Suppose in column V_i we have a vertex in some row/edge which we denote by v^i_ω, here ω is a t-bit string. First we note that applying db_j with $j \neq i$ will change ω as it will flip the j-th bit. There are two cases; either we have already applied db_i in which case v^i_ω appears in exactly one more row above the current row, or we have not applied db_i yet in which case there is no v^i_ω in an earlier row. It will appear exactly once in a later row since applying db_i will not change ω. For (3), choose again some vertex v^i_ω in V_i and denote by e the row which appears first in the table containing v^i_ω. The other row/edge which contains v^i_ω is $e' := \mathrm{db}_i(e)$. Now, let v^j_τ be a vertex in V_j with $j \neq i$ and $v^j_\tau \in e$, i.e. it is in the same row as v^i_ω. There are two cases; either $j > i$ in which case $e = \mathrm{db}_j(e'')$ where e'' is the other (unique) edge containing v^j_τ. Or $j < i$ and the other edge which contains v^j_τ is $e'' := \mathrm{db}_j(e)$. In either case, a moments thought shows that $e' \cap e'' = \emptyset$. □
The next ingredient is the following lemma.

2.3.5. Lemma. *For a t-player free XOR game G with game tensor T, we have that*

$$\beta^*(G) \leq \|T\|_{H(t)}.$$

Proof:
For convenience, we write the hypergraph $H(t)$ in a slightly different way. Recall that the sets $X_1, \ldots, X_t$ are the set of questions in the the game G and $V_1 \cup \cdots \cup V_t$ is the vertex set of $H(t)$. For $i, j \in [t]$, let $\phi_i \colon V_i \to X_i$ and we define an operation Δ_j on such maps in the same way as above, i.e.

$$(\Delta_j \phi_i)(v^i_\omega) = \phi_i(\Delta_j v^i_\omega) \text{ for } j \neq i$$
$$(\Delta_i \phi_i)(v^i_\omega) = \phi_i(v^i_\omega).$$

Also, using the same symbol, we define on functions $T \colon X_1 \times \cdots \times X_t \to \mathbb{C}$

$$\Delta_j T(\phi_1(v^1_{\omega^1}), \ldots, \phi_t(v^t_{\omega^t})) := T(\phi_1(v^1_{\omega^1}), \ldots, \phi_t(v^t_{\omega^t})) T^*((\Delta_j\phi_1)(v^1_{\omega^1}), \ldots, (\Delta_j\phi_t)(v^t_{\omega^t})),$$

one could think of this operation as a kind of multiplicative derivative. If T were an operator-valued map, we still define it in this way. It is then not hard to see that

$$\Delta_1 \ldots \Delta_t T(\phi_1(v^1_{0^t}), \ldots, \phi_t(v^t_{0^t})) = \prod_{(v^1_{\omega^1}, \ldots, v^t_{\omega^t}) \in E(H(t))} T(\phi_1(\omega^1), \ldots, \phi_t(\omega^t)),$$

using the table as a description of $H(t)$. So we can write

$$\|T\|_{H(t)} = |\,\mathbb{E}\,\Delta_1 \dots \Delta_t T(\phi_1(v_{0^t}^1), \dots, \phi_t(v_{0^t}^t))|^{1/|E|},$$

where the expectation is taken over all maps $\phi_i \colon V_i \to X_i$ with the particular distribution given in Definition 2.3.3.
Now let us look at the bias of a two-player game G with game tensor T and finite-dimensional strategies A, B

$$\eta = \| \mathop{\mathbb{E}}_{(x,y)\in X\times Y} T(x,y)A(x)B(y)\|_{\mathrm{op}}.$$

We will do the example of two players to clarify the idea and will prove it in general afterwards. First we use Proposition 2.3.1, the Cauchy-Schwarz inequality, to peel off strategy B

$$\begin{aligned}\eta &= \| \mathop{\mathbb{E}}_{y\in Y} \big(\mathop{\mathbb{E}}_{x\in X} T(x,y)A(x)\big)B(y)\|_{\mathrm{op}} \\ &\le \| \mathop{\mathbb{E}}_{y} \big(\mathop{\mathbb{E}}_{x} T(x,y)A(x)\big)\big(\mathop{\mathbb{E}}_{x} T(x,y)A(x)\big)^*\|_{\mathrm{op}}^{1/2} \;\; \| \mathop{\mathbb{E}}_{y} B(y)^*B(y)\|_{\mathrm{op}}^{1/2} \\ &\le \| \mathop{\mathbb{E}}_{y,x,x'} T(x,y)T(x',y)A(x)A(x')^*\|_{\mathrm{op}}^{1/2}.\end{aligned}$$

In the second inequality we used that operator norm of strategies are smaller or equal to 1. We will use the Cauchy-Schwarz inequality one more time, now to peel off strategy A

$$\begin{aligned}\eta &\le \| \mathop{\mathbb{E}}_{x,x'} (\mathop{\mathbb{E}}_{y} T(x,y)T(x',y))A(x)A(x')^*\|_{\mathrm{op}}^{1/2} \\ &\le \| \mathop{\mathbb{E}}_{x,x'} (\mathop{\mathbb{E}}_{y} T(x,y)T(x',y))(\mathop{\mathbb{E}}_{y} T(x,y)T(x',y))^*\|_{\mathrm{op}}^{1/4} \;\; \| \mathop{\mathbb{E}}_{x,x'} (A(x)A(x')^*)^*A(x)A(x')^*\|_{\mathrm{op}}^{1/4} \\ &\le |\mathop{\mathbb{E}}_{x,x',y,y'} T(x,y)T(x',y)T(x,y')T(x',y')|^{1/4}.\end{aligned}$$

Now, to see that this last expression is equal to $\|T\|_{H(2)}$, we write the expectation in a different way. Instead of writing $\mathbb{E}_{x,x'}$ we write $\mathbb{E}_{\phi\colon V\to X}$ where $V = \{v_0, v_1\}$ is a vertex set, so that $x = \phi(v_0)$ and $x' = \phi(v_1)$. Similarly, instead of $\mathbb{E}_{y,y'}$ we write $\mathbb{E}_{\psi\colon W\to Y}$ where $W = \{w_0, w_1\}$ and we view $H(2)$ to be on this vertex sets. Then, we evaluate T on the edges of $H(2)$, so

$$\begin{aligned}&|\mathop{\mathbb{E}}_{x,x',y,y'} T(x,y)T(x',y)T(x,y')T(x',y')|^{1/4} \\ &\qquad\qquad = |\mathop{\mathbb{E}}_{\phi\colon V\to X,\psi\colon W\to Y} \prod_{(v,w)\in E(H(2))} T(\phi(v),\psi(w))|^{1/4}.\end{aligned}$$

In general, for t players, the proof is as follows

$$\begin{aligned}\eta :=& \| \mathbb{E}\, T(\phi_1(v_{0^t}^1), \ldots, \phi_t(v_{0^t}^t)) \left(\prod_{i\in[t-1]} A_i(\phi_i(v_{0^t}^i)) \right) A_t(\phi_t(v_{0^t}^t)) \|_{\mathrm{op}} \\ &\leq \| \mathbb{E}\, T(\phi_1(v_{0^t}^1), \ldots, \phi_t(v_{0^t}^t)) T(\phi_1(v_{0^{t-1}1}^1), \ldots, \phi_t(v_{0^t}^t)) \prod_{i\in[t-1]} A_i(\phi_i(v_{0^t}^i)) A_i(\phi_i(v_{0^{t-1}1}^i))^* \|_{\mathrm{op}}^{1/2} \\ &= \| \mathbb{E}\, \Delta_t T(\phi_1(v_{0^t}^1), \ldots, \phi_t(v_{0^t}^t)) \prod_{i\in[t-1]} \Delta_t A_i(\phi_i(v_{0^t}^i)) \|_{\mathrm{op}}^{1/2}.\end{aligned}$$

Now assume that we have applied the Cauchy-Schwarz inequality $1 < n < t$ times to peel off the last n operators and we have obtained the expression

$$\eta \leq \| \mathbb{E}\, \Delta_{t-n+1} \cdots \Delta_t T(\phi_1(v_{0^t}^1), \ldots, \phi_t(v_{0^t}^t)) \prod_{i\in[t-n]} \Delta_{t-n+1} \cdots \Delta_t A_i(\phi_i(v_{0^t}^i)) \|_{\mathrm{op}}^{1/2^n}.$$

Now apply Cauchy-Schwarz inequality to remove the operator $\Delta_{t-n+1} \cdots \Delta_t A_{t-n}(\phi_i(v_{0^t}^i))$ so that we obtain

$$\eta \leq \| \mathbb{E}\, \Delta_{t-n} \cdots \Delta_t T(\phi_1(v_{0^t}^1), \ldots, \phi_t(v_{0^t}^t)) \prod_{i\in[t-n-1]} \Delta_{t-n} \cdots \Delta_t A_i(\phi_i(v_{0^t}^i)) \|_{\mathrm{op}}^{1/2^{n+1}}.$$

This completes the induction. Putting $n = t - 1$ we have the inequality

$$\eta \leq | \mathbb{E}\, \Delta_1 \cdots \Delta_t T(\phi_1(v_{0^t}^1), \ldots, \phi_t(v_{0^t}^t)) |^{1/2^t}.$$

We have shown that for each $m \in \mathbb{N}$, that $\beta^*(G, m) \leq \|T\|_{H(t)}$. This implies the result. □

We are now ready to give a proof of Theorem 2.1.2.

Proof of Theorem 2.1.2:
We assume $\beta^*(G) > \eta$. Lemma 2.3.5 immediately implies $\|T\|_{H(t)} \geq \eta$. To construct a classical strategy, we choose an edge $e^* = (v_1^*, \ldots, v_t^*) \in E(H(t))$. Any choice of edge works for our argument. We have that $H(t)$ is 2-regular (by Proposition 2.3.4), so denote by e_i^* the unique edge different from e^* such that $v_i^* \in e_i^*$. Write $e_i^* = (v_1^{(i)}, \ldots, v_i^*, \ldots, v_t^{(i)})$ and $V_i' := V_i \setminus \{v_i^*\}$. Using Proposition 2.3.4 we

see that $v_j^* \notin e_i^*$ whenever $i \neq j$. Then

$$\begin{aligned}
\eta^{2^t} &\leq |\mathop{\mathbb{E}}_{\phi_i\colon V_i \to X_i} \prod_{(v_1,\ldots,v_t)\in E} T(\phi_1(v_1),\ldots,\phi_t(v_t))| \\
&= |\mathop{\mathbb{E}}_{\phi_i\colon V_i' \to X_i}[\prod_{(v_1,\ldots,v_t)\in E\setminus\{e^*,e_1^*,\ldots,e_t^*\}} T(\phi_1(v_1),\ldots,\phi_t(v_t)) \\
&\mathop{\mathbb{E}}_{\phi_i^*\colon \{v_i^*\}\to X_i} T(\phi_1^*(v_1^*),\ldots,\phi_t(v_t^*))T(\phi_1^*(v_1^*),\ldots,\phi_t(v_t^{(1)}))\cdots T(\phi_1(v_1^{(t)}),\ldots,\phi_t^*(v_t^*))]| \\
&\leq \mathop{\mathbb{E}}_{\phi_i\colon V_i' \to X_i} |\mathop{\mathbb{E}}_{\phi_i^*\colon \{v_i^*\}\to X_i} T(\phi_1^*(v_1^*),\ldots,\phi_t^*(v_t^*))T(\phi_1^*(v_1^*),\ldots,\phi_t(v_t^{(1)})) \\
&\cdots T(\phi_1(v_1^{(t)}),\ldots,\phi_t^*(v_t^*))|.
\end{aligned}$$

Let us explain the second and third line in detail. Write $V_i = V_i' \cup \{v_i^*\}$. Any map $\phi_i\colon V_i \to X_i$ can be given by two maps $\phi_i'\colon V_i' \to X_i$ and $\phi_i^*\colon \{v_i^*\} \to X_i$ by defining $\phi_i(v)$ to be $\phi_i'(v)$ when $v \in V_i'$ and otherwise equal to $\phi_i^*(v)$. It can then be seen that

$$\mathop{\mathbb{E}}_{\phi_i\colon V_i\to X_i}(\text{some expression}) = \mathop{\mathbb{E}}_{\phi_i'\colon V_i'\to X_i}[\mathop{\mathbb{E}}_{\phi_i^*\colon \{v_i^*\}\to X_i}(\text{some expression})].$$

After this we use the triangle inequality. By the averaging principle we see that there exist specific choices of maps $\phi_i\colon V_i' \to X_i$ such that

$$|\mathop{\mathbb{E}}_{\phi_i^*\colon \{v_i^*\}\to X_i} T(\phi_1^*(v_1^*),\ldots,\phi_t^*(v_t^*))T(\phi_1^*(v_1^*),\ldots,\phi_t(v_t^{(1)}))\cdots T(\phi_1(v_1^{(t)}),\ldots,\phi_t^*(v_t^*))| > \eta^{2^t}.$$

The expectation over t-tuples of maps $\phi_i^*\colon \{v_i^*\} \to X_i$ is the same as the expectation over t-tuples $x_i^* \in X_i$ and by defining

$$a_i(x_i^*) := T(\phi_1(v_1^{(i)}),\ldots,x_i^*,\ldots,\phi_t(v_t^{(i)}))$$

we see that

$$|\mathop{\mathbb{E}}_{x_1^*,\ldots,x_k^*} T(x_1^*,\ldots,x_k^*)\prod_{i=1}^k a_i(x_i^*)| \geq \eta^{2^t},$$

in other words, the classical bias is at least η^{2^t}. □

2.4 Linear forms game

Before we will go in to the details of the proof of Theorem 2.1.3, we briefly discuss some concepts from higher-order Fourier analysis. The reference for this subsection is [Tao12].

2.4.1 Preliminaries

Recall the notation as in Section 1.1. Let G be a finite abelian group and $f\colon G \to \mathbb{C}$ a complex-valued function on G. Recall the definition of the Gowers U^s-norm in Definition 1.1. Now let $\psi_0, \dots, \psi_t\colon G^d \to G$ be affine linear forms, i.e. maps that take the form $\psi_i(g_1, \dots, g_d) = c_i + \sum_{j=1}^d c_{ij} g_j$ where $c_i \in G$ and $c_{ij} \in \mathbb{Z}$.

2.4.1. Definition. Let $\{\psi_0, \dots, \psi_t\}$ be a system of affine linear forms. We say that the system has Cauchy-Schwarz complexity at most s if for any $0 \leq i \leq t$ one can partition $\{\psi_0, \dots, \psi_t\} \setminus \{\psi_i\}$ into $s+1$ classes (empty classes are allowed) such that ψ_i does not lie in the affine linear span (over $\mathbb{Q}$) of the forms in any of these classes. The Cauchy-Schwarz complexity of the system is defined to be the least such s or ∞ if no such s exists.

If $\psi\colon G^d \to G$ is an affine linear form, we denote by $\dot{\psi}\colon \mathbb{Q}^d \to \mathbb{Q}$ the map induced by its integer coefficients. The characteristic of G is defined to be least order of all non-identity elements. Here is an equivalent formulation of Cauchy-Schwarz complexity in terms of change of variables.

2.4.2. Proposition. *Let $\{\psi_0, \dots, \psi_t\}$ be a system of affine linear forms $G^d \to G$. Suppose that the characteristic of G is sufficiently large depending on the coefficients of $\psi_0, \dots, \psi_t$. Then the system has Cauchy-Schwarz complexity at most s if and only if for every $0 \leq i \leq t$ one can find a linear change of variables $\vec{x} = L_i(y_1, \dots, y_{s+1}, z_1, \dots, z_d)$ on $\mathbb{Q}^d$ such that the form $\dot{\psi}_i(L_i(y_1, \dots, y_{s+1}, z_1, \dots, z_d))$ has non-zero $y_1, \dots, y_{s+1}$ coefficients, but all other forms $\dot{\psi}_j(L_i(y_1, \dots, y_{s+1}, z_1, \dots, z_d))$ with $j \neq i$ have at least one vanishing $y_1, \dots, y_{s+1}$ coefficient.*

We also need the Gowers inverse theorem 1.4.4.

2.4.2 Quantum/classical bias ratio for line games

We will now continue with line games. Line games, as discussed in the introduction, fall inside a larger class of games which we will describe first. For this, let Γ be a finite abelian group, $m \geq 1$ an integer and let $\psi_0, \dots, \psi_t\colon \Gamma^m \to \Gamma$ for $i = 1, \dots, t$ be $t+1$ affine linear forms, i.e.

$$\psi_i(g_1, \dots, g_m) = c_i + \sum_{j=1}^m c_{ij} g_j$$

where $(g_1, \dots, g_m) \in \Gamma^m$, $c_i \in \Gamma$ and $c_{ij} \in \mathbb{Z}$.

2.4.3. Definition. A t-player linear forms game is given by the above data together with a game map $\rho\colon \Gamma \to \{0, 1\}$ as follows. The referee samples a uniform random point g from Γ^m and sends $\psi_i(g)$ to player i (players are numbered from 1 to t). The winning criterion is given by $\rho(\psi_0(g))$.

Let G be such a game. The classical bias is given by

$$\beta(G) = \max_{a_i \colon \Gamma \to \{\pm 1\}} |\mathop{\mathbb{E}}_{g \in \Gamma^m} (-1)^{\rho(\psi_0(g))} \prod_{i=1}^{t} a_i(\psi_i(g))|.$$

The quantum bias is

$$\beta^*(G) = \sup_{N \geq 1, A_i \colon \Gamma \to \mathrm{Obs}^{\pm 1}(\mathbb{C}^N)} \| \mathop{\mathbb{E}}_{g \in \Gamma^m} (-1)^{\rho(\psi_0(g))} \otimes_{i=1}^{t} A_i(\psi_i(g))\|_{\mathrm{op}}.$$

We will use again that $\lim_{n\to\infty} \beta^*(G,n) = \beta^*(G)$ where

$$\beta^*(G,n) = \max_{N \leq n, A_i \colon \Gamma \to \mathrm{Obs}^{\pm 1}(\mathbb{C}^N)} \| \mathop{\mathbb{E}}_{g \in \Gamma^m} (-1)^{\rho(\psi_0(g))} \prod_{i=1}^{t} A_i(\psi_i(g))\|_{\mathrm{op}},$$

where the maximization is over A_i such that $[A_i, A_j] = 0$ for $i \neq j$.

2.4.4. Remark. A t-player line game as discussed in the introduction falls inside this framework where the finite group is $\Gamma = \mathbb{F}_p^n$ and the linear forms $\psi_i \colon (\mathbb{F}_p^n)^2 \to \mathbb{F}_p$ are given by $\psi_0(x,y) = y$ and $\psi_i(x,y) = x + (i-1)y$ for $i = 1, \ldots, t$.

The main technical theorem of this section is the following.

2.4.5. Theorem. *Let G be a game as above. If the Cauchy-Schwarz complexity of $\{\psi_0, \ldots, \psi_t\}$ is at most s, we then have the inequality*

$$\beta^*(G) \leq \|(-1)^{\rho}\|_{U^{s+1}(\Gamma)}. \tag{2.6}$$

This theorem is a corollary of the following two results that are immediate generalizations from the "commutative" setting. See [Tao12] for more details.

2.4.6. Lemma (Second Gowers-Cauchy-Schwarz inequality for operators). *For a function $f \colon \Gamma \to \mathbb{C}$ and maps $A_i \colon \Gamma^m \to \mathcal{L}(\mathbb{C}^N)$ for $i \in [m]$ such that $\|A_i(g)\|_{\mathrm{op}} \leq 1$ for any $g \in \Gamma^m$, A_i is independent of the i-th coordinate of g and $[A_i(g), A_j(h)] = 0$, $[A_i(g)^*, A_j(h)] = 0$ for all $i \neq j$ and $g, h \in \Gamma^m$, we have that*

$$\| \mathop{\mathbb{E}}_{(g_1,\ldots,g_m) \in \Gamma^m} f(a_1 g_1 + \cdots + a_m g_m) \prod_{i=1}^{m} A_i(g_1, \ldots, g_m)\|_{\mathrm{op}} \leq \|f\|_{U^m(\Gamma)},$$

where a_i are non-zero integers such that the characteristic of Γ exceeds all of them.

Proof:
We will prove this by induction. For $m = 1$ we have

$$\| \mathop{\mathbb{E}}_{g \in \Gamma} f(ag) A(g)\|_{\mathrm{op}} \leq | \mathop{\mathbb{E}}_{g \in \Gamma} f(ag)| = | \mathop{\mathbb{E}}_{g \in \Gamma} f(g)| = \|f\|_{U^1(\Gamma)}.$$

Here we used that A is independent of g. Assume we have proven the statement up to some integer $m \geq 1$. Then

$$\eta := \| \mathop{\mathbb{E}}_{(g_1,\dots,g_{m+1})\in\Gamma^{m+1}} f(a_1g_1+\cdots+a_{m+1}g_{m+1}) \prod_{i=1}^{m+1} A_i(g_1,\dots,g_{m+1})\|_{\mathrm{op}}$$
$$= \| \mathop{\mathbb{E}}_{(g_2,\dots,g_{m+1})\in\Gamma^{m}} A_1(g_2,\dots,g_{m+1}) \mathop{\mathbb{E}}_{g_1\in\Gamma} f(a_1g_1+\cdots+a_{m+1}g_{m+1}) \prod_{i=2}^{m+1} A_i(g_1,\dots,g_{m+1})\|_{\mathrm{op}},$$

we have done nothing, just rearranged and used the fact that A_1 is independent of g_1. Now write $F(g_2,\dots,g_{m+1}) := \mathbb{E}_{g_1\in\Gamma} f(a_1g_1+\cdots+a_{m+1}g_{m+1}) \prod_{i=2}^{m+1} A_i(g_1,\dots,g_{m+1})$ so that

$$\begin{aligned}\eta &= \| \mathop{\mathbb{E}}_{(g_2,\dots,g_{m+1})\in\Gamma^{m}} A_1(g_2,\dots,g_{m+1})F(g_2,\dots,g_{m+1})\|_{\mathrm{op}} \\ &\leq \| \mathop{\mathbb{E}}_{(g_2,\dots,g_{m+1})\in\Gamma^{m}} F(g_2,\dots,g_{m+1})F(g_2,\dots,g_{m+1})^*\|_{\mathrm{op}}^{1/2}.\end{aligned}$$

Here we used Proposition 2.3.1 and we used the fact that $\|A_1(g)\|_{\mathrm{op}} \leq 1$ for any $g \in \Gamma^{m+1}$. Recall that $\Delta_h(f)(x) = f(x)f(x+h)^*$. Then

$$\begin{aligned}\eta &\leq \| \mathop{\mathbb{E}}_{g_1,g_1',g_2,\dots,g_{m+1}} f(a_1g_1+\cdots+a_{m+1}g_{m+1})f(a_1g_1'+\cdots+a_{m+1}g_{m+1})^* \\ &\qquad \times \prod_{i=2}^{m+1} A_i(g_1,\dots,g_{m+1})A_i(g_1',\dots,g_{m+1})^*\|_{\mathrm{op}}^{1/2} \\ &\leq (\mathop{\mathbb{E}}_{g_1,h_1} \| \mathop{\mathbb{E}}_{(g_2,\dots,g_{m+1})\in\Gamma^m} \Delta_{h_1} f(a_1g_1+\cdots+a_{m+1}g_{m+1}) \\ &\qquad \times \prod_{i=2}^{m+1} A_i(g_1+h_1,\dots,g_{m+1})A_i(g_1,\dots,g_{m+1})^*\|_{\mathrm{op}})^{1/2} \\ &\leq (\mathop{\mathbb{E}}_{g_1,h_1} (\mathop{\mathbb{E}}_{h_2,\dots,h_{m+1},z\in\Gamma} \Delta_{h_{m+1}}\dots\Delta_{h_1} f(a_1g_1+z))^{1/2^m})^{1/2} \\ &\leq (\mathop{\mathbb{E}}_{g_1,h_1} (\mathop{\mathbb{E}}_{h_2,\dots,h_{m+1},z\in\Gamma} \Delta_{h_{m+1}}\dots\Delta_{h_1} f(a_1g_1+z)))^{1/2^{m+1}} \\ &= (\mathop{\mathbb{E}}_{h_1,h_2,\dots,h_{m+1},z\in\Gamma} \Delta_{h_{m+1}}\dots\Delta_{h_1} f(z))^{1/2^{m+1}} = \|f\|_{U^{m+1}(\Gamma)}.\end{aligned}$$

In the third line we used triangle inequality to get the expectation in g_1, h_1 outside the norm. In the fifth line we used the induction hypothesis to upper bound the expression in the previous line with the Gowers norm. We then use in the sixth line Jensen's inequality. □

2.4.7. Proposition (Generalized von Neumann inequality). *Let $f\colon \Gamma \to \mathbb{C}$ be a function, $\{\psi_0,\dots,\psi_t\}$ a system of affine linear forms of Cauchy-Schwarz complexity s, $A_i\colon \Gamma^m \to \mathcal{L}(\mathbb{C}^N)$ for $i \in [t]$ such that $\|A_i(g)\|_{\mathrm{op}} \leq 1$ for any $g \in \Gamma^m$*

and $[A_i(g), A_j(h)] = 0$, $[A_i(g)^*, A_j(h)] = 0$ *for all* $i \neq j$ *and* $g, h \in \Gamma^m$. *Also assume the characteristic of* Γ *is sufficiently large depending on the coefficients of the affine linear forms. Then we have the inequality*

$$\| \mathop{\mathbb{E}}_{g\in\Gamma^m} f(\psi_0(g)) \prod_{i=1}^{t} A_i(\psi_i(g))\|_{\mathrm{op}} \leq \|f\|_{U^{s+1}(\Gamma)}.$$

Proof:
The system of affine linear forms has Cauchy-Schwarz complexity s so we can partition the forms $\{\psi_1, \ldots, \psi_t\}$ into $s+1$ classes $\mathcal{A}_1, \ldots, \mathcal{A}_{s+1}$ such that ψ_0 is not an affine linear combination of any forms in any class $\mathcal{A}_i$ for any i (over $\mathbb{Q}$). So one can find a linear change of variables using Proposition 2.4.2

$$(g_1, \ldots, g_m) \mapsto (h_1, \ldots, h_m) + y_1 v_1 + \cdots + y_{s+1} v_{s+1}$$

with the property that $\psi_0(y_j v_j) = a_j y_j$ where a_j is a non-zero integer and $v_j \in \mathbb{Z}^m$, but if $\psi_i \in \mathcal{A}_j$, then $\psi_i(y_j v_j) = 0$, this is where we need the large characteristic hypothesis. Now we define

$$\tilde{A}_k(g_1, \ldots, g_m) := \prod_{j\in\mathcal{A}_k} A_j(\psi_j(g_1, \ldots, g_m)).$$

Note that $\tilde{A}_i$ is independent of its i-th coordinate and has operator norm smaller than 1. We then have

$$\begin{aligned}
&\| \mathop{\mathbb{E}}_{g\in\Gamma^m} f(\psi_0(g_1, \ldots, g_m)) \prod_{i=1}^{t} A_i(\psi_i(g_1, \ldots, g_m))\|_{\mathrm{op}} \\
&\quad = \| \mathop{\mathbb{E}}_{g\in\Gamma^m} f(\psi_0(g_1, \ldots, g_m)) \prod_{i=1}^{s+1} \tilde{A}_i(g_1, \ldots, g_m)\|_{\mathrm{op}} \\
&\quad = \| \mathop{\mathbb{E}}_{h\in\Gamma^m} \mathop{\mathbb{E}}_{y_1,\ldots,y_{s+1}\in\Gamma} f(\psi_0(h) + a_1 y_1 + \cdots + a_{s+1} y_{s+1}) \prod_{i=1}^{s+1} \tilde{A}_i(h, y_1, \ldots, y_{s+1})\|_{\mathrm{op}} \\
&\quad \leq \mathop{\mathbb{E}}_{h\in\Gamma^m} \|f\|_{U^{s+1}(\Gamma)} = \|f\|_{U^{s+1}(\Gamma)}.
\end{aligned}$$

In the third line we used the linear change of variables just described. Then we used the triangle inequality together with Lemma 2.4.6 where we need the large characteristic hypothesis. □

2.4.8. Remark. If f takes values in $\{\pm 1\}$ and A_i are ± 1-valued observables, then the inequality says that the quantum bias of such games is bounded from above by the Gowers norm of the game tensor f.

Proposition 2.4.7 is in full generality, i.e. for any abelian group the inequality holds. However, we will now restrict ourselves to the case where $\Gamma = \mathbb{F}_p^n$, where p is prime and $n \geq 1$ as it will make many things easier. We can then use the Gowers inverse theorem 1.4.4. Let us start giving the proof of Theorem 2.1.3. A t-player line game is given by a map $\tau\colon \mathbb{F}_p^n \to \{0,1\}$ which stands for the predicate together with a system of linear forms $\psi_0, \psi_i\colon (\mathbb{F}_p^n)^2 \to \mathbb{F}_p^n$ which are given by

$$\psi_0(x,y) = y \text{ and } \psi_i(x,y) = x + (i-1)y \text{ for } i = 1,\dots,t.$$

Note that the Cauchy-Schwarz complexity of this system is at most $t-1$. For the bias of the game, it is more convenient to look at $f := (-1)^\tau$. We also need the following lemma, provided kindly to us by Shravas Rao.

2.4.9. Lemma. *Let $P\colon \mathbb{F}_p^n \to \mathbb{F}_p$ be a polynomial of degree $d-1$ and $p \geq d$. Then there exist d polynomials $P_i\colon \mathbb{F}_p^n \to \mathbb{F}_p$, $i = 0,\dots,d-1$, such that*

$$P(y) = \sum_{i=0}^{d-1} P_i(x+iy).$$

Proof:
The polynomial P can be represented as

$$P(x_1,\dots,x_n) = \sum_{i=0}^{d-1} T_i(x,\dots,x), \quad x = (x_1,\dots,x_n),$$

where each $T_i\colon (\mathbb{F}_p^n)^i \to \mathbb{F}_p$ is an i-linear form. We will show that for each linear form T_i we can find $\alpha_0,\dots,\alpha_{d-1}$ such that

$$T_i(y,\dots,y) = \sum_{j=0}^{d-1} \alpha_j T_i(x+jy,\dots,x+jy) \tag{2.7}$$

and this will be enough to construct $P_0,\dots,P_{d-1}$. By linearity, we can rewrite the right hand side as follows,

$$\begin{aligned}&\sum_{j=0}^{d-1} \sum_{s\in\{0,1\}^i} \alpha_j T_i((1-s_1)x + s_1 jy,\dots,(1-s_i)x + s_i jy)\\ &\quad = \sum_{j=0}^{d-1} \sum_{s\in\{0,1\}^i} \alpha_j j^{|s|} T_i((1-s_1)x + s_1 y,\dots,(1-s_i)x + s_i y),\end{aligned}$$

where $|s|$ denotes the Hamming weight of s. Then 2.7 holds, if for $0 \leq k < i$

$$\sum_{j=0}^{d-1} \alpha_j j^k = 0 \text{ and } \sum_{j=0}^{d-1} \alpha_j j^i = 1.$$

As $d \leq p$ the $d \times d$ Vandermonde matrix associated with the sequence $1, j, \ldots, j^i$ is invertible, hence there exist unique $\alpha_0, \ldots, \alpha_{d-1}$ satisfying the above equations which concludes the proof. □

The following lemma will help us later in converting complex strategies into ± 1-strategies.

2.4.10. Lemma. *For any* $z \in \{u \in \mathbb{C} \colon |u| = 1\}$

$$z = \frac{\pi}{2} \mathop{\mathbb{E}}_{w} [\operatorname{sgn}(\Re(z\overline{w}))\, w],$$

where $w \in \{u \in \mathbb{C} \colon |u| = 1\}$ *is taken uniformly at random.*

Proof:
Write $z = e^{i\psi}$. Then

$$\begin{aligned}\mathop{\mathbb{E}}_{w}[\mathrm{sgn}(\Re(z\overline{w}))w] &= \frac{1}{2\pi}\int_0^{2\pi} \mathrm{sgn}(\Re(e^{-i(\phi-\psi)}))e^{i\phi}d\phi \\ &= \frac{1}{2\pi}\int_0^{2\pi} \mathrm{sgn}(\Re(e^{-i\chi}))e^{i\chi+i\psi}d\chi \\ &= \frac{z}{2\pi}\int_0^{2\pi} \mathrm{sgn}(\Re(e^{-i\chi}))e^{i\chi}d\chi \\ &= \frac{2z}{\pi}.\end{aligned}$$

□

Proof of Theorem 2.1.3:
By Theorem 2.4.5 and the hypothesis that the game has entangled value $\varepsilon > 0$ implies that $\|f\|_{U^t} \geq \varepsilon$. Then by the Gowers inverse theorem 1.4.4 and assumption that $p > t$, there exists a constant $\delta = \delta(t, \varepsilon, p) > 0$ and a polynomial of degree at most $t-1$ such that

$$|\mathop{\mathbb{E}}_{x\in\mathbb{F}_p^n} f(x)\omega^{P(x)}| > \delta,$$

where ω is a p-th root of unity. We now want to convert this presence of structure into a classical strategy. First by Lemma 2.4.9 we can find t polynomials P_i for $i = 1, \ldots, t$ such that

$$P(y) = \sum_{i=1}^{t} P_i(x + (i-1)y).$$

This implies

$$|\mathop{\mathbb{E}}_{x,y\in\mathbb{F}_p^n} f(\psi_0(x,y))\omega^{P(\psi_0(x,y))}| = |\mathop{\mathbb{E}}_{x,y\in\mathbb{F}_p^n} f(\psi_0(x,y))\prod_{i=1}^{t}\omega^{P_i(\psi_i(x,y))}| > \delta.$$

The polynomials are not classical strategies yet, we can turn them into ± 1-strategies using Lemma 2.4.10 at a loss of a factor $2^t/\pi^t$. □

2.4.3 Parallel repetition

Let $f\colon \Gamma \to \{\pm 1\}$ be a function, representing the predicate. We want to consider k-fold XOR parallel repetition. The predicate for this is $f^k\colon \Gamma^k \to \{1, -1\}$ defined by

$$f^k(g_1, \ldots, g_k) := \prod_{i=1}^{k} f(g_i).$$

2.4.11. Lemma. *We have that*

$$\|f^k\|_{U^{s+1}(\Gamma^k)} = \|f\|^k_{U^{s+1}(\Gamma)}.$$

Proof:
This follows immediately from the definition of the Gowers norm. □

Let $\{\psi_0, \ldots, \psi_t\}$ be linear forms $\Gamma^m \to \Gamma$ which together with f define the game G. The linear forms corresponding with k-fold XOR parallel repetition are denoted by $\{\psi_0^k, \ldots, \psi_t^k\}$ which are maps $(\Gamma^m)^k \to \Gamma^k$ and are given by

$$\psi_i^k(g^1, \ldots, g^k) := (\psi_i(g^1), \ldots, \psi_i(g^k)), \text{ where } g^i \in \Gamma^m.$$

Note that if $\{\psi_0, \ldots, \psi_t\}$ has Cauchy-Schwarz complexity at most s, then the Cauchy-Schwarz complexity of $\{\psi_0^k, \ldots, \psi_t^k\}$ is also at most s. Denote by $G^{\oplus k}$ the k-fold XOR parallel repetition, then we have as an immediate consequence of Theorem 2.4.5 together with Lemma 2.4.11 the following upper bound

$$\beta^*(G^{\oplus k}) \leq \|f\|^k_{U^{s+1}(\Gamma)}.$$

If G is an XOR game, denote by G^k the k-fold parallel repetition. If S is a strategy (classical or quantum) for a game G, denote by $\omega(G, S)$ the winning probability using strategy S. Also denote by $\varepsilon(G, S) := 2\omega(G, S) - 1$ the bias of this strategy. To prove Lemma 2.1.4, we use the following lemma, which is a straightforward generalization of the 2-player version in [CSUU08] (Lemma 8 in that paper) to any number of players.

2.4.12. Lemma. *Let G be an XOR game assume. Let S be any strategy for G^k. For each $M \subset [k]$, we denote by S_M the following strategy for the XOR parallel repetition $\oplus_{i\in M} G$: (1) Run strategy S, yielding answers $a_i^1, \ldots, a_i^k$ for player $i = 1, \ldots, t$. (2) Player i outputs $\sum_{j\in M} a_i^j \mod 2$. We then have*

$$\omega(G^k, S) = \frac{1}{2^k} \sum_{M\subset[k]} \varepsilon(\oplus_{i\in M} G, S_M).$$

Proof of Lemma 2.1.4.:
Let S be the quantum strategy that achieves the maximum winning probability of the game G^k. We then use Lemma 2.4.12,

$$\begin{aligned}
\omega^*(G^k) = \omega(G^k, S) &= \frac{1}{2^k} \sum_{M\subset[k]} \varepsilon(\oplus_{i\in M} G, S_M) \\
&\leq \frac{1}{2^k} \sum_{M\subset[k]} \beta^*(G^{\oplus|M|}) = \frac{1}{2^k} \sum_{l=0}^{k} \beta^*(G^{\oplus l}) \binom{k}{l} \\
&\leq \frac{1}{2^k} \sum_{l=0}^{k} \binom{k}{l} \|f\|_{U^{s+1}(\Gamma)} = \left(\frac{1 + \|f\|_{U^{s+1}(\Gamma)}}{2}\right)^k.
\end{aligned}$$

□

Chapter 3

Quasirandom quantum channels

This chapter is based on the paper [BBLM20] which is joint work with Tom Bannink, Jop Briët and Hans Maassen. Recall the notation and basic concepts on quasirandom graphs in Section 1.3.

3.1 Introduction

In a seminal work [CGW89], Chung, Graham and Wilson — building on work of Thomason [Tho87a, Tho87b] — proved that several seemingly distinct notions of quasirandomness for graphs are equivalent. In particular, they identified seven properties found in random graphs with high probability, that always coexist simultaneously in any large dense graph. Two of these properties are spectral expansion and uniformity. A question of Chung and Graham [CG02] on the equivalence of these two properties in *sparse* graphs resulted in a line of research culminating in recent work of Conlon and Zhao [CZ17], which introduced a surprising new item to the armory of combinatorics: the famous Grothendieck inequality [Gro53]. In this chapter, we draw a parallel line in the context of quantum information theory, where quantum channels take the place of graphs. In addition, we give a streamlined proof of the main result of [CZ17] and show that the use of Grothendieck's inequality yields an optimal constant. Similarly, we show that the non-commutative Grothendieck inequality gives an optimal constant in the quantum setting.

Expander Mixing Lemma. For a regular graph G, denote by $\varepsilon(G)$ its uniformity parameter and $\lambda(G)$ its expansion parameter (see Section 1.3 for definitions). A basic result known as the Expander Mixing Lemma [HLW06] shows that for any regular graph G we have $\varepsilon(G) \leq \lambda(G)$, which is to say that spectral expansion implies uniformity. A sequence G_n of d_n-regular graphs is called *dense* if $d_n \geq \Omega(n)$, and *sparse* if $d_n/n \longrightarrow 0$. It was shown in [CGW89] that in the dense case, a converse to the Expander Mixing Lemma $\varepsilon(G_n) \leq o(1) \Rightarrow \lambda(G_n) \leq o(1)$

also holds. In contrast, Krivelevich and Sudakov [KS06] showed that this is false for sparse graphs, thereby answering the question posed in [CG02]. Their counterexample is not regular, however (and a later one from [BN04] is not connected). But in [CZ17] it was shown that even regular sparse graphs (where $d_n \leq o(n)$) can simultaneously satisfy $\varepsilon(G_n) \leq o(1)$ and $\lambda(G_n) \geq \Omega(1)$. Surprisingly, Kohayakawa, Rödl, and Schacht [KRS16] showed that Cayley graphs over abelian groups, including sparse ones, do again admit such a converse. Cayley graphs are an important class of regular graphs that include for instance the famous Ramanujan graphs of Margulis [Mar88] and Lubotzky, Phillips and Sarnak [LPS88]. Conlon and Zhao [CZ17] generalized this to all Cayley graphs and showed that this implies the same for all vertex-transitive graphs in general, for which they showed that $\lambda(G) \leq 4K_G\varepsilon(G)$, where $1.6769\ldots \leq K_G < 1.7822\ldots$ is the famous *Grothendieck constant*, whose exact value is currently unknown; the bounds shown here are the best known and were shown by Davie and Reeds (independently) in [Dav84, Ree91] and Braverman et al. in [BMMN13], respectively. Spectral expansion and uniformity are thus equivalent notions of quasirandomness for dense graphs and vertex-transitive graphs.

Quasirandomness in quantum information theory. A transition matrix, such as the normalized adjacency matrix of a graph, maps probability vectors[1] to probability vectors. A natural non-commutative generalization of a transition matrix is a quantum channel, a completely positive trace preserving linear map $\Phi : M_n(\mathbb{C}) \to M_n(\mathbb{C})$; see Section 1.1 for formal definitions. Quantum channels are the most general operations on quantum systems that are physically realizable. They encapsulate the "classical" transition matrices by restricting them to diagonal matrices whose diagonals form probability vectors; we discuss this in more detail in Section 3.3. Recall that in quantum information theory, general linear maps from $M_n(\mathbb{C})$ to itself are referred to as superoperators. Since superoperators are in one-to-one correspondence with bilinear forms on $M_n(\mathbb{C}) \times M_n(\mathbb{C})$, they also appear in the context of (generalizations of) Bell inequalities from physics in the form of quantum XOR games [RV15, CJPPG15], as well as in combinatorial optimization [NRV14]. The graph-theoretic concepts mentioned above have natural analogues for superoperators, which we discuss next.

In independent work, Hastings [Has07] and Ben-Aroya, Schwartz and Ta-Schma [BST10] introduced *quantum expanders* as a special class of quantum channels defined analogously to spectral expanders. For a superoperator Φ, the expansion parameter is given by

$$\lambda(\Phi) = \|\Phi - \Pi\|_{S_2 \to S_2} = \sup\big\{\|(\Phi - \Pi)(X)\|_{S_2} : \|X\|_{S_2} \leq 1\big\}, \tag{3.1}$$

where $\Pi : X \mapsto \frac{1}{n}\mathrm{Tr}(X)\mathrm{Id}$ is the projection onto the identity, $\|X\|_{S_2} = \sqrt{\langle X, X\rangle}$

[1] We use the convention of writing probability vectors as *column* vectors intead of row vectors.

is the Frobenius (or Schatten-2) norm and $\langle X, Y\rangle = \frac{1}{n}\mathrm{Tr}(Y^*X)$ is the normalized trace inner product. A quantum channel is an expander if $\lambda(\Phi)$ is much smaller than 1. Also quantum expanders found many applications, one of which is again randomness reduction, where randomness takes on the form of random unitary matrices. Since a k-qubit unitary requires 4^k real parameters, sampling one from the uniform distribution (Haar probability measure) is very expensive. A 1-design is a fixed collection of unitaries $U_1, \ldots, U_m$ such that the superoperator $\Phi : X \mapsto \frac{1}{m}\sum_{i=1}^m U_i X U_i^*$ exactly effects the projection Π, thus mimicking in a finite way the Haar measure on $U(n)$. Quantum expanders can be used to construct *approximate* 1-designs, meaning that $\Phi(X)$ and $\Pi(X)$ are close in trace distance[2] instead of precisely equal. Another application is in cryptography where Ambainis and Smith [AS04b] used quantum expanders to construct short quantum one-time pads. It was shown in [Has07] that truly random quantum channels (given by independent Haar-uniform U_i as described above) are quantum expanders with high probability, supporting the idea that this is a notion of quasirandomness.

In this work we introduce a natural notion of uniformity for superoperators, informally given by how well they mimic the action of Π on projectors on subspaces, which may be thought of as generalizations of vertex subsets in graphs. This is similar to Hasting's notion of edge expansion for quantum channels [Has07]. In particular, we say that Φ is ε-uniform if for any two subspaces $V, W \subseteq \mathbb{C}^n$ with associated projections P_V, P_W, it holds that

$$|\langle P_V, (\Phi - \Pi)(P_W)\rangle| \leq \varepsilon. \tag{3.2}$$

Let $\varepsilon(\Phi)$ denote the smallest ε for which this holds. As we show in Section 3.3.3, the parameters $\lambda(\Phi)$ and $\varepsilon(\Phi)$ reduce to their graphical analogs under a suitable embedding of graphs into quantum channels.

Finally, also symmetry, which in the graph-theoretic context takes the form of vertex transitivity, is an important property of quantum channels. In particular, *irreducibly covariant* quantum channels, which turn out to generalize vertex-transitive graphs (see Section 3.3), play an important role in questions about the capacity of quantum channels as noisy transmitters of quantum information [Hol06]. A now famous result of Hastings [Has09] shows that the minimum output capacity in general does not have the intuitively natural property of being sub-additive under tensor products. However, it was shown earlier by Holevo [Hol02], that the capacity is additive for the subclass of irreducibly covariant quantum channels.

Summary of our results. In this work we make a first step in the study of the equivalence of quasirandom properties for quantum channels, or superoperators in

[2]The trace distance is the distance induced by the Schatten-1 norm, defined in Section 1.1.

general, and show optimality in the case of vertex-transitive graphs and covariant quantum channels.

- (Section 3.3.2) Our main result shows that under irreducible covariance, expansion and uniformity are equivalent for superoperators. In particular, while a simple analogue of the classical Expander Mixing Lemma implies that $\varepsilon(\Phi) \leq \lambda(\Phi)$ in general, we show using a non-commutative version of Grothendieck's inequality due to Haagerup [Haa85], that for this class of superoperators, also $\lambda(\Phi) \leq 2\pi^2\varepsilon(\Phi)$ always holds. This implies the same result for vertex-transitive graphs with $\mathbb{C}$-weighted edges, essentially proved in [CZ17] with the factor 2 replaced by the *complex* Grothendieck constant $1.3380\ldots \leq K_G^{\mathbb{C}} \leq 1.4049\ldots$.

- (Section 3.3.3) We show that a construction of sparse regular graphs from [CZ17] can be embedded to give a sequence of quantum channels Φ_n that are not irreducibly covariant and for which it holds that $\varepsilon(\Phi_n) \leq o(1)$ and $\lambda(\Phi_n) \geq \Omega(1)$.

- (Section 3.3.4) We show that for *randomizing* channels, a notion introduced in [Aub09], the two notions of quasirandomness are also equivalent. This can be interpreted as a generalization of the same statement for dense graphs proved in [CGW89].

- (Section 3.4.1) We show that the result of [CZ17] cannot be improved in the sense that the factors $4K_G$ and $\pi^2 K_G^{\mathbb{C}}$ are optimal in the case of vertex-transitive graphs with $\mathbb{R}$-weighted and $\mathbb{C}$-weighted edges, respectively.

- (Section 3.4.2) Our work leaves open whether the factor $2\pi^2$ in our main result is optimal. However, our proof consists of two steps, the first of which gives a factor 2 and the second a factor π^2, and we show these steps are individually optimal. We prove that the first step is optimal by showing that an example of Haagerup and Ito [HI95] for the non-commutative Grothendieck inequality is irreducibly covariant, which uses some representation theory of $\mathrm{SO}(n)$. The optimality of the second step follows directly from a result of [CZ17].

3.2 Preliminaries

For a compact set S, write $C(S)$ for the set of continuous functions from S to $\mathbb{C}$. For a compact group Γ, write $\mathbb{E}_{g\in\Gamma}$ for the the integral with respect to the (unique) Haar probability measure on Γ.

Recall that $M_n(\mathbb{C})$ is the set of complex $n \times n$ matrices and we write $U(n)$ for the set of unitary matrices $\{X \in M_n(\mathbb{C}) : X^*X = \mathrm{Id}\}$. Here, all maps of

the form $\Phi : M_n(\mathbb{C}) \to M_n(\mathbb{C})$ are linear and referred to as superoperators. A superoperator Φ is *unital* if $\Phi(\mathrm{Id}) = \mathrm{Id}$.

We normalize inner products so that for $x, y \in \mathbb{C}^n$ we define $\langle y, x\rangle = \mathbb{E}_{i\in[n]} \overline{y_i} x_i$ and for matrices $X, Y \in M_n(\mathbb{C})$ we have $\langle Y, X\rangle = \frac{1}{n}\mathrm{Tr}[Y^* X]$.

3.2.1. Proposition. *Let $p \geq 1$ and let $X \in M_n(\mathbb{C})$. Then*

$$\|X\|_{S_p} \geq \|(X_{11}, \ldots, X_{nn})\|_{L_p}.$$

Proof:
For a vector $x \in \mathbb{C}^n$, denote by $\mathrm{Diag}(x)$ the $n \times n$ matrix with x on the diagonal and for a matrix X denote by $\mathrm{diag}(X)$ the matrix where we set the off-diagonal elements to 0. A small computation shows that

$$\mathop{\mathbb{E}}_{s\in\{\pm 1\}^n} \mathrm{Diag}(s)\, X\, \mathrm{Diag}(s) = \mathrm{diag}(X).$$

Since the Schatten-p norms are invariant under conjugation with a unitary matrix, applying the above with the triangle inequality gives

$$\|(X_{11}, \ldots, X_{nn})\|_{L_p} = \|\mathrm{diag}(X)\|_{S_p} \leq \mathop{\mathbb{E}}_{s\in\{\pm 1\}^n} \|\mathrm{Diag}(s)\, X\, \mathrm{Diag}(s)\|_{S_p} = \|X\|_{S_p}.$$

□

For $q \in [1, \infty]$, define $q' \in [1, \infty]$ to be its dual given by $\frac{1}{q} + \frac{1}{q'} = 1$. For $p, q \in [1, \infty]$, a matrix $A \in M_n(\mathbb{C})$ and a superoperator $\Phi : M_n(\mathbb{C}) \to M_n(\mathbb{C})$, define

$$\|A\|_{L_p \to L_q} = \sup\{|\langle y, Ax\rangle| : \|x\|_{L_p} \leq 1,\ \|y\|_{L_{q'}} \leq 1\}$$
$$\|\Phi\|_{S_p \to S_q} = \sup\{|\langle Y, \Phi(X)\rangle| : \|X\|_{S_p} \leq 1,\ \|Y\|_{S_{q'}} \leq 1\}.$$

If G is a d-regular graph on n vertices with normalized adjacency matrix A, then $\lambda(G) = \|A - \frac{1}{n}J\|_{L_2 \to L_2}$, where J is the all-ones matrix. Also recall from (3.1) that for a superoperator Φ the expansion parameter is $\lambda(\Phi) = \|\Phi - \Pi\|_{S_2 \to S_2}$.

Also define the *cut norms* by

$$\|A\|_{\mathrm{cut}} = \max\{|\langle y, Ax\rangle| : x, y \in \{0, 1\}^n\}$$
$$\|\Phi\|_{\mathrm{cut}} = \sup\{|\langle Y, \Phi(X)\rangle| : X, Y \text{ projectors}\}.$$

It is then not hard to see that if G is a d-regular graph on n vertices with normalized adjacency matrix A, then $\varepsilon(G) = \|A - \frac{1}{n}J\|_{\mathrm{cut}}$. Similarly, we have $\varepsilon(\Phi) = \|\Phi - \Pi\|_{\mathrm{cut}}$.

We have the following relation between these norms, the proof of which is a simple generalization of the same result from [CZ17] for matrices.

3.2.2. Lemma. *For any superoperator Φ, we have $\|\Phi\|_{\text{cut}} \leq \|\Phi\|_{S_\infty \to S_1} \leq \pi^2 \|\Phi\|_{\text{cut}}$ and π^2 is the best possible constant.*

Proof:
First note that the cut norm as defined above can also be written as

$$\|\Phi\|_{\text{cut}} = \sup\{|\langle Y, \Phi(X)\rangle| : X, Y \succeq 0\,,\ \|X\|_{S_\infty}, \|Y\|_{S_\infty} \leq 1\}, \tag{3.3}$$

because the set $\{X : X \succeq 0, \|X\|_{S_\infty} \leq 1\}$ is the convex hull of the set of projectors. Hence, by linearity the supremum in (3.3) will always be attained by projectors.

The first inequality of the lemma follows by dropping the positive semidefinite constraint. For the second inequality, let z be a complex number of norm 1, and w a uniform random complex number of norm 1. Then

$$z = \pi\, \mathbb{E}_w[w\, 1_{\{\Re(z\bar{w}) \geq 0\}}\,].$$

This follows from Lemma 2.4.10. We have that

$$\|\Phi\|_{S_\infty \to S_1} = \sup\{|\langle Y, \Phi(X)\rangle| : \|X\|_{S_\infty}, \|Y\|_{S_\infty} \leq 1\}.$$

The set of matrices X such that $\|X\|_{S_\infty} \leq 1$ is the convex hull of the set of unitary matrices, so by linearity we can assume that the supremum in $\|\Phi\|_{S_\infty \to S_1}$ is obtained by unitary X, Y. Unitary matrices are diagonalizable, so write $X = UAU^*$ and $Y = VBV^*$ with U, V unitary and A, B diagonal. Let $u, w \in \mathbb{C}$, $|u| = |w| = 1$ be uniform random complex numbers and define diagonal matrices A', B' as $A'_{ii}(w) = 1_{\{\Re(A_{ii}\bar{w}) \geq 0\}}$ and $B'_{ii}(u) = 1_{\{\Re(B_{ii}\bar{u}) \geq 0\}}$. By the above we have $A = \pi\, \mathbb{E}_w[wA'(w)]$ and similar for B, so we have $X = \pi\, \mathbb{E}_w[wUA'(w)U^*]$ and $Y = \pi\, \mathbb{E}_u[uVB'(u)V^*]$. Now, $UA'(w)U^*$ and $VB'(u)V^*$ are projections for all values of w and u, as required in the definition of the cut norm. Therefore

$$\begin{aligned}
\|\Phi\|_{S_\infty \to S_1} = |\langle Y, \Phi(X)\rangle| &= \pi^2 |\mathbb{E}_{u,w} \bar{u} w \langle VB'(u)V^*, \Phi(UA'(w)U^*)\rangle| \\
&\leq \pi^2 \mathbb{E}_{u,w} |\langle VB'(u)V^*, \Phi(UA'(w)U^*)\rangle| \\
&\leq \pi^2 \mathbb{E}_{u,w} \|\Phi\|_{\text{cut}} \\
&= \pi^2 \|\Phi\|_{\text{cut}},
\end{aligned}$$

completing the first part of the proof. Conlon and Zhao show that π^2 is the best possible constant in the commutative case, using the matrix $A \in M_n(\mathbb{C})$ given by $A_{st} = e^{2\pi i(s-t)/n}$. This matrix satisfies $\|A\|_{L_\infty \to L_1} = n$ and one can show $\|A\|_{\text{cut}} = (\pi^{-2} + o(1))n$. By Theorem 3.3.7 in Section 3.3.3, their example can be embedded into a superoperator with the same norms so π^2 is also the best possible constant here. □

Define the *Grothendieck norm* of a matrix $A \in M_n(\mathbb{C})$ by

$$\|A\|_G := \sup\Big\{\Big|\frac{1}{n}\sum_{i,j=1}^{n} A_{ij}\langle x_i, y_j\rangle\Big| : d \in \mathbb{N},\ x_i, y_j \in \mathbb{C}^d,\ \|x_i\|_{L_2} \le 1,\ \|y_j\|_{L_2} \le 1\Big\}.$$

Then, the *complex Grothendieck constant* is given by

$$K_G^{\mathbb{C}} := \sup\Big\{\frac{\|A\|_G}{\|A\|_{L_\infty \to L_1}} : n \in \mathbb{N},\ A \in M_n(\mathbb{C})\Big\}.$$

The current best upper and lower bounds on $K_G^{\mathbb{C}}$ are 1.4049 [Haa87] and 1.338 [Dav84], respectively. The real version of the Grothendieck constant, denoted by K_G and mentioned in the introduction, is obtained by replacing the underlying field in the above quantities by the reals.

Some basic group theory. Given a graph $G = (V, E)$, a permutation $\pi : V \to V$ is an *automorphism* of G if for all $u, v \in V$, we have $\{\pi(u), \pi(v)\} \in E \Leftrightarrow \{u, v\} \in E$. The automorphisms of G form a group under composition, which we call $\mathrm{Aut}(G)$. Then, G is said to be *vertex-transitive* if for every $u, v \in V$, there is a $\pi \in \mathrm{Aut}(G)$ such that $\pi(u) = v$. For superoperators, we have the following analogous definitions. A unitary representation of a group Γ on $\mathbb{C}^n$ is a homomorphism from Γ to $U(n)$ and it is irreducible if the only subspaces of $\mathbb{C}^n$ that are left invariant by the group action are the zero-dimensional subspace and $\mathbb{C}^n$ itself.

3.2.3. Definition (Irreducible covariance). A superoperator $\Phi : M_n(\mathbb{C}) \to M_n(\mathbb{C})$ is *irreducibly covariant* if there exist a compact group Γ and continuous irreducible unitary representations $U, V \colon \Gamma \to U(n)$ such that for all $g \in \Gamma$ and $X \in M_n(\mathbb{C})$, we have

$$\Phi(U(g)XU^*(g)) = V(g)\Phi(X)V^*(g).$$

3.3 Converse expander mixing lemmas

In this section, we prove the "converse expander mixing lemmas" announced in the first and third bullet in the introduction as well as the examples announced in the second bullet. As a warm-up, we start with a proof of the commutative case due to Conlon and Zhao, which we reprove in a slightly different manner analogous to how we will prove the non-commutative case.

3.3.1 Commutative case

In the following, let S be a compact set and Γ be a compact group acting continuously and transitively on S. The Haar probability measure on Γ induces a

measure on S (by pullback) according to which the L_p-norm (for $p \in [1,\infty)$) and inner product of $f, g \in C(S)$ are given by

$$\|f\|_{L_p} = \Big(\mathop{\mathbb{E}}_{\pi\in\Gamma} \big|f\big(\pi(s_0)\big)\big|^p \Big)^{\frac{1}{p}} \qquad \text{and} \qquad \langle f, g\rangle = \mathop{\mathbb{E}}_{\pi\in\Gamma} \overline{f\big(\pi(s_0)\big)} g\big(\pi(s_0)\big), \quad (3.4)$$

where (by transitivity) s_0 can be taken to be some arbitrary but fixed element of S. We lift the action of Γ on S to an action on $C(S)$ by precomposition, that is, for any function $f \in C(S)$ and element $\pi \in \Gamma$, define the function f^π by $f^\pi(s) := f(\pi(s))$. Furthermore, for a linear map $A\colon C(S) \to C(S)$ define A^π by $A^\pi f := (Af^\pi)^{\pi^{-1}}$ and say that A is transitive covariant with respect to Γ if for any $\pi \in \Gamma$ we have $A^\pi = A$.[3] We sometimes omit the group and simply say A is *transitive covariant* if such a group Γ exists.

In [CZ17], the following result is proved (over the real numbers) for the case $S = [n]$, in which case transitive covariant linear maps A are simply $n \times n$ matrices which commute with the permutation matrices of a transitive subgroup Γ of S_n. However, their proof easily implies the more general version below.

3.3.1. Theorem (Conlon–Zhao). *Let S be as above and let $A : C(S) \to C(S)$ be a linear map that is transitive covariant with respect to Γ. Then,*

$$\|A\|_{L_2\to L_2} \le K_G^{\mathbb{C}} \|A\|_{L_\infty \to L_1} .$$

Here we give a somewhat more streamlined proof of this result based on a well-known factorization version of Grothendieck's inequality [Gro53] (see also [Pis12]), which will serve as a stepping stone to the proof of the non-commutative case.[4] In our setting the inequality asserts the following

3.3.2. Theorem (Commutative Grothendieck inequality (factorization)). *Let S be as above and let $A : C(S) \to C(S)$ be a linear map. Then, there exist probability measures λ, ν on S such that for all $f, g \in C(S)$, we have*

$$|\langle g, Af\rangle| \le K_G^{\mathbb{C}} \|A\|_{L_\infty\to L_1} \left(\int_S |f(s)|^2 \, d\lambda(s) \right)^{1/2} \left(\int_S |g(s)|^2 d\nu(s) \right)^{1/2} .$$

Proof of Theorem 3.3.1:
It follows from the triangle inequality and transitivity that

$$|\langle g, Af\rangle| \le \mathop{\mathbb{E}}_{\pi\in\Gamma} |\langle g, A^\pi f\rangle| = \mathop{\mathbb{E}}_{\pi\in\Gamma} |\langle g^\pi, Af^\pi\rangle|.$$

[3] In general one says A is *covariant* with respect to Γ, but we say *transitive* to emphasize that we require Γ to act transitively on S.

[4] The main difference is that in [CZ17], the result is first proved for weighted Cayley graphs, after which it is shown that this implies the result for transitive covariant matrices.

By Theorem 3.3.2 and the AM-GM inequality there are probability measures λ, ν on S such that the above right-hand side is at most

$$\frac{K_G^{\mathbb{C}}\|A\|_{L_\infty\to L_1}}{2}\mathop{\mathbb{E}}_{\pi\in\Gamma}\left(\int_S |f^\pi(s)|^2 d\lambda(s) + \int_S |g^\pi(s)|^2 d\nu(s)\right)$$
$$= \frac{K_G^{\mathbb{C}}\|A\|_{L_\infty\to L_1}}{2}(\|f\|_{L_2}^2 + \|g\|_{L_2}^2),$$

where we switched the order of the integrals (using Tonelli's theorem) and the expression (3.4) for the L_2 norm. Thus, when $\|f\|_{L_2} = \|g\|_{L_2} = 1$ we have shown that $\|A\|_{L_2\to L_2} \le K_G^{\mathbb{C}}\|A\|_{L_\infty\to L_1}$. □

3.3.2 Non-commutative case

Our main technical result is as follows.

3.3.3. Theorem. *Let $\Phi : M_n(\mathbb{C}) \to M_n(\mathbb{C})$ be an irreducibly covariant superoperator. Then, $\|\Phi\|_{S_\infty\to S_1} \le \|\Phi\|_{S_2\to S_2} \le 2\|\Phi\|_{S_\infty\to S_1}$.*

Since the supremum in $\|\Phi\|_{S_\infty\to S_1}$ is taken over X, Y with S_∞-norm equal to 1, the first inequality of the theorem follows from the fact that $\|X\|_{S_2} \le \|X\|_{S_\infty}$. As projectors have Schatten-∞ norm 1, the first inequality also easily implies the analogue of the Expander Mixing Lemma, that is, $\varepsilon(\Phi) \le \lambda(\Phi)$, where $\lambda(\Phi)$ and $\varepsilon(\Phi)$ are as in (3.1) and (3.2), respectively; note that when Φ is irreducibly covariant, so is $\Phi - \Pi$. The second inequality is proved at the end of this section and in Section 3.4.2 we show that the factor 2 in the theorem is optimal. With Lemma 3.2.2, which relates the uniformity parameter $\varepsilon(\Phi)$ to $\|\Phi - \Pi\|_{S_\infty\to S_1}$, Theorem 3.3.3 then immediately gives the following result stated in the introduction.

3.3.4. Corollary (Converse Quantum Expander Mixing Lemma). *Let Φ be an irreducibly covariant superoperator. Then, $\lambda(\Phi) \le 2\pi^2\varepsilon(\Phi)$.*

In this non-commutative setting we use the following analog of Theorem 3.3.2 (a factorization version of the non-commutative Grothendieck inequality), proved by Haagerup in [Haa85]; see also [Pis12]. A density matrix is a positive semidefinite matrix with trace equal to 1.

3.3.5. Theorem (Haagerup). *Let $\Phi\colon M_n(\mathbb{C}) \to M_n(\mathbb{C})$ be a superoperator. Then, there exist density matrices $\rho_1, \rho_2, \sigma_1, \sigma_2$ such that for any $X, Y \in M_n(\mathbb{C})$, we have*

$$|\langle Y, \Phi(X)\rangle| \le \|\Phi\|_{S_\infty\to S_1}\left(\mathrm{Tr}[\rho_1 X^*X] + \mathrm{Tr}[\rho_2 XX^*]\right)^{1/2}\left(\mathrm{Tr}[\sigma_1 Y^*Y] + \mathrm{Tr}[\sigma_2 YY^*]\right)^{1/2}. \tag{3.5}$$

We also use the following lemma.

3.3.6. Lemma. *Let Γ be a compact group. Then, a representation $U : \Gamma \to U(n)$ is irreducible if and only if for any $X \in M_n(\mathbb{C})$, we have*

$$\mathop{\mathbb{E}}_{g\in\Gamma} U(g)XU(g)^* = \mathrm{Tr}(X)\frac{1}{n}\mathrm{Id}.$$

Proof:
By Schur's lemma, if U is an irreducible representation, then for $T \in M_n(\mathbb{C})$

$$\Big[\forall g \in \Gamma \quad U(g)TU(g)^* = T\Big] \iff \Big[\exists \lambda \in \mathbb{C} \quad T = \lambda\,\mathrm{Id}\Big].$$

Let $T_X = \mathbb{E}_{g\in\Gamma} U(g)XU(g)^*$, by the group structure we have $U(g)T_XU(g)^* = T_X$ for all $g \in \Gamma$. Therefore, if U is irreducible then $T_X = \lambda_X\,\mathrm{Id}$. By taking the trace, it follows that $\lambda_X = \mathrm{Tr}(X)/n$. In the other direction, if U is reducible then there exists a projector P onto an irreducible subspace that is left invariant, i.e. $U(g)PU(g)^* = P$ for all $g \in \Gamma$, so $T_P \neq \lambda\mathrm{Id}$. □

Proof of Theorem 3.3.3:
Denote by Γ and $U, V\colon \Gamma \to U(n)$ the group and irreducible representations such that Φ is irreducibly covariant with respect to Γ (see Definition 3.2.3). For any $X, Y \in M_n(\mathbb{C})$ write $X_g = U(g)XU^*(g)$ and $Y_g = V(g)YV^*(g)$, then we have

$$|\langle Y, \Phi(X)\rangle| = \mathop{\mathbb{E}}_{g\in\Gamma} |\langle Y_g, \Phi(X_g)\rangle|.$$

By Theorem 3.3.5 and the AM-GM inequality, there exist density matrices $\rho_1, \rho_2, \sigma_1, \sigma_2$ such that the right-hand side is bounded from above by

$$\frac{1}{2}\|\Phi\|_{S_\infty\to S_1} \mathop{\mathbb{E}}_{g\in\Gamma} \left(\mathrm{Tr}[\rho_1 X_g^* X_g] + \mathrm{Tr}[\rho_2 X_g X_g^*] + \mathrm{Tr}[\sigma_1 Y_g^* Y_g] + \mathrm{Tr}[\sigma_2 Y_g Y_g^*]\right).$$

By Lemma 3.3.6 we have

$$\mathop{\mathbb{E}}_{g\in\Gamma} X_g^* X_g = \mathop{\mathbb{E}}_{g\in\Gamma} U(g)X^*XU^*(g) = \frac{1}{n}\mathrm{Tr}[X^*X]\mathrm{Id} = \|X\|_{S_2}^2\mathrm{Id}.$$

Let ρ be a density matrix, then $\mathbb{E}_{g\in\Gamma}\,\mathrm{Tr}[\rho X_g^* X_g] = \|X\|_{S_2}^2$. The same holds for $\mathbb{E}_{g\in\Gamma}\,\mathrm{Tr}[\rho X_g X_g^*]$ but with U, and for Y with V, so we see that the above quantity is equal to

$$\|\Phi\|_{S_\infty\to S_1}\left(\|X\|_{S_2}^2 + \|Y\|_{S_2}^2\right).$$

If $\|X\|_{S_2} = \|Y\|_{S_2} = 1$ we obtain $\|\Phi\|_{S_2\to S_2} \leq 2\|\Phi\|_{S_\infty\to S_1}$. □

3.3.3 Embedding graphs into quantum channels

In this subsection, we elucidate the claim that quantum channels generalize graphs and prove the result stated in the second bullet in the introduction, namely that there are non-irreducible quantum channels for which a converse expander mixing lemma does not hold.

We consider the following embeddings. For $A \in M_n(\mathbb{C})$, define the superoperator $\Phi_A : M_n(\mathbb{C}) \to M_n(\mathbb{C})$ as

$$\Phi_A(X) = \sum_{i,j} A_{ij} X_{jj} E_{ii}, \tag{3.6}$$

where E_{ij} is the matrix with a single 1 at position (i, j). When A is a transition matrix, i.e., its column sums are 1, then it is not hard to see that Φ_A is completely positive and trace preserving and that $\Phi_{\frac{1}{n}J} = \Pi$. Several other ways exist to create quantum expanders from expander graphs, see for example [HH09] and [Har08], but as we show below, our embedding given above carries over all relevant properties of the graph we consider here.

Conlon and Zhao [CZ17] give an infinite sequence of d-regular graphs G_n that are $o(1)$-uniform but for which $\lambda(G_n) \geq 1/2$. Combined with the following proposition, this immediately gives the result stated in the second bullet in the introduction.

3.3.7. Proposition. *Let $A \in M_n(\mathbb{C})$ and $p, q \in [1, \infty]$. Then, for Φ_A as in* (3.6)*, we have*

$$\|\Phi_A - \Pi\|_{S_p \to S_q} = \|A - \frac{1}{n}J\|_{L_p \to L_q} \quad \textit{and} \quad \|\Phi_A - \Pi\|_{\text{cut}} = \|A - \frac{1}{n}J\|_{\text{cut}}.$$

Proof:
Let $B = A - \frac{1}{n}J$, then $\Phi_A - \Pi = \Phi_B$. By compactness and definition of $\|\cdot\|_{S_p \to S_q}$ we can assume there is an $X \in M_n(\mathbb{C})$ such that $\|\Phi_B\|_{S_p \to S_q} = \|\Phi_B(X)\|_{S_q} / \|X\|_{S_p}$. Write $X = \text{diag}(x) + X_{\text{other}}$ where $x \in \mathbb{C}^n$ is the diagonal of X, and X_{other} are the off-diagonal entries. Note that by definition of Φ_B we have $\Phi_B(X) = \Phi_B(\text{diag}(x)) = \text{diag}(Bx)$. By definition of Schatten norms, $\|\text{diag}(x)\|_{S_p} = \|x\|_{L_p}$ and by Proposition 3.2.1 we have $\|X\|_{S_p} \geq \|x\|_{L_p}$. We have

$$\|B\|_{L_p \to L_q} \geq \frac{\|Bx\|_{L_q}}{\|x\|_{L_p}} \geq \frac{\|\text{diag}(Bx)\|_{S_q}}{\|X\|_{S_p}} = \frac{\|\Phi_B(X)\|_{S_q}}{\|X\|_{S_p}} = \|\Phi_B\|_{S_p \to S_q}$$

Now let $y \in \mathbb{C}^n$ be such that $\|B\|_{L_p \to L_q} = \|By\|_{L_q} / \|y\|_{L_p}$. Then

$$\|\Phi_B\|_{S_p \to S_q} \geq \frac{\|\Phi_B(\text{diag}(y))\|_{S_q}}{\|\text{diag}(y)\|_{S_p}} = \frac{\|\text{diag}(By)\|_{S_q}}{\|y\|_{L_p}} = \frac{\|By\|_{L_q}}{\|y\|_{L_p}} = \|B\|_{L_p \to L_q}.$$

This proves the first part.

The cut norm of a matrix takes the supremum over $x, y \in \{0,1\}^n$. Instead we can relax this to $x, y \in [0,1]^n$, since by linearity the supremum will always be attained by the extreme points. Similarly, for the superoperator case, we use Equation (3.3). Then, there exist $x, y \in [0,1]^n$ such that $\|B\|_{\text{cut}} = |\langle Bx, y\rangle|$. We have $\text{diag}(x), \text{diag}(y) \succeq 0$ and $\|\text{diag}(x)\|_{S_\infty}, \|\text{diag}(y)\|_{S_\infty} \leq 1$. Therefore

$$\|\Phi_B\|_{\text{cut}} \geq |\langle \text{diag}(y), \Phi_B(\text{diag}(x))\rangle| = |\langle \text{diag}(y), \text{diag}(Bx)\rangle| = |\langle y, Bx\rangle| = \|B\|_{\text{cut}}.$$

In the other direction, let $X, Y \in M_n(\mathbb{C})$ such that $X, Y \succeq 0$ and with operator norm at most 1. Define x, y to be the diagonals of X, Y, i.e. $x_i = X_{ii}$ and $y_i = Y_{ii}$. By Proposition 3.2.1 we have $\|x\|_{L_\infty}, \|y\|_{L_\infty} \leq 1$. Since $X, Y \succeq 0$ we know all diagonal entries of X and Y are real and non-negative, so we have $x, y \in [0,1]^n$. We conclude

$$\|B\|_{\text{cut}} \geq |\langle y, Bx\rangle| = |\langle \text{diag}(y), \text{diag}(Bx)\rangle| = |\langle Y, \Phi_B(X)\rangle| = \|\Phi_B\|_{\text{cut}},$$

completing the proof. □

Note that $\|A - \frac{1}{n}J\|_{L_2 \to L_2}$ is the second largest eigenvalue in absolute value of the matrix A, so spectral expansion is preserved under the embedding of graphs into quantum channels. Also, uniformity is preserved since the cut-norm does not change.

The following proposition shows that the embedding (3.6) preserves transitivity. This shows that our Theorem 3.3.3 generalizes the main result of [CZ17], albeit with a slightly worse constant.

3.3.8. Proposition. *For any $A \in M_n(\mathbb{C})$, A is vertex transitive if and only if Φ_A is irreducibly covariant.*

Proof:
Suppose A is vertex transitive. Let $\pi \in \text{Aut}(A)$ be a permutation and $P_\pi \in M_n(\mathbb{C})$ be the associated permutation matrix, so that $P_\pi A P_\pi^* = A$. Then,

$$\begin{aligned}
\Phi_A(P_\pi X P_\pi^*) &= \sum_{i,j} A_{ij} (P_\pi X P_\pi^*)_{jj} E_{ii} \\
&= \sum_{i,j} A_{ij} X_{\pi^{-1}(j)\pi^{-1}(j)} E_{ii} \\
&= \sum_{i,j} A_{i\pi(j)} X_{jj} E_{ii} \\
&= \sum_{i,j} A_{\pi(i)\pi(j)} X_{jj} E_{\pi(i)\pi(i)} \\
&= \sum_{i,j} A_{\pi(i)\pi(j)} X_{jj} (P_\pi E_{ii} P_\pi^*) = P_\pi \Phi_A(X) P_\pi^*.
\end{aligned}$$

This shows that for all $\pi \in \text{Aut}(A)$ we have $\Phi_A(P_\pi X P_\pi^*) = P_\pi \Phi_A(X) P_\pi^*$.

Let $U(1) = \{c \in \mathbb{C} : |c| = 1\}$ be the complex unit circle. For $\alpha \in (U(1))^n$, define $U_\alpha := \operatorname{diag}(\alpha)$. We have $U_\alpha E_{ii} U_\alpha^* = |\alpha_i|^2 E_{ii} = E_{ii}$ and $(U_\alpha X U_\alpha^*)_{ii} = |\alpha_i|^2 X_{ii} = X_{ii}$. Therefore

$$\Phi_A(U_\alpha X U_\alpha^*) = \sum_{i,j} A_{ij} (U_\alpha X U_\alpha^*)_{jj} E_{ii} = \sum_{i,j} A_{ij} X_{jj} U_\alpha E_{ii} U_\alpha^* = U_\alpha \Phi_A(X) U_\alpha^*.$$

We combine these two observations as follows. First we have that

$$\left(\mathop{\mathbb{E}}_{\alpha \in U(1)^n} U_\alpha X U_\alpha^* \right)_{ij} = \mathop{\mathbb{E}}_{\alpha \in U(1)^n} \alpha_i X_{ij} \overline{\alpha_j} = \int_0^{2\pi} \int_0^{2\pi} e^{i\theta_i} X_{ij} e^{-i\theta_j} \, d\theta_i d\theta_j = X_{ii} \delta_{ij}$$

If A is vertex transitive then for all $x \in \mathbb{C}^n$ we have $\mathbb{E}_{\pi \in \mathrm{Aut}(A)} P_\pi \operatorname{diag}(x) P_\pi^* = (\mathbb{E}_i \, x_i) \operatorname{Id}$. Therefore

$$\mathop{\mathbb{E}}_{\substack{\pi \in \mathrm{Aut}(A) \\ \alpha \in U(1)^n}} (P_\pi U_\alpha) X (P_\pi U_\alpha)^* = \mathop{\mathbb{E}}_{\pi \in \mathrm{Aut}(A)} P_\pi \left(\mathop{\mathbb{E}}_{\alpha \in U(1)^n} U_\alpha X U_\alpha^* \right) P_\pi^* = \frac{\operatorname{Tr}(X)}{n} \operatorname{Id}.$$

Letting $G \subset M_n(\mathbb{C})$ be the subgroup generated by the U_α and P_π for $\pi \in \mathrm{Aut}(A)$, we see that for any $g \in G$

$$\Phi_A(gXg^*) = g \Phi_A(X) g^*$$

and by the previous equation and Lemma 3.3.6, G acts irreducibly on $\mathbb{C}^n$ (and it is unitary). This proves Φ is irreducibly covariant with respect to the group G with equal representations.

For the other direction, let $U : G \to U(n)$ be the irreducible representation such that Φ_A is irreducibly covariant, i.e. $\Phi_A(U(g) X U^*(g)) = U(g) \Phi_A(X) U^*(g)$ for all $g \in G$. Define $P_g \in M_n(\mathbb{C})$ as $(P_g)_{ij} = |U(g)_{ij}|^2$ so that $(U(g) E_{jj} U(g)^*)_{ii} = (P_g)_{ij}$. Then

$$\begin{aligned} A_{kl} = \operatorname{Tr}[E_{kk} \Phi_A(E_{ll})] &= \operatorname{Tr}[U(g) E_{kk} U(g)^* \; \Phi_A(U(g) E_{ll} U(g)^*)] \\ &= \sum_{ij} A_{ij} (P_g)_{jl} (P_g)_{ik} = (P_g^T A P_g)_{kl}, \end{aligned}$$

showing $P_g^T A P_g = A$. Since $U(g)$ is unitary, P_g is doubly stochastic so by Birkhoff's Theorem P_g is a convex combination of permutation matrices, i.e., $P_g = \mathbb{E}_i \, \Pi_i$ for some (not necessarily uniform) probability distribution and where Π_i is a permutation matrix. We have

$$A_{kl} = (P_g^T A P_g)_{kl} = \mathop{\mathbb{E}}_i \mathop{\mathbb{E}}_j (\Pi_i^T A \Pi_j)_{kl} = \mathop{\mathbb{E}}_i \mathop{\mathbb{E}}_j A_{\pi_i(k) \, \pi_j(l)}.$$

Since A is $\{0,1\}$-valued, it follows that if $A_{kl} = 1$ then all elements of the convex combination on the right-hand side must be 1, and if $A_{kl} = 0$ then all elements

of the right hand side must be 0. Therefore, for all i we have $\Pi_i^T A \Pi_i = A$. By irreducibility, we have for all k, l that

$$\frac{1}{n} = \frac{\mathrm{Tr}[E_{kk}]}{n}\mathrm{Id}_{ll} = \left(\mathop{\mathbb{E}}_{g\in G} U(g)E_{kk}U^*(g)\right)_{ll} = \mathop{\mathbb{E}}_{g\in G} |U(g)_{lk}|^2 ,$$

showing $\mathbb{E}_{g\in G}(P_g)_{lk} = 1/n$. It follows that there is a $g \in G$ such that $(P_g)_{lk} > 0$. Decomposing P_g into permutation matrices shows there is a $\Pi \in \mathrm{Aut}(A)$ such that $\Pi_{lk} = 1$. This holds for all k, l, proving the lemma. □

3.3.4 Randomizing superoperators

We prove the following analogue of one of the results from [CGW89] showing that for any d-regular graph G, we have $\lambda(G) \leq (2\varepsilon(G)/\delta^2)^{1/4}$, where $\delta = d/n$ is the edge density. This in particular establishes a tight relation between spectral expansion and uniformity for sequences of graphs with $\delta_n \geq \Omega(1)$. For $A \in M_n(\mathbb{C})$, we have $\|A\|_{L_1\to L_\infty} = n \sup_{ij} |A_{ij}|$, and for an n-vertex d-regular graph with normalized adjacency matrix A we have $\sup_{ij} |A_{ij}| = \frac{1}{d}$ so $\|A - J/n\|_{\ell_1\to\ell_\infty} = \frac{1}{\delta} - 1$ with J being the all-ones matrix. Therefore, a sequence of graphs with normalized adjacency matrices A_n is dense exactly when $\|A_n - J_n/n\|_{L_1\to L_\infty} \leq \mathcal{O}(1)$, where J_n is the all-ones n by n matrix .

Let Π be the projector onto the identity matrix. A superoperator Φ is said to be *η-randomizing* if $\|\Phi - \Pi\|_{S_1\to S_\infty} \leq \eta$, which when $\eta \leq \mathcal{O}(1)$, may thus be seen as an analogue of density. Note that by Theorem 3.3.7 the embedding of any dense graph is $\mathcal{O}(1)$-randomizing.

3.3.9. Proposition. *Let $\Phi : M_n(\mathbb{C}) \to M_n(\mathbb{C})$ be a superoperator that is $\mathcal{O}(1)$-randomizing. Then, $\lambda(\Phi) \leq \mathcal{O}(\varepsilon(\Phi)^{1/4})$.*

To prove Proposition 3.3.9, we require the following lemma.

3.3.10. Lemma. *Let $\Phi : M_n(\mathbb{C}) \to M_n(\mathbb{C})$ be a superoperator and define $C = \|\Phi\|_{S_1\to S_\infty}$. Then we have $\|\Phi\|_{S_2\to S_2} \leq \left(C^3\|\Phi\|_{S_\infty\to S_1}\right)^{1/4}$.*

Proof:
Note that by definition of C we have $|\langle Q, \Phi(P)\rangle| \leq C\|Q\|_{S_1}\|P\|_{S_1}$. Let $X, Y \in M_n(\mathbb{C})$ be such that $\langle Y, \Phi(X)\rangle = \|\Phi\|_{S_2\to S_2}$ with $\|X\|_{S_2} = \|Y\|_{S_2} = 1$. Write $X = \frac{1}{n}\sum_{i=1}^n \lambda_i P_i$ and $Y = \frac{1}{n}\sum_{i=1}^n \mu_i Q_i$ with P_i, Q_i rank-1 matrices with $\|Q_i\|_{S_1} =$

$\|P_i\|_{S_1} = 1$. We have $\|\lambda\|_{L_2} = \|\mu\|_{L_2} = 1$ and by applying Cauchy-Schwarz twice,

$$\begin{aligned}
|\langle Y, \Phi(X)\rangle|^4 &= \Big| \mathop{\mathbb{E}}_{ij} \lambda_i \mu_j \langle Q_j, \Phi(P_i)\rangle \Big|^4 \\
&\le \Big(\mathop{\mathbb{E}}_{i} \lambda_i^2 \Big)^2 \Big(\mathop{\mathbb{E}}_{i} \big| \mathop{\mathbb{E}}_{j} \mu_j \langle Q_j, \Phi(P_i)\rangle \big|^2 \Big)^2 \\
&= \Big(\mathop{\mathbb{E}}_{i,j,j'} \mu_j \mu_{j'} \langle Q_j, \Phi(P_i)\rangle \langle P_i, \Phi^*(Q_{j'})\rangle \Big)^2 \\
&\le \Big(\mathop{\mathbb{E}}_{j,j'} \mu_j^2 \mu_{j'}^2 \Big) \Big(\mathop{\mathbb{E}}_{j,j'} \Big| \mathop{\mathbb{E}}_{i} \langle Q_j, \Phi(P_i)\rangle \langle P_i, \Phi^*(Q_{j'})\rangle \Big|^2 \Big) \\
&= \mathop{\mathbb{E}}_{i,i',j,j'} \langle Q_j, \Phi(P_i)\rangle \langle P_i, \Phi^*(Q_{j'})\rangle \langle Q_{j'}, \Phi(P_{i'})\rangle \langle P_{i'}, \Phi^*(Q_j)\rangle,
\end{aligned}$$

where all indices are averaged from 1 to n. Now we see

$$\begin{aligned}
|\langle Y, \Phi(X)\rangle|^4 &\le \mathop{\mathbb{E}}_{i,j} \langle Q_j, \Phi(P_i)\rangle \Big\langle \mathop{\mathbb{E}}_{j'} \langle Q_{j'}, \Phi(P_i)\rangle Q_{j'}, \Phi\big(\mathop{\mathbb{E}}_{i'} \langle P_{i'}, \Phi^*(Q_j)\rangle P_{i'} \big) \Big\rangle \\
&\le \mathop{\mathbb{E}}_{i,j} |\langle Q_j, \Phi(P_i)\rangle| \, \|\Phi\|_{S_\infty \to S_1} \, \big\| \mathop{\mathbb{E}}_{j'} \langle Q_{j'}, \Phi(P_i)\rangle Q_{j'} \big\|_{S_\infty} \, \big\| \mathop{\mathbb{E}}_{i'} \langle P_{i'}, \Phi^*(Q_j)\rangle P_{i'} \big\|_{S_\infty} \\
&\le \mathop{\mathbb{E}}_{i,j} |\langle Q_j, \Phi(P_i)\rangle| \, \|\Phi\|_{S_\infty \to S_1} \, \max_{j'} |\langle Q_{j'}, \Phi(P_i)\rangle| \, \max_{i'} |\langle Q_j, \Phi(P_{i'})\rangle| \\
&\le C^3 \|\Phi\|_{S_\infty \to S_1}.
\end{aligned}$$

□

Proof of Theorem 3.3.9:
Let $\Pi(X) = \frac{1}{n}\mathrm{Tr}[X]\mathrm{Id}$ be the projector on to the identity. By assumption, we have $\|\Phi - \Pi\|_{\mathrm{cut}} = \varepsilon(\Phi)$. Define $C = \|\Phi - \Pi\|_{S_1 \to S_\infty}$. Using Lemma 3.2.2 and Lemma 3.3.10 applied to $\Phi - \Pi$ we find $\|\Phi - \Pi\|_{S_2 \to S_2} \le (C^3 \pi^2 \varepsilon(\Phi))^{1/4}$. □

3.4 Optimality of constants

3.4.1 Commutative case

In this section we prove the fourth bullet point in our introduction. Theorem 3.3.1 shows that $K_G^{\mathbb{C}}$ bounds the ratio of the $L_2 \to L_2$ and $L_\infty \to L_1$ norms, and Lemma 3.2.2 (the matrix version) shows that π^2 bounds the ratio of the $L_\infty \to L_1$ norm and the cut norm. We now prove the optimality of the combined inequality.

Let $S^{m-1} = \{x \in \mathbb{C}^m : \|x\|_{L_2} = 1\}$ denote the $(m-1)$-dimensional unit sphere endowed with its Haar probability measure μ.

3.4.1. Theorem. *For any $\varepsilon > 0$ there exist positive integers m, k and a transitive covariant linear map $M : C(S^{m-1} \times [k]) \to C(S^{m-1} \times [k])$ such that $\|M\|_{L_2 \to L_2} \ge (\pi^2 K_G^{\mathbb{C}} - \varepsilon)\|M\|_{\mathrm{cut}}$.*

The optimality of π^2 between the $L_\infty \to L_1$ norm and the cut norm is already covered in Lemma 3.2.2. We show that $K_G^{\mathbb{C}}$ is optimal in the sense that Theorem 3.3.1 cannot be improved (despite the fact that the exact value of the Grothendieck constant $K_G^{\mathbb{C}}$ is unknown). We do this in Lemma 3.4.2 below. Then in Theorem 3.4.3 we show that any map can be lifted to one on a bigger space with appropriately bounded cut norm. The combination of these lemmas proves our theorem.

In the introduction we also mentioned the optimal constant $4K_G$ in the case where the field is $\mathbb{R}$ instead of $\mathbb{C}$. The proofs below still apply in this case, with only small modifications.

3.4.2. Lemma. *For any $\varepsilon > 0$ there exists a positive integer m and a transitive covariant linear map $B : C(S^{m-1}) \to C(S^{m-1})$ such that*

$$\|B\|_{L_2 \to L_2} \geq (K_G^{\mathbb{C}} - \varepsilon)\|B\|_{L_\infty \to L_1}.$$

Proof:
By definition of the Grothendieck constant, for any $\varepsilon > 0$ there exists an $n \in \mathbb{N}$ and a linear map $A \in M_n(\mathbb{C})$ such that $\|A\|_{\mathrm{G}} \geq (K_G^{\mathbb{C}} - \varepsilon)\|A\|_{L_\infty \to L_1}$. This map A might not be transitive covariant, so from it we will now construct a transitive covariant linear map $B : C(S^{2n-1}) \to C(S^{2n-1})$ such that $\|B\|_{L_\infty \to L_1} \leq \|A\|_{L_\infty \to L_1}$ and $\|B\|_{L_2 \to L_2} \geq \|A\|_{\mathrm{G}}$. This idea is based on a lemma found in [Bri11].

Let $x^i, y^j \in S^{2n-1}$ be the vectors that attain the Grothendieck norm for A, which can always be assumed to be $2n$-dimensional since there are only $2n$ of them, so

$$\|A\|_G = \Big|\frac{1}{n}\sum_{i,j} A_{ij}\langle x^i, y^j\rangle\Big|.$$

Define the map B by

$$\langle f, B(g)\rangle = \frac{1}{n}\sum_{i,j} A_{ij}\int_{U(2n)} f(Ux^i)g(Uy^j)dU.$$

To bound $\|B\|_{L_\infty \to L_1}$ we have to bound $|\langle f, B(g)\rangle|$ for $f, g : S^{2n-1} \to [-1, 1]$. By the triangle inequality,

$$|\langle f, B(g)\rangle| \leq \int_{U(2n)} \Big|\frac{1}{n}\sum_{i,j} A_{ij} f(Ux^i)g(Uy^j)\Big|dU \leq \int_{U(2n)} \|A\|_{L_\infty \to L_1} dU \leq \|A\|_{L_\infty \to L_1}.$$

Now for each $i \in [2n]$ let $f_i \in C(S^{2n-1})$ be given by $f_i(x) = x_i$ (i.e. the i-th

coordinate). Then,

$$\begin{aligned}\frac{1}{2n}\sum_{i=1}^{2n}\langle f_i, B(f_i)\rangle &\le \frac{1}{2n}\sum_{i=1}^{2n}\|B\|_{L_2\to L_2}\|f_i\|_{L_2}^2\\ &= \|B\|_{L_2\to L_2}\int_{S^{2n-1}}\frac{1}{2n}\sum_{i=1}^{2n}x_i^2 d\mu(x)\\ &= \|B\|_{L_2\to L_2}.\end{aligned}$$

On the other hand,

$$\frac{1}{2n}\sum_{i=1}^{2n}\langle f_i, B(f_i)\rangle = \frac{1}{n}\sum_{i,j}A_{ij}\int_{U(2n)}\langle Ux^i, Uy^j\rangle dU = \frac{1}{n}\sum_{i,j}A_{ij}\langle x^i, y^j\rangle = \|A\|_G,$$

so we conclude $\|B\|_{L_2\to L_2}\ge\|A\|_G$. We will show B is transitive covariant with respect to $\Gamma = U(2n)$. To show B is invariant, we have to prove that for all $V\in U(2n)$ we have $\langle f^V, B(g^V)\rangle = \langle f, B(g)\rangle$. Indeed,

$$\begin{aligned}\langle f^V, B(g^V)\rangle &= \frac{1}{n}\sum_{i,j}A_{ij}\int_{U(2n)}f(VUx^i)g(VUy^j)dU\\ &= \frac{1}{n}\sum_{i,j}A_{ij}\int_{U(2n)}f(U'x^i)g(U'y^j)dU' = \langle f, B(g)\rangle,\end{aligned}$$

which completes the proof. □

3.4.3. Lemma. *Let S be any compact set and let $B : C(S)\to C(S)$ be a linear map. For any $\varepsilon>0$ there exists a $k\in\mathbb{N}$ and a linear map $M : C(S\times[k])\to C(S\times[k])$ such that*

$$\frac{\|M\|_{\text{cut}}}{\|M\|_{L_2\to L_2}}\le\big(\frac{1}{\pi^2}+\varepsilon\big)\frac{\|B\|_{L_\infty\to L_1}}{\|B\|_{L_2\to L_2}}$$

and if B is transitive covariant then so is M.

Proof:
We will choose k large enough, to be determined later. For any $f,g\in C(S\times[k])$ define $f^i\in C(S)$ as $f^i(s) := f(s,i)$, and similar for g^i. Define $\omega = e^{2\pi i/k}$. Define a linear map $M : C(S\times[k])\to C(S\times[k])$ as

$$\big(M(f)\big)(t,j) := \frac{1}{k}\sum_{i=1}^{k}\omega^{i-j}B(f^i)(t),\quad \text{for } t\in S \text{ and } j\in[k].$$

We then have

$$\langle g, M(f)\rangle_{S\times[k]} = \frac{1}{k^2}\Big\langle \sum_i \omega^i g^i, B\big(\sum_j \omega^j f^j\big)\Big\rangle_S$$

where one factor of $\frac{1}{k}$ comes from our normalization of the inner product. This implies

$$\big|\langle g, M(f)\rangle_{S\times[k]}\big| \leq \|B\|_{L_\infty\to L_1}\Big\|\frac{1}{k}\sum_{i=1}^k \omega^i g^i\Big\|_{L_\infty}\Big\|\frac{1}{k}\sum_{j=1}^k \omega^j f^j\Big\|_{L_\infty}. \tag{3.7}$$

If $f, g \in C(S\times[k])$ are the $[0,1]$-valued functions that attain the cut norm of M, then by (3.7)

$$\|M\|_{\mathrm{cut}} \leq \Big(\frac{1}{\pi^2}+\varepsilon\Big)\|B\|_{L_\infty\to L_1},$$

where we used Theorem 3.4.4 to bound $\Big\|\frac{1}{k}\sum_{i=1}^k \omega^i g^i\Big\|_{L_\infty}$.

Let $u, v \in C(S)$ with $\|u\|_{L_2} = \|v\|_{L_2} = 1$ be such that $\|B\|_{L_2\to L_2} = \langle v, B(u)\rangle_S$. Now define $f_{(u)}, g_{(v)} \in C(S\times[k])$ as $f_{(u)}(s,i) := \omega^{-i}u(s)$ and $g_{(v)}(s,i) := \omega^{-i}v(s)$, which also have L_2-norm equal to 1. We then see

$$\|M\|_{L_2\to L_2} \geq \big\langle g_{(v)}, M(f_{(u)})\big\rangle_{S\times[k]} = \langle v, B(u)\rangle_S = \|B\|_{L_2\to L_2}.$$

The combination of these observations completes the first part of the proof. Now assume B is transitive covariant with respect to Γ, so $B(f^\pi)(\pi^{-1}(s)) = B(f)(s)$ for all $s \in S$ and $\pi \in \Gamma$. Define a new group Γ' as the cartesian product $\Gamma' = \Gamma \times \mathbb{Z}_k$. For $(\pi, m) \in \Gamma'$ define the action $(\pi, m) : S\times[k] \to S\times[k]$ as $(\pi,m)(s,i) = (\pi(s), i+m)$. By entering $f^{(\pi,m)}$ into the definition of M it follows that $M^{(\pi,m)} = M$, so M is transitive covariant with respect to Γ', completing the proof. □

3.4.4. Lemma. *Let $\varepsilon > 0$, then there exists a $k_0 \in \mathbb{N}$ such that for all $k \geq k_0$ and $x \in [0,1]^k$ we have*

$$\Big|\frac{1}{k}\sum_{j=1}^k e^{2\pi i\, j/k} x_j\Big| \leq \frac{1}{\pi} + \varepsilon.$$

Proof:
First let k_0 be arbitrary, to be determined later and $k \geq k_0$. Define $y \in [-1,1]^k$ as $y_i = 2x_i - 1$, then

$$\Big|\frac{1}{k}\sum_{j=1}^k e^{2\pi i\, j/k} x_j\Big| = \frac{1}{2}\Big|\frac{1}{k}\sum_{j=1}^k e^{2\pi i\, j/k} y_j\Big| = \frac{1}{2}e^{2\pi i\phi}\frac{1}{k}\sum_{j=1}^k e^{2\pi i\, j/k} y_j.$$

In the first equality we used that $\sum_{j=1}^{k} e^{2\pi i\, j/k} = 0$. In the second equality we used that there exists a ϕ such that the full expression becomes real and positive. Since $e^{i\theta} = \cos(\theta) + i\sin(\theta)$ and the full expression is real, we know the sin component vanishes and therefore

$$\frac{1}{2}\frac{1}{k}\sum_{j=1}^{k} e^{2\pi i(\phi+j/k)} y_j = \frac{1}{2}\frac{1}{k}\sum_{j=1}^{k} \cos(2\pi(\phi+j/k)) y_j.$$

Now note that $\cos(2\pi(\phi+j/k))y_j \leq \big|\cos(2\pi(\phi+j/k))\big|$ and hence

$$\frac{1}{2}\frac{1}{k}\sum_{j=1}^{k} \big|\cos(2\pi(\phi+j/k))\big| \quad \overset{k\to\infty}{\longrightarrow} \quad \frac{1}{2}\int_0^1 \big|\cos\big(2\pi(\phi+x)\big)\big| dx = \frac{1}{\pi}.$$

This completes the proof. □

3.4.2 Non-commutative case

In the non-commutative case we show optimality of Theorem 3.3.3. By Theorem 3.2.2, the factor π^2 between the cut-norm and $S_\infty \to S_1$-norm is also optimal. In contrast with the commutative case, our work leaves the optimality of the combined inequality in Theorem 3.3.4 as an open problem. Straightforward analogues of the techniques employed in Theorem 3.4.3 did not follow through in the non-commutative case.

3.4.5. Proposition. *For any $\varepsilon > 0$, there exists a positive integer n and an irreducibly covariant superoperator $\Phi : M_n(\mathbb{C}) \to M_n(\mathbb{C})$ such that $\|\Phi\|_{S_2\to S_2} \geq (2-\varepsilon)\|\Phi\|_{S_\infty\to S_1}$.*

One of the forms of the non-commutative Grothendieck inequality, equivalent to Theorem 3.3.5, is the following [Pis12]. Let $\Phi : M_n(\mathbb{C}) \to M_n(\mathbb{C})$ be a linear map and $x_i, y_j \in M_n(\mathbb{C})$ finite sets of matrices. Then,

$$\Big|\sum_i \langle x_i, \Phi(y_i)\rangle\Big| \leq K'_G \|\Phi\|_{S_\infty\to S_1} \left(\frac{\|\sum_i x_i^* x_i\| + \|\sum_i x_i x_i^*\|}{2} \cdot \frac{\|\sum_i y_i^* y_i\| + \|\sum_i y_i y_i^*\|}{2}\right)^{1/2} \tag{3.8}$$

where $K'_G \leq 2$ and the norms on the right hand side are operator norms $\|\cdot\|_{S_\infty}$. To show tightness, i.e. $K'_G \geq 2$, Haagerup and Itoh [HI95] (see [Pis12] for a survey) gave an explicit family of operators for which (3.8) gives a lower bound of K'_G approaching 2. We will show that slight modifications of these operators are irreducibly covariant, which proves Proposition 3.4.5. It is instructive to repeat their construction. The proof uses techniques familiar in the context of the antisymmetric Fock space, but our proof is self contained.

3.4.6. Lemma ([HI95]). *For each $n \in \mathbb{N}$ there exists a $d \in \mathbb{N}$ and a linear map $\Phi : M_d(\mathbb{C}) \to M_d(\mathbb{C})$ with sets of matrices $\{x_i\}$, $\{y_i\}$ such that* (3.8) *yields $K'_G \geq (2n+1)/(n+1)$.*

Proof:
Let $H = \mathbb{C}^{2n+1}$ and consider the antisymmetric k-fold tensor product $H^{\wedge k}$ which is a linear subspace of the k-fold tensor product $H^{\otimes k}$. A basis of $H^{\wedge k}$ is formed by vectors $e_{i_1} \wedge e_{i_2} \wedge \cdots \wedge e_{i_k}$ with $i_1 < \cdots < i_k$ where the e_i are standard basis vectors of H. Here $\wedge$ is the wedge product or exterior product, which has the property $x \wedge y = -y \wedge x$ and is given by $x \wedge y = x \otimes y - y \otimes x$, for $x, y \in H$. We will consider $k = n$ and $k = n+1$ so that the dimension of $H^{\wedge k}$ is $d = \binom{2n+1}{n}$ for both $k = n$ and $k = n+1$.

For $1 \leq i \leq (2n+1)$, define $c_i : H^{\wedge n} \to H^{\wedge(n+1)}$ as $c_i(x) := e_i \wedge x$, which physicists call the fermionic creation operator. Its adjoint $c_i^* : H^{\wedge(n+1)} \to H^{\wedge n}$ is known as the annihilation operator. By the antisymmetric property, $c_i(x) = 0$ whenever e_i was present in x, i.e., when $x = e_i \wedge x'$. The operator $c_i c_i^*$, also known as the number operator, is a projector onto the space spanned by basis vectors in which e_i is present. The operator $c_i^* c_i$ is a projector onto the space where e_i is *not* present. Since there are always $n+1$ vectors present in $H^{\wedge(n+1)}$ and $n+1$ vectors *not* present in $H^{\wedge n}$, we have

$$\sum_{i=1}^{2n+1} c_i c_i^* = (n+1)\mathrm{Id}_{H^{\wedge(n+1)}} \quad \text{and} \quad \sum_{i=1}^{2n+1} c_i^* c_i = (n+1)\mathrm{Id}_{H^{\wedge n}}.$$

We will now argue that

$$\langle c_i, c_j \rangle := \frac{1}{d}\,\mathsf{Tr}(c_i^* c_j) = \delta_{i,j}\frac{n+1}{2n+1}, \tag{3.9}$$

$$\|\sum_{i=1}^{2n+1} \alpha_i c_i\|_{S_1} = \|\alpha\|_{L_2}\frac{n+1}{\sqrt{2n+1}} \qquad \text{for } \alpha \in \mathbb{C}^{2n+1}. \tag{3.10}$$

The $\delta_{i,j}$ in (3.9) follows because $\langle x, c_i^* c_j x\rangle = 0$ for any $x = e_{k_1} \wedge \cdots \wedge e_{k_n}$ when $i \neq j$. The factor $\frac{n+1}{2n+1}$ follows by taking the trace of one of the sums above and noting that by symmetry in i, every term of the sum must have the same trace. To prove (3.10), first note that for any unitary $U \in \mathrm{U}(2n+1)$ we have

$$U^{\otimes(n+1)} \cdot c_i \cdot (U^{\otimes n})^{-1} = \sum_j U_{ji} c_j, \tag{3.11}$$

which can be shown by proving it for all basis states:

$$\begin{aligned} U^{\otimes(n+1)}c_i(U^{\otimes n})^{-1}(e_{k_1}\wedge ...\wedge e_{k_n}) &= U^{\otimes(n+1)}c_i(U^{-1}e_{k_1}\wedge ...\wedge U^{-1}e_{k_n}) \\ &= U^{\otimes(n+1)}(e_i\wedge U^{-1}e_{k_1}\wedge ...\wedge U^{-1}e_{k_n}) \\ &= (Ue_i\wedge e_{k_1}\wedge ...\wedge e_{k_n}) \\ &= (\sum_j U_{ji}e_j\wedge e_{k_1}\wedge ...\wedge e_{k_n}) \\ &= \sum_j U_{ji}c_j(e_{k_1}\wedge ...\wedge e_{k_n}). \end{aligned}$$

The trace-norm is unitarily invariant, so (3.11) implies $\|c_i\|_{S_1} = \|\sum_j U_{ji}c_j\|_{S_1}$. Since $c_i^*c_i$ is a projector, we have $\sqrt{c_i^*c_i} = c_i^*c_i$ and hence $\|c_i\|_{S_1} = \frac{1}{d}\,\mathsf{Tr}(c_i^*c_i)$. Now let $\alpha\in\mathbb{C}^{2n+1}$ with $\sum_i|\alpha_i|^2 = 1$, then there is a unitary $U\in\mathrm{U}(2n+1)$ such that the i-th row of U is α. Note that $\|\alpha\|_{L_2} = 1/\sqrt{2n+1}$ since we use normalized Ll_2-norms, which implies (3.10).

Since the dimensions of $H^{\wedge n}$ and $H^{\wedge(n+1)}$ are equal, we can identify the space of linear maps $L(H^{\wedge n}, H^{\wedge(n+1)})$ with $M_d(\mathbb{C})$ (by choosing bases for $H^{\wedge n}$ and $H^{\wedge(n+1)}$), and define the following operator $\Phi : M_d(\mathbb{C})\to M_d(\mathbb{C})$,

$$\Phi(x) = \sum_{i=1}^{2n+1}\langle c_i, x\rangle\, c_i.$$

Consider (3.8) for Φ with $x_i = y_i = c_i$. For the left hand side, note that by (3.9) we have

$$\Big|\sum_{j=1}^{2n+1}\langle c_j, \Phi(c_j)\rangle\Big| = \Big|\sum_{i,j=1}^{2n+1}\langle c_i, c_j\rangle\,\langle c_j, c_i\rangle\Big| = \frac{(n+1)^2}{2n+1}.$$

For the right-hand side of (3.8), we require $\|\Phi\|_{S_\infty\to S_1} = \sup_{\|x\|_{S_\infty}=1}\|\Phi(x)\|_{S_1}$. For any $x\in M_d(\mathbb{C})$, define $v^{(x)}\in\mathbb{C}^{2n+1}$ as $v_i^{(x)} = \langle c_i, x\rangle$. Note that $\|v\|_{L_2} = \sup_{\|\alpha\|_{L_2}=1}|\langle v,\alpha\rangle|$. First apply (3.10) to obtain

$$\|\Phi(x)\|_{S_1} = \|\sum_{i=1}^{2n+1}\langle c_i, x\rangle c_i\|_{S_1} = \|v^{(x)}\|_{L_2}\frac{n+1}{\sqrt{2n+1}} = \sup_{\|\alpha\|_{L_2}=1}|\langle v^{(x)},\alpha\rangle|\frac{n+1}{\sqrt{2n+1}}.$$

Using (3.10) again, we compute $\sup_{\|x\|_{S_\infty}=1}|\langle v^{(x)},\alpha\rangle|$ for arbitrary α with $\|\alpha\|_{L_2} = 1$,

$$\begin{aligned} \sup_{\|x\|_{S_\infty}=1}|\langle v^{(x)},\alpha\rangle| &= \sup_{\|x\|_{S_\infty}=1}\frac{1}{2n+1}\big|\langle x, \sum_i\alpha_i c_i\rangle\big| \\ &= \frac{1}{2n+1}\|\sum_i\alpha_i c_i\|_{S_1} = \frac{n+1}{(2n+1)\sqrt{2n+1}}. \end{aligned}$$

We obtain $\|\Phi\|_{S_\infty \to S_1} = (n+1)^2/(2n+1)^2$. Now (3.8) yields $\frac{(n+1)^2}{2n+1} \leq K'_G \frac{(n+1)^2}{(2n+1)^2} \cdot (n+1)$ and therefore $\frac{2n+1}{n+1} \leq K'_G$. □

We use the following fact from [FH13, Theorem 19.14], about the representations of the odd-dimensional complex special orthogonal groups on wedge products of *complex* vector spaces.

3.4.7. Lemma. *Let $n, k \in \mathbb{N}$, $N := 2n+1$ and let $R_k : \mathrm{SO}(N, \mathbb{C}) \to \mathrm{GL}((\mathbb{C}^N)^{\wedge k})$ be given by $A \mapsto A^{\otimes k}$. This representation is irreducible.*

Below, we actually need that the *real* special orthogonal group $\mathrm{SO}(N, \mathbb{R})$ acts irreducibly on the same anti-symmetric space. Fortunately, this is implied by Lemma 3.4.7; see [FH13, pp. 439]. We will also use the fact that R_k and R_{N-k} are *unitarily* equivalent to each other. This is the content of the following proposition [Sim96, Proposition IX.10.4].

3.4.8. Proposition. *For positive integer n and $N = 2n+1$ and $k \in \{1, \ldots, N\}$, let R_k be the representation as in lemma 3.4.7. Then, there exists an isometry $V_k \colon (\mathbb{C}^N)^{\wedge k} \to (\mathbb{C}^N)^{\wedge (N-k)}$ such that*

$$V_k R_k(A) = R_{N-k}(A) V_k, \quad \forall A \in \mathrm{SO}(N, \mathbb{R}).$$

Proof of Proposition 3.4.5:
Let d be the dimension of $(\mathbb{C}^N)^{\wedge n}$ and let $\Phi : M_d(\mathbb{C}) \to M_d(\mathbb{C})$ be as in the proof of Lemma 3.4.6. For each $k \in \mathbb{N}$, let $R_k : \mathrm{SO}(N, \mathbb{R}) \to \mathrm{GL}(H^{\wedge k})$ be the representation $A \mapsto A^{\otimes k}$, which is irreducible by Theorem 3.4.7. Define, for notational convenience, $\pi := R_{n+1}$ and $\rho := R_n$. We first show that for all $A \in \mathrm{SO}(N, \mathbb{R})$, we have

$$\Phi(\pi(A) x \rho^*(A)) = \pi(A)\, \Phi(x)\, \rho^*(A). \tag{3.12}$$

For the left-hand side, note that

$$\begin{aligned}
\Phi(\pi(A) x \rho^*(A)) &= \sum_i \big\langle c_i, \pi(A) x \rho^*(A) \big\rangle\, c_i \\
&= \sum_i \big\langle \pi(A)^* c_i \rho(A), x \big\rangle\, c_i \\
&= \sum_i \Big\langle \sum_j A_{ij} c_j, x \Big\rangle\, c_i \\
&= \sum_{ij} A_{ij} \langle c_j, x \rangle\, c_i,
\end{aligned}$$

where we used (3.11) from the proof of Lemma 3.4.6 and noting that $\mathrm{SO}(N,\mathbb{R}) \subset U(N)$ is a subgroup. Using (3.11) again for the right-hand side, we have

$$\begin{aligned}\pi(A)\ \Phi(x)\ \rho^*(A) &= \sum_i \langle c_i, x\rangle\ \pi(A) c_i \rho^*(A) \\ &= \sum_i \langle c_i, x\rangle \sum_j A_{ji} c_j \\ &= \sum_{ij} A_{ij} \langle c_j, x\rangle\ c_i.\end{aligned}$$

which proves (3.12).

Define a new superoperator $\Phi'\colon M_d(\mathbb{C}) \to M_d(\mathbb{C})$ by

$$\Phi'(x) = \Phi(xV^*)V,$$

where $V := V_{n+1}$ is the isometry as in Proposition 3.4.8 (we view V as a matrix in $M_d(\mathbb{C})$ by choosing basis). We first note that this Φ' might also be used in Lemma 3.4.6 to show that the non-commutative Grothendieck constant is 2, since Schatten-norms are unitarily invariant. Hence, if we show that Φ' is irreducibly covariant, we are done. This follows from the following computation, where we use (3.12) and the fact that $V\pi(A) = \rho(A)V$ for all $A \in \mathrm{SO}(N,\mathbb{R})$:

$$\begin{aligned}\Phi'\big(\pi(A)x\pi(A)^*\big) &= \Phi\big(\pi(A)x\pi(A)^*V^*\big)V \\ &= \Phi\big(\pi(A)xV^*\rho(A)^*\big)V \\ &\overset{(3.12)}{=} \pi(A)\ \Phi(xV^*)\ \rho(A)^*V \\ &= \pi(A)\ \Phi(xV^*)\ V\pi(A)^* \\ &= \pi(A)\ \Phi'(x)\ \pi^*(A),\end{aligned}$$

where the second-last line follows since $\rho(A)^* = V\pi(A)^*V^*$. Hence, Φ' is irreducibly covariant with respect to the irreducible representation π of $\mathrm{SO}(N,\mathbb{R})$. □

3.4.3 Discussion

We have shown that for irreducibly covariant superoperators Φ the inequality $\lambda(\Phi) \leq 2\pi^2\varepsilon(\Phi)$ holds, see Corollary 3.3.4. But we have not resolved the question whether the constant $2\pi^2$ is optimal in this inequality. However, we have shown that the inequalities in Lemma 3.2.2 and Theorem 3.3.3 have optimal constants of π^2 and 2 respectively. But it is not clear that the constant $2\pi^2$ in Corollary 3.3.4 is optimal, which is a combination of Lemma 3.2.2 and Theorem 3.3.3. The two families of superoperators for which we have shown optimality of Lemma 3.2.2 and Theorem 3.3.3, are different. We could prove optimality of Corollary 3.3.4 by combining these two families of superoperators, but this appears to be non-trivial and is left as an open problem for future work.

Chapter 4

Stabilizer rank and higher-order Fourier analysis

This chapter is based on the paper [Lab21].

4.1 Introduction

The Gottesman-Knill Theorem [Got98, NC02] states that any quantum circuit consisting of Clifford gates can be efficiently classically simulated. The Clifford group on n qubits is generated by the Hadamard gate H, the $\pi/4$ phase gate S, and the entangling CNOT gate. In particular, this means that circuits consisting only of Clifford gates cannot provide a computational advantage over classical computers. We can promote such circuits to universal quantum computers by having access to a non-Clifford gate or (equivalently) a "magic state" [BK05]. It is widely believed that universal quantum computers cannot be efficiently simulated by classical computers: state-of-the-art simulators using modern-day supercomputers are only able to simulate a few dozen qubits [CZH$^+$18, HS17, PGN$^+$17, SSAG16]. So it has to be this magic state that fuels the computational hardness of simulation by classical computers. It is therefore important to understand how much this resource costs in terms of free (efficiently simulatable) resources. These costs are quantified by "measures of magic" [LW20] an example of which is *stabilizer rank*, first introduced in [BSS16]. Here, the free resources are states obtained from the canonical all-zero state by applying only Clifford operations, which are the well-known stabilizer states. To increase our understanding of non-stabilizerness, or the amount of "magic" a quantum state has, a valid approach might be to find a different characterization of these objects. This might introduce new techniques in analyzing measures of magic. It is well known that stabilizer states are characterized by quadratic forms defined on affine subspaces [DDM03, HDDM05, Gro06]. Here we observe that these objects are so-called *nonclassical quadratic phase functions* defined on affine subspaces

which are well-studied objects in higher-order Fourier analysis.

Higher-order Fourier analysis, a still nascent area of mathematics, grew out of a Fourier-analytic proof of Szemerédi's theorem [Sze75] by Gowers [Gow98]. Whereas in Fourier analysis one studies how functions correlate with characters, in higher-order Fourier analysis one studies correlations with functions that resemble polynomials (where characters correspond to linear functions). These polynomial-like functions are known as *polynomial phase functions*.

It turns out that Boolean functions giving the (unnormalised) amplitudes of graph states are examples of quadratic phase functions. In general, quadratic phase functions can be defined using "multiplicative derivatives": for $h \in \mathbb{F}_2^n$, the multiplicative derivative of $f\colon \mathbb{F}_2^n \to \mathbb{C}$ in direction h is

$$\Delta_h f(x) := f(x+h)\overline{f(x)}.$$

Nonclassical polynomial phase functions of degree d are those functions that are constant after taking $d+1$ multiplicative derivatives. It is not difficult to check that graph states, whose amplitude function always has the form $f(x) = (-1)^{q(x)}$ where q is a quadratic polynomial, satisfies this property with $d = 2$. These are referred to as the *classical* quadratic phase functions. However, the nonclassical quadratic phase functions are not exhausted by these examples. It turns out that stabilizer states correspond to functions in this broader class. This establishes a surprising link between higher-order Fourier analysis and quantum information theory. It was shown in [DDM03] (see also [BG16]) that stabilizer states are *quadratic forms* taking values in $\mathbb{Z}_8$ defined on affine subspaces. We will see that they are nonclassical quadratic phase functions on affine subspaces.

Let p be an odd prime and consider qudits of dimension p. Then, the amplitudes of n-qudit stabilizer states are also quadratic phase functions defined on affine subspaces of $\mathbb{F}_p^n$ [HDDM05], see also [Gro06]. It is interesting to note that for primes $p > 2$ the n-qudit stabilizer states are given by classical quadratic polynomials while there are no nonclassical polynomials of degree two, contrary to the case $p = 2$.

Stabilizer rank. *Stabilizer rank* is a measure of magic which was recently extensively analyzed by Bravyi et al. [BBC+19]. The stabilizer rank of a quantum state $|\psi\rangle$, denoted $\chi(|\psi\rangle)$, is the minimal number r such that $|\psi\rangle$ can be written as a linear combination of r stabilizer states. As is well known, any circuit $\mathcal{C}$ consisting of Clifford gates and n copies of the T-gate, given by $|0\rangle\langle 0| + e^{i\pi/4}|1\rangle\langle 1|$, can be implemented using Clifford operations on the n-qubit magic state $|T\rangle^{\otimes n}$, where $|T\rangle = \frac{|0\rangle + e^{i\pi/4}|1\rangle}{\sqrt{2}}$. Then, the stabilizer rank of $|T\rangle^{\otimes n}$ upper bounds the simulation cost of the circuit $\mathcal{C}$. Bravyi, Smith and Smolin [BSS16] showed that the stabilizer rank of the n-qubit magic state is $\Omega(\sqrt{n})$. Very recently Peleg, Shpilka, and Volk [PSV21] showed a lower bound of $\Omega(n)$ for the stabilizer rank of the

n-qubit magic state, which is a quadratic improvement. Here we give the same lower bound but use completely different techniques (from higher-order Fourier analysis) and generalize to qudits of any prime dimension.

Adding any non-Clifford gate to the Clifford gate set could in principle promote it to universal quantum computation. However, we use the generalization of the T gate for qudits from [HV12]. Let us call this gate U (defined below in Section 4.3) and define $|+\rangle := \frac{|0\rangle+|1\rangle+\cdots+|p-1\rangle}{\sqrt{p}}$. Then, the single qudit magic state over $\mathbb{F}_p$ is defined to be

$$|\psi_U\rangle = U|+\rangle.$$

Our main result is the following.

4.1.1. Theorem. *Let p be any prime and let $|\psi_U\rangle^{\otimes n}$ be the n-qudit magic state over $\mathbb{F}_p$. We have that $\chi(|\psi_U\rangle^{\otimes n}) \geq \Omega(n)$.*

This result generalizes [PSV21], but our techniques are completely different and use explicitly tools from higher-order Fourier analysis. Roughly speaking, we show that the function giving the amplitudes of the n-qudit magic state is a cubic nonclassical polynomial phase function for which the polynomial "in the phase" has high rank (see Definition 4.2.5). We then prove that the lower bound for this rank is also a lower bound for the stabilizer rank. In this step, we use a lemma from [PSV21, Claim 3.3] to get a handle on the affine subspaces that appear from the stabilizer states. Apart from this lemma, the techniques are different.

The techniques used here might pave the way to super-linear lower bounds for decompositions in terms of stabilizer states defined on the full space $\mathbb{F}_p^n$.

4.2 Techniques

In this section, we introduce all the definitions necessary for the proof of our main result. Our main result generalizes the recent result of [PSV21], but we use completely different techniques that we explain here as well. Recall the notation from Section 1.1.

4.2.1 Stabilizer states

In [BG16], a succinct representation of n-qubit stabilizer states is given in terms of quadratic forms on affine subspaces of $\mathbb{F}_2^n$. For this, they introduced the following definition. For an affine subspace $H \subset \mathbb{F}_p^n$, we write $L(H) = \{x - y \colon x, y \in H\}$.

4.2.1. Definition. For an affine subspace $H \subset \mathbb{F}_2^n$, a map $Q \colon H \to \frac{1}{8}\mathbb{Z}/\mathbb{Z}$ is called a quadratic form if $\Delta_{h_1}\Delta_{h_2}Q(x)$ is independent of $x \in H$ for all $h_1, h_2 \in L(H)$.

4.2.2. Theorem ([BG16]). *For any n-qubit stabilizer state $|\phi\rangle$, there exists a unique affine subspace $H \subset \mathbb{F}_2^n$ and a quadratic form $Q\colon H \to \frac{1}{8}\mathbb{Z}/\mathbb{Z}$ such that*

$$|\phi\rangle = 2^{-\dim H/2} \sum_{x \in H} e(Q(x))|x\rangle. \tag{4.1}$$

In [BG16] quadratic forms are considered that are maps $Q\colon H \to \frac{1}{8}\mathbb{Z}/\mathbb{Z}$. We will show that the only way that such a map has the property $\Delta_{h_1}\Delta_{h_2}Q(x)$ being independent of $x \in H$ for all $h_1, h_2 \in L(H)$, is if Q actually takes values in $\frac{1}{4}\mathbb{Z}/\mathbb{Z} \subset \frac{1}{8}\mathbb{Z}/\mathbb{Z}$. We will also see that such functions are *nonclassical* polynomials of degree two in the literature of higher-order Fourier analysis. Making this explicit link with higher-order Fourier analysis allows us to import tools from that theory to use in lower bounding the stabilizer rank of quantum states.

Next, qudit stabilizer states where the dimension of the qudit is an odd prime p are somewhat simpler as they are given by quadratic polynomials taking values in $\mathbb{F}_p$ on affine subspaces in $\mathbb{F}_p^n$. Let $\omega = e^{2\pi i/p}$ be a p-th root of unity. A map $Q\colon H \to \mathbb{F}_p$ is a quadratic polynomial on an affine subspace $H \subset \mathbb{F}_p^n$ if $\Delta_{h_1}\Delta_{h_2}Q(x)$ is independent of $x \in H$ for all $h_1, h_2 \in L(H)$.

4.2.3. Theorem ([HDDM05], see also [Gro06]). *Let p be an odd prime and $|\phi\rangle$ an n-qudit stabilizer state where the dimension of the qudit is p. Then, there is an affine subspace $H \subset \mathbb{F}_p^n$ and a quadratic polynomial $Q\colon H \to \mathbb{F}_p$ such that*

$$|\phi\rangle = p^{-\dim(H)/2} \sum_{x \in H} \omega^{Q(x)}|x\rangle. \tag{4.2}$$

4.2.4. Definition. For an n-qudit quantum state $|\psi\rangle$ define its stabilizer rank, denoted $\chi(|\psi\rangle)$, to be the minimal number r needed to write $|\psi\rangle$ as a linear combination of r stabilizer states.

4.2.2 Rank of nonclassical polynomials

In this section we introduce a notion of rank for nonclassical polynomials. The main reference is again [HHL19]. Recall the definition of nonclassical polynomials in Section 1.4.

4.2.5. Definition. Let $P\colon \mathbb{F}_p^n \to \mathbb{T}$ be a nonclassical polynomial. For an integer $d \geq 1$, we define the d-rank, denoted by $\mathrm{rank}_d(P)$, to be the minimal number r such that there are nonclassical polynomials $Q_1, \ldots, Q_r$ all of degree at most $d-1$ and a function $\Gamma\colon \mathbb{T}^r \to \mathbb{T}$ such that

$$\Gamma(Q_1(x), \ldots, Q_r(x)) = P(x).$$

If $d = 1$, the d-rank will be ∞ if P is non-constant and 0 otherwise. The rank of P, denoted $\mathrm{rank}(P)$, is the $\deg(P)$-rank of P.

The next result is a standard application of Fourier analysis and can be thought of as a kind of "inverse theorem". We give the proof here for convenience.

4.2.6. Proposition ([HHL19]). *Let $d \geq 2$, $P\colon \mathbb{F}_p^n \to \mathbb{T}$ be a nonclassical polynomial and $r := \mathrm{rank}_d(P)$. Then there exists a nonclassical polynomial Q of degree at most $d-1$ such that*

$$|\langle e(P), e(Q)\rangle| \geq p^{-(1+\lceil (d-1)/(p-1)\rceil)r}.$$

Proof:
Let $\Gamma\colon \mathbb{T}^r \to \mathbb{T}$ be a map and $Q_1, \dots, Q_r$ be nonclassical polynomials of degree at most $d-1$ such that

$$P(x) = \Gamma(Q_1(x), \dots, Q_r(x)).$$

We can assume that Γ is a map with domain $G := \prod_{i=1}^r \frac{1}{p^{k_i+1}}\mathbb{Z}/\mathbb{Z}$, where k_i is the depth of Q_i. Let $\widehat{G} = \prod_{i=1}^r \mathbb{Z}_{p^{k_i+1}}$ be the dual of G, so that the Fourier decomposition of $e(\Gamma)$ becomes

$$e(\Gamma(z)) = \sum_{\alpha \in \widehat{G}} \widehat{\Gamma}_\alpha e(\langle \alpha, z\rangle).$$

The Fourier decomposition of $e(\Gamma)$ gives a "higher-order" Fourier decomposition of $e(P)$: for $\alpha \in \widehat{G}$ define $Q_\alpha(x) := \sum_{i=1}^r \alpha_i Q_i(x)$, then

$$e(P(x)) = \sum_{\alpha \in \widehat{G}} \widehat{\Gamma}_\alpha e(Q_\alpha(x)).$$

So

$$1 = |\langle e(P), e(P)\rangle| \leq \sum_{\alpha \in \widehat{G}} |\langle e(P), e(Q_\alpha)\rangle|.$$

Indeed, there is an α^* such that

$$|\langle e(P), e(Q_{\alpha^*})\rangle| \geq |\widehat{G}|^{-1}.$$

Since the degree of the polynomials Q_i is at most $d-1$, it follows that $k_i(p-1) \leq d-1$. This implies that

$$|\widehat{G}| = p^{k_1+1+\cdots+k_r+1} \leq p^{(1+\lceil (d-1)/(p-1)\rceil)r}.$$

□

The next lemma tells us how the rank changes if we restrict a polynomial to an affine subspace.

4.2.7. Lemma ([HHL19]). *Let $P\colon \mathbb{F}_p^n \to \mathbb{T}$ be a polynomial of degree $d \geq 2$ and $r := \operatorname{rank}(P)$. Let $U \subset \mathbb{F}_p^n$ be an affine subspace of codimension k and define P' to be the restriction of P to U. If $r > pk + 1$ then P' is a polynomial of degree d and $\operatorname{rank}(P') \geq r - pk$.*

Proof:
We will prove the statement for $k = 1$. The general case will follow after repeated application of the proof for $k = 1$.

Since rank and degree do not change under invertible affine linear transformations, we can assume without loss of generality that $U = \{x \in \mathbb{F}_p^n \colon x_n = 0\}$. Let $\pi\colon \mathbb{F}_p^n \to \mathbb{F}_p^n$ be the projection onto U, so $\pi(x_1, \ldots, x_n) = (x_1, \ldots, x_{n-1}, 0)$. Define $P'' = P - P' \circ \pi$. For $x \in U$ we have $P''(x) = 0$. Let $a \in \mathbb{F}_p \setminus \{0\}$ and $h_a = (0, \ldots, 0, a) \in \mathbb{F}_p^n$. We have that $\Delta_{h_a} P''$ has degree at most $d - 1$ and that $\Delta_{h_a} P''(y) = P''(y + h)$ for all $y \in U$. So P'' agrees with a polynomial Q_a of degree at most $d - 1$ on $U + h_a$. This implies there is a function $\Gamma\colon \mathbb{T}^{p+1} \to \mathbb{T}$ such that $P(x) = \Gamma(|x_n|/p, P'(x), Q_1(x), \ldots, Q_{p-1}(x))$.

Now, if P' has degree at most $d - 1$, then $\operatorname{rank}(P) \leq p + 1 < r$, which is a contradiction. If P' has rank $< r - p$, we get that $\operatorname{rank}(P) < r$ which is again a contradiction. □

Next, we define the *Fourier rank* of a function.

4.2.8. Definition. Let $f\colon \mathbb{F}_p^n \to \mathbb{C}$ be a function and $d \geq 2$. The degree-d Fourier rank of f, denoted $\operatorname{frank}_d(f)$, is the minimal r such that there are polynomials $Q_1, \ldots, Q_r$ of degree at most $d - 1$ such that

$$f(x) = \sum_{i=1}^{r} c_i e(Q_i(x)). \tag{4.3}$$

The following lemma relates the notion of rank of a polynomial and its Fourier rank.

4.2.9. Lemma. *Let $P\colon \mathbb{F}_p^n \to \mathbb{T}$ be a polynomial and $d \geq 2$. Then*

$$\operatorname{frank}_d(e(P)) \geq \operatorname{rank}_d(P).$$

Proof:
Denote by r the degree d Fourier rank of $e(P)$. So there are polynomials $Q_1, \ldots, Q_r$ of degree at most $d - 1$ such that we have a decomposition

$$e(P(x)) = \sum_{i=1}^{r} c_i e(Q_i(x)). \tag{4.4}$$

We will now define a function $\Gamma\colon \mathbb{T}^r \to \mathbb{T}$ as follows. Let $Q\colon \mathbb{F}_p^n \to \mathbb{T}^r$ be defined by $Q(x) = (Q_1(x), \ldots, Q_r(x))$. The map Γ on the image of Q is defined by

$$e(\Gamma(Q_1(x), \ldots, Q_r(x))) = \sum_{i=1}^{r} c_i e(Q_i(x)).$$

For a point z in the complement of the image of Q, we (arbitrarily) set $\Gamma(z) = 0$. But the Γ we just defined has the property that $\Gamma(Q_1(x), \ldots, Q_r(x)) = P(x)$ by (4.8). Hence $r \geq \mathrm{rank}_d(P)$, proving the statement. □

For $d = 2$, Sanyal [San19] shows that $\mathrm{frank}_2(e(P)) \geq \mathrm{rank}_2(P)^2$. This is quadratically better than the above lemma. See Section 4.5 for a discussion for the case $d > 2$.

In the other direction, we have the following lemma.

4.2.10. Lemma. *Let $d \geq 2$ and $P\colon \mathbb{F}_p^n \to \mathbb{T}$ a polynomial. Then*

$$\mathrm{frank}_d(e(P)) \leq p^{(1+\lceil (d-1)/(p-1) \rceil)\mathrm{rank}_d(P)}.$$

Proof:
Let $r := \mathrm{rank}_d(P)$. Then there is a function $\Gamma\colon \mathbb{T}^r \to \mathbb{T}$ and polynomials $Q_1, \ldots, Q_r$ of degree at most $d-1$ such that $\Gamma(Q_1(x), \ldots, Q_r(x)) = P(x)$. The Fourier expansion of Γ gives us a degree $d-1$ Fourier expansion of P, namely

$$\begin{aligned} e(\Gamma(Q_1(x), \ldots, Q_r(x))) &= \sum_{\alpha \in \widehat{G}} \widehat{\Gamma}(\alpha) e(Q_\alpha(x)) \\ &= e(P(x)), \end{aligned}$$

where $Q_\alpha(x) = \sum_{i=1}^r \alpha_i Q_i(x)$ and $\widehat{G} = \prod_{i=1}^r \mathbb{Z}_{p^{k_i+1}}$ is the dual of the group $G = \prod_{i=1}^r \frac{1}{p^{k_i+1}}\mathbb{Z}/\mathbb{Z}$ and k_i is the depth of Q_i. Since there are at most

$$|\widehat{G}| \leq p^{(1+\lceil (d-1)/(p-1) \rceil)r}$$

Fourier coefficients in the above expression, the result follows. □

4.3 Magic states in prime dimension

In this section we will give the explicit form of the magic state that we obtain by using the generalization of the T gate in odd prime dimensions [HV12]. We will then proceed to show that these magic states have exponentially small correlation with quadratic phase functions.

4.3.1 Generalization of the T gate

The generalization of the T gate, and hence the corresponding magic state, curiously enough depends on whether the prime dimension p is equal to three or $p > 3$.

- But first, let us consider the $p = 2$ case. In this case, the T-gate is given by $T = \left(\begin{smallmatrix} 1 & 0 \\ 0 & e^{i\pi/4} \end{smallmatrix}\right)$ and the corresponding single qubit magic state is given by $T|+\rangle = |T\rangle = \frac{|0\rangle + e^{i\pi/4}|1\rangle}{\sqrt{2}}$. Hence, the n-qubit magic state is

$$|T\rangle^{\otimes n} = \frac{1}{2^{n/2}} \sum_{x \in \mathbb{F}_2^n} e(|x|/8)|x\rangle. \tag{4.5}$$

  Here, the polynomial $P\colon \mathbb{F}_2^n \to \mathbb{T}\colon x \mapsto |x|/8$ is a nonclassical polynomial of degree three. This follows immediately from Proposition 1.4.8; see also Example 1.4.7. We will see a similar phenomenon in other prime dimensions.

- In the case that $p = 3$, let $\xi = e^{2\pi i/9}$ be a ninth-root of unity. The generalization of the T-gate for $p = 3$, denoted by U, is given by

$$U = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \xi & 0 \\ 0 & 0 & \xi^2 \end{pmatrix}.$$

  The corresponding single qutrit magic state is then

$$|\psi_U\rangle = U|+\rangle = \frac{|0\rangle + \xi|1\rangle + \xi^2|3\rangle}{\sqrt{3}}.$$

  The n-qutrit magic state in this case is

$$|\psi_U\rangle^{\otimes n} = \frac{1}{3^{n/2}} \sum_{x \in \mathbb{F}_3^n} e(|x|/9)|x\rangle. \tag{4.6}$$

  The polynomial $P\colon \mathbb{F}_3^n \to \mathbb{T}\colon x \mapsto |x|/9$ is again a nonclassical polynomial of degree three, which follows from Proposition 1.4.8.

- The case $p > 3$ are all similar, but different from the previous two cases. In this case, let $\omega = e^{2\pi i/p}$ be a p-th root of unity and $P\colon \mathbb{F}_p \to \mathbb{F}_p$ be a classical polynomial of degree three. Define

$$U := \sum_{x \in \mathbb{F}_p} \omega^{P(x)} |x\rangle\langle x|.$$

  This gate is then a non-clifford gate [HV12] and could serve as a generalization of the T gate. The condition that P has degree three is important: if

P has degree two, the gate U would be a Clifford gate. The corresponding n-qudit magic state is then

$$|\psi_U\rangle^{\otimes n} = \frac{1}{p^{n/2}} \sum_{x \in \mathbb{F}_p^n} \omega^{P_n(x)} |x\rangle, \tag{4.7}$$

where $P_n(x) = \sum_{i=1}^n P(x_i)$. Unlike the $p = 2, 3$ case, the polynomial P_n is a *classical* polynomial of degree three. Coincidentally, all cubic nonclassical polynomials for $p > 3$ are classical.

4.3.2 Correlation with quadratic phase functions

We will now show that for the n-qudit magic states defined as above, the correlation with quantum states whose amplitudes are given by a quadratic phase functions is exponentially small. We will give the proof for the $p = 2$ case, since the cases $p > 2$ are similar.

We need the following basic lemma, see [LMS08]. We will state and prove it here for convenience.

4.3.1. Lemma. *For any two functions $f, g \colon \mathbb{F}_p^n \to \mathbb{T}$, we have*

$$|\langle e(f), e(g)\rangle|^4 \leq \mathop{\mathbb{E}}_h |\langle e(\Delta_h f), e(\Delta_h g)\rangle|^2.$$

Proof:
By the Cauchy-Schwarz inequality,

$$\begin{aligned}
\sqrt{\mathop{\mathbb{E}}_h |\langle e(\Delta_h f), e(\Delta_h g)\rangle|^2} &\geq |\mathop{\mathbb{E}}_h \langle e(\Delta_h f), e(\Delta_h g)\rangle| \\
&= |\mathop{\mathbb{E}}_{x,h} e(f(x+h) - f(x) - (g(x+h) - g(x)))| \\
&= |\langle e(f), e(g)\rangle|^2.
\end{aligned}$$

□

In other words, we can compute the correlation between two phase functions by computing the correlations between their derivatives and taking their average. If the derivatives are easier to work with, this will become useful.

We will now show that the polynomial phase function $e(|x|/8)$, giving the amplitudes of the n-qubit magic state (see (4.5)), has exponentially small correlation with quadratic phase functions. First we compute the derivative of P; let $h \in \mathbb{F}_2^n$.

Then

$$\begin{aligned} P(x+h) - P(x) &= |x+h|/8 - |x|/8 = \sum_i |x_i + h_i|/8 - |x|/8 \\ &= \sum_i (|x_i| + |h_i| - 2|x_i||h_i|)/8 - |x|/8 \\ &= |h|/8 - |x \circ h|/4. \end{aligned}$$

Let $f\colon x \mapsto |x \circ h|/4$. We need the magnitudes of the Fourier coefficients of f. We have for $\alpha \in \mathbb{F}_2^n$,

$$\widehat{f}(\alpha) = \mathop{\mathbb{E}}_x e^{i\pi(|x\circ h| - 2|\alpha \circ x|)/2},$$

so

$$\begin{aligned} |\widehat{f}(\alpha)|^2 &= \mathop{\mathbb{E}}_{x,y} e^{i\pi(|x\circ h| - |y\circ h| - 2|\alpha\circ x| + 2|\alpha \circ y|)/2} \\ &= \mathop{\mathbb{E}}_{y,z} e^{i\pi(|y\circ h| + |z\circ h| - 2|y\circ z\circ h| - |y\circ h| - 2|\alpha\circ z|)/2} \\ &= \mathop{\mathbb{E}}_{y,z} e^{i\pi(|z\circ h| - 2|y\circ z\circ h| - 2|\alpha\circ z|)/2} \\ &= \mathop{\mathbb{E}}_z e^{i\pi(|z\circ h| - 2|\alpha\circ z|)/2} \mathop{\mathbb{E}}_y (-1)^{|y\circ z\circ h|}. \end{aligned}$$

The expectation over y is 0 unless $h_i = 1 \Rightarrow z_i = 0$ in which case it is equal to 1. So continuing where we left off

$$\begin{aligned} &= \frac{1}{2^n} \sum_{z : h_i = 1 \Rightarrow z_i = 0} e^{i\pi(|z\circ h| - 2|\alpha\circ z|)/2} \\ &= \frac{1}{2^n} \sum_{z : h_i = 1 \Rightarrow z_i = 0} (-1)^{|\alpha \circ z|} \end{aligned}$$

which is 0 unless $h_i = 0 \Rightarrow \alpha_i = 0$, in which case it equals $2^{-|h|}$. So the non-zero Fourier coefficients of f are at those $\alpha \in \mathbb{F}_2^n$ for which it holds that $h_i = 0 \Rightarrow \alpha_i = 0$.

We will now show that there is exponentially small correlation between $e(P)$ and any nonclassical polynomial phase function $e(Q)$ of degree two. By Lemma 4.3.1 and Parseval, we have that

$$\begin{aligned}|\langle e(P), e(Q)\rangle|^4 &\leq \mathop{\mathbb{E}}_h |\langle e(\Delta_h P), e(\Delta_h Q)\rangle|^2 \\ &= \mathop{\mathbb{E}}_h |\langle \widehat{e(\Delta_h P)}, \widehat{e(\Delta_h Q)}\rangle|^2 \\ &\leq \mathop{\mathbb{E}}_h 2^{-|h|} \\ &= \frac{1}{2^n}\sum_{k=0}^{n} 2^{-k}|\{h \in \mathbb{F}_2^n \colon |h| = k\}| \\ &= \left(\frac{3}{4}\right)^n.\end{aligned}$$

In the third line, we used that $\Delta_h Q$ is a degree-one classical polynomial, which means that $e(\Delta_h Q)$ has only one non-zero Fourier coefficient.

Similarly, for $p \geq 3$, the cubic phase function giving the amplitudes of the n-qudit magic states given by (4.6) for $p = 3$ and (4.7) for $p > 3$ have exponentially small correlation with any quadratic phase function.

4.3.2. Proposition. *Let $P \colon \mathbb{F}_p^n \to \mathbb{T}$ be the polynomial in the phase of the amplitudes in either Equation (4.5), (4.6) or (4.7). Then, for any nonclassical polynomial $Q \colon \mathbb{F}_p^n \to \mathbb{T}$ of degree at most two,*

$$|\langle e(P), e(Q)\rangle| \leq 2^{-cn},$$

for some $c > 0$ depending on p.

4.4 Stabilizer rank of the n-qudit magic state

Here we prove Theorem 4.1.1, our main result.

The following claim from [PSV21] is needed to get a handle on the affine subspaces that appear with the stabilizer states in a stabilizer decomposition.

4.4.1. Claim ([PSV21]). *Let p be a prime and $H_1, \ldots, H_r \subset \mathbb{F}_p^n$ be a collection of affine subspaces and assume $r \leq n/2$. Then there exists an affine subspace U of dimension at least $n - 2r$ and a subset $S \subset [r]$ such that for all $x \in U$*

$$1_{H_i}(x) = \begin{cases} 1 & \text{if } i \in S \\ 0 & \text{otherwise.} \end{cases}$$

Proof:
Let $E \colon \mathbb{F}_p^n \to \{0,1\}^r$ be the map given by $x \mapsto (1_{H_1}(x), \ldots, 1_{H_r}(x))$. By the pigeonhole principle, there is $\alpha \in \{0,1\}^r$ such that $|E^{-1}(\alpha)| \geq p^n 2^{-r} \geq p^{n-r}$. Denote by $S \subset [r]$ the set of indices i such that $\alpha_i = 1$. It is clear that

$E^{-1}(\alpha) = \cap_{i\in S} H_i \setminus \cup_{i\notin S} H_i$. Now define $V = \cap_{i\in S} H_i$, this is an affine subspace of dimension at least $n-r$. Pick an arbitrary $x_0 \in E^{-1}(\alpha)$, so $x_0 \notin H_i$ for $i \notin S$. This implies that $\forall i \notin S$ there is an affine equation h_i such that $h_i(x_0) = 1$ but $h_i(x) = 0$ for all $x \in H_i$. The affine subspace we are looking for is

$$U := \{x \in V : \forall i \notin S \quad h_i(x) = 1\}.$$

Note that U is not empty since $x_0 \in U$. Since we only add at most r extra equations, the dimension of U is at least $n - 2r$. □

Proof of Theorem 4.1.1 Let $P : \mathbb{F}_p^n \to \mathbb{T}$ be the nonclassical polynomial of degree three given by the corresponding n-qudit magic state, depending on the prime p. For $p = 2$, $p = 3$ and $p > 3$ it is given by the (polynomials in the phase of the) amplitudes of Equation (4.5), (4.6) and (4.7) respectively. Let $|\psi\rangle$ be the corresponding n-qudit magic state, i.e.

$$|\psi\rangle = p^{-n/2} \sum_{x\in\mathbb{F}_p^n} e(P(x))|x\rangle.$$

Denote by r the stabilizer rank of $|\psi\rangle$, so there is a decomposition

$$|\psi\rangle = \sum_{i=1}^{r} c_i |\phi_i\rangle,$$

for some constants c_i and stabilizer states $|\phi_i\rangle$. Each such ϕ_i is defined on an affine subspace $H_i \subset \mathbb{F}_p^n$. Let $C > 0$ be a large enough constant[1]. If $r > n/C$, we are done. So assume $r \leq n/C$. Then, we have that

$$\begin{aligned} e(P(x)) = p^{n/2}\langle x|\psi\rangle &= p^{n/2} \sum_{i=1}^{r} c_i \langle x|\phi_i\rangle \\ &= \sum_{i=1}^{r} c_i' e(Q_i(x)) 1_{H_i}(x), \end{aligned}$$

where each Q_i is a (nonclassical) quadratic polynomial on H_i and $c_i' = p^{(n-\dim(H_i))/2} c_i$. Then by Claim 4.4.1 (using $r \leq n/C$), there is an affine subspace U of dimension cn for some $c \geq 0.99$ and a non-empty subset $S \subset [r]$ such that $\forall x \in U$

$$1_{H_i}(x) = \begin{cases} 1 & \text{if } i \in S \\ 0 & \text{otherwise.} \end{cases}$$

[1] Given the prime p and the constant c from Proposition 4.3.2, the constant C should satisfy $C > 2p/c$.

Let $A\colon \mathbb{F}_p^{cn} \to \mathbb{F}_p^n$ be an affine linear map such that $U = \{A(y)\colon y \in \mathbb{F}_p^{cn}\}$. Let $P'\colon \mathbb{F}_p^{cn} \to \mathbb{T}$ be the polynomial given by $P'(y) = P(A(y))$. Then the restriction of the above decomposition of $e(P(x))$ to U implies that

$$e(P'(y)) = \sum_{i \in S} c_i' e(Q_i'(y)), \tag{4.8}$$

where $Q_i'(y) := Q_i(A(y))$ and we have that Q_i' is a nonclassical polynomial of degree at most two. By Propositions 4.3.2 and 4.2.6 the polynomial P has high rank: $\operatorname{rank}(P) \geq \Omega(n)$. By Lemma 4.2.7 it follows that P' is still cubic and we have $\operatorname{rank}(P') \geq \Omega(n)$ (this uses that C is a large enough constant). But (4.8) is a decomposition in terms of nonclassical polynomial phase functions of degree at most two. By Lemma 4.2.9 we have $|S| \geq \Omega(n)$ so that $r \geq \Omega(n)$.

4.5 Discussion

From the above proof, we can immediately conclude that it is not possible to get super-linear lower bounds on the stabilizer rank of n-qudit magic states. This is due to the use of Claim 4.4.1. However, there is no obvious obstruction to get super-linear lower bounds on the number of stabilizer states needed in a decomposition of n-qudit magic states where all the stabilizer states are defined on the full space $\mathbb{F}_p^n$. The graph states (classical quadratic polynomials) are for example in this set. The possibility of such a super-linear lower bound hinges on the relationship between the rank of a polynomial and its Fourier rank: the d-rank of a polynomial P on n variables is at most n, whereas the degree-d Fourier rank of a polynomial is at most p^{dn}. Lemma 4.2.9 only shows that $\operatorname{frank}_d(e(P)) \geq \operatorname{rank}_d(P)$. Is this relation optimal, or can we expect much better?

4.5.1. PROBLEM. *Let $d \geq 2$ and $P\colon \mathbb{F}_p^n \to \mathbb{T}$ be a polynomial. Is it true that*

$$\operatorname{frank}_d(e(P)) \geq \omega(\operatorname{rank}_d(P))?$$

A positive answer to this question would not only show that the n-qubit magic state needs super-linear many stabilizer states defined on the full space $\mathbb{F}_p^n$, but would also have implications in another field. For this, let us consider the case $p = 2$.

Let $\operatorname{AND}(x) = |x_1 x_2 \cdots x_n|/2$ be the (classical) polynomial giving the AND function. The "quadratic uncertainty principle" [FHH+14] is a conjecture that states that any decomposition

$$e(\operatorname{AND}(x)) = \sum_{i=1}^{r} c_i e(Q_i(x)), \tag{4.9}$$

where Q_i are classical quadratic polynomials, must have $r \geq 2^{\Omega(n)}$. Note that Definitions 4.2.5 and 4.2.8 of rank and frank are still valid if we only allow *classical* polynomials, which we will denote by rank' and frank'. The best known lower bound is $\mathrm{frank}'_3(e(\mathrm{AND})) \geq n/2$ [Wil18]. The proof of this uses the Chevalley-Warning theorem in an elegant way. It shows, implicitly, that the 3-rank' of the AND function is at least $n/2$, i.e. near maximal rank. This way of looking at it fits very well with Lemma 4.2.9. What is actually shown in [Wil18] is the same lower bound for the NOR function, i.e. $\mathrm{NOR}(x) = |1 + x_1| \cdots |1 + x_n|/2$. Lower bounds on the NOR function imply the same lower bounds on the AND function (it is the same function in a different basis). The proof also works if we allow polynomials of degree at most a constant.

4.5.2. Theorem ([Wil18], generalized). *Let* $\mathrm{NOR}\colon \mathbb{F}_2^n \to \mathbb{T}$ *be the function defined above and let* $d \geq 3$ *be an integer (constant). Then*

$$\mathrm{frank}'_d(e(\mathrm{NOR})) \geq n/(d-1).$$

Proof:
We will show that $\mathrm{rank}'_d(\mathrm{NOR}) \geq n/(d-1)$. Having shown this, the result follows immediately from Lemma 4.2.9 (which also holds for rank' and frank').

Let $r = \mathrm{rank}'_d(\mathrm{NOR})$. So there are (classical) polynomials $Q_1, \ldots, Q_r$ of degree at most $d-1$ such that there is a function $\Gamma\colon \mathbb{T}^r \to \mathbb{T}$ such that

$$\mathrm{NOR}(x) = \Gamma(Q_1(x), \ldots, Q_r(x)).$$

We may assume without loss of generality that $Q_i(0) = 0$ for all $i = 1, \ldots, r$. Assume that $r < n/(d-1)$. By the Chevalley-Warning theorem, the polynomials $Q_1, \ldots, Q_r$ have another common root $x \neq (0, \ldots, 0)$. This contradicts the fact that $\mathrm{NOR}(x) = 1/2$ if and only if $x = (0, \ldots, 0)$. □

Since the NOR function has $\mathrm{rank}'_3(\mathrm{NOR}) \geq n/2$, a positive answer to Problem 4.5.1 would show that the number of quadratic polynomials needed in (4.9) is super-linear. As noted before, Sanyal [San19] showed that $\mathrm{frank}'_2(e(P)) \geq \mathrm{rank}'_2(P)^2$. An analogue of this result for $d > 2$ would quadratically improve the best lower bound on $\mathrm{rank}'_3(\mathrm{NOR})$.

Chapter 5

High-entropy dual functions over finite fields

This chapter is based on the paper [BL21] which is joint work with Jop Briët.

5.1 Introduction

For $k \geq 2$, integer vector $\mathbf{i} = (i_1, \dots, i_k) \in \mathbb{Z}_{\geq 0}^k$ and finite abelian group G, the associated set of *order-k dual functions* is given by

$$\Delta_{\mathbf{i}} = \big\{\phi : y \mapsto \mathbb{E}_{x \in G} f_1(x + i_1 y) \cdots f_k(x + i_k y) : \ f_i : G \to \mathbb{D}\big\},$$

where $\mathbb{D}$ denotes the complex unit disc. For example, if $A \subseteq G$ is a subset, $\mathbf{i} = (0, 1, 2)$ and $f_i = \mathbb{1}_A$ for each $i \in [3]$, then $\phi(y)$ is the fraction of three-term arithmetic progressions in A with common difference y.

For applications in additive combinatorics and higher-order Fourier analysis, it is desirable to understand to what extent dual functions can be approximated by simpler functions. If $k = 2$, it follows from the Fourier inversion formula that one has the simple decomposition in terms of the characters:

$$\phi(y) = \sum_{\chi \in \widehat{G}} \alpha_\chi \chi\big((i_2 - i_1)y\big), \tag{5.1}$$

where $\|\alpha\|_{\ell_1} \leq 1$. Similar decompositions exist for higher-order dual functions thanks to inverse theorems for the Gowers uniformity norms. Inverse theorems roughly show that if f has large U^k-norm, then f correlates with a function $\psi : G \to \mathbb{D}$ akin to a polynomial of degree at most $k - 1$, see Section 1.4 for the precise statement of the Gowers inverse theorem when G is a vector space over a finite field. Here the "linear" ψ are precisely the characters. What exactly the "higher-order characters" are depends on the group G. For finite vector spaces $\mathbb{F}_p^n$ with $p \geq k$, they are the polynomial phase functions

$$\psi(x) = e^{2\pi i P(x)/p},$$

where $P \in \mathbb{F}_p[x_1, \dots, x_n]$ is a polynomial of degree at most $k-1$ [TZ10]. When $p < k$, one has to consider the larger class of non-classical polynomials [TZ12]. For the cyclic group $\mathbb{Z}_N$, they are the $(k-1)$-step nilsequences (of bounded complexity) [GTZ12]. Combined with the Hahn-Banach theorem, these inverse theorems imply that the decomposition (5.1) generalizes for larger k in terms of higher-order characters of degree at most $k-1$ up-to small L_1-error [Gow10]. Recall from Section 1.5 that, in the finite-field setting, this amounts to the following:

5.1.1. Proposition. *Let $p \geq k+1$ be a prime and let $G = \mathbb{F}_p^n$. Then, for any $\varepsilon > 0$ and $\mathbf{i} \in \mathbb{Z}_{\geq 0}^k$, there is an $M = M(\varepsilon, k, p) > 0$ such that any dual function $\phi \in \Delta_{\mathbf{i}}$ can be decomposed as*

$$\phi = \sum_{i=1}^{r} \alpha_i \psi_i + \tau, \tag{5.2}$$

where $\alpha_1, \dots, \alpha_r \in \mathbb{C}$ satisfy $|\alpha_1| + \cdots + |\alpha_r| \leq M$, $\psi_1, \dots, \psi_r$ are polynomial phases of degree at most $k-1$ and $\|\tau\|_{L_1} \leq \varepsilon$.

While a decomposition theorem like this (in particular over $\mathbb{Z}_N$) can be useful in higher-order Fourier analysis [Gow10], for other applications in additive combinatorics it is preferable to have more precise control over the error function τ in (5.2). A natural finite-field analog of a question raised by Frantzikinakis in [Fra16, Problem 1] (see also [Alt20]) asks if this error function can be bounded *everywhere*, that is, if Proposition 5.1.1 still holds with $\|\tau\|_{L_\infty} \leq \varepsilon$. The apparent expectation of a positive answer to Frantzikinakis's question motivated conjectures on a poorly-understood probabilistic variant of Szemerédi's theorem on arithmetic progressions (cf. Section 5.1.1). Our main result, however, shows that in the finite-field setting, the answer is negative.

5.1.2. Theorem. *For infinitely many primes p, there is a $k = k(p) \in \mathbb{N}$ and an integer vector $\mathbf{i} \in \mathbb{Z}_{\geq 0}^k$ such that (5.2) cannot hold with $\|\tau\|_{L_\infty} \leq \varepsilon$.*

Special cases of Theorem 5.1.2 show that for $k = 3$ and $p = 2^t - 1$ a Mersenne prime, the decomposition (5.2) requires polynomial phases of degree at least t for fixed ε, M and $\|\tau\|_{L_\infty} \leq \varepsilon$. The largest known Mersenne prime as of January 2018 has $t = 77,232,917$ [GIM].

5.1.1 Locally decodable codes and random Szemerédi

The examples behind Theorem 5.1.2 originate from constructions of special types of error-correcting codes called *locally decodable codes* (LDCs). These codes have the property that any single encoded message symbol can be retrieved from a codeword with good probability by reading only a tiny number of codeword symbols, even if the codeword is partially corrupted. LDCs originated in complexity

theory [BK95, AS98, ALM+98] and cryptography [CGKS98] and were defined in the context of channel coding in [KT00]. They have since found many other applications in computer science and mathematics, for instance in fault tolerant distributed storage systems [GHSY12] and Banach space geometry [BNR12]. We refer to [Yek12, Gop18] for extensive surveys.

Despite their ubiquity, LDCs are poorly understood. Of particular interest is the tradeoff between the codeword length N as a function of message length k when the *query complexity*—the number of probed codeword symbols—and alphabet size are constant. The Hadamard code is a 2-query LDC of length $N = 2^{O(k)}$ and this length is optimal in the 2-query regime [KdW04]. For $q \geq 3$, the best-known lower bounds show that any q-query LDC has at least polynomial length $k^{1+1/(\lceil q/2 \rceil - 1) - o(1)}$ [KdW04, Woo07]. The family of Reed-Muller codes, which generalize the Hadamard code, were for a long time the best-known examples, giving q-query LDCs of length $\exp(O(k^{1/(q-1)}))$.

In a breakthrough result, Yekhanin [Yek08] constructed an entirely new family of vastly shorter LDCs. For each Mersenne prime $p = 2^t - 1$, he gave a 3-query LDC of length $N \leq \exp(O(k^{1/t}))$. The construction uses a family of k homomorphisms from $\mathbb{F}_p^n$ to the multiplicative subgroup of $\mathbb{F}_{2^t}$. The homomorphisms are constructed using a family of *matching vectors* $(u_i, v_i)_{i \in [k]}$, which are pairs of orthogonal vectors in $\mathbb{F}_p^n$ such that the inner products $\langle u_i, v_j \rangle$ with $i \neq j$ belong to a special subset of $\mathbb{F}_p^*$. It is this construction that forms the basis for Theorem 5.1.2.

Subsequently, Efremenko [Efr12] constructed much larger matching vector families over $\mathbb{Z}_m^n$ for composite moduli m and used Yekhanin's framework to give the first 3-query LDCs of subexponential length $N \leq \exp(\exp(O\sqrt{\log k \log \log k}))$. But huge gaps persist between the best-known upper and lower bounds for constant-query LDCs.

In contrast with other combinatorial objects such as expander graphs, the probabilistic method has so far not been successfully used to beat the best explicit LDC constructions. In [BDG19], a probabilistic framework was given that could in principle yield best-possible LDCs, albeit non-constructively. A special instance of this framework connects LDCs with a probabilistic version of Szemerédi's theorem alluded to above. The setup for this is as follows:

For a finite abelian group G of size $N = |G|$, let $D \subseteq G$ be a random subset where each element is present with probability ρ independently of all others. For $k \geq 3$ and $\varepsilon \in (0, 1)$, let E be the event that every subset $A \subseteq G$ of size $|A| \geq \varepsilon |G|$ contains a proper k-term arithmetic progression with common difference in D. If $\rho = 1$, then it follows from the Density Hales–Jewett Theorem [FK91] that E holds with probability 1 provided N is large enough in terms of k and ε. It is an open problem to determine the smallest value of ρ — which we will denote by ρ_k — such that $\Pr[E] \geq \frac{1}{2}$. This value will depend on ε too, but we will suppress this in the notation and assume ε is a fixed constant. It is also assumed that N is large enough so that ρ_k exists.

In [BDG19] it is shown that there exist k-query LDCs of message length $\Omega(\rho_k N)$ and codeword length $O(N)$. As such, Szemerédi's theorem with random differences, in particular lower bounds on ρ_k, can be used to show the existence of LDCs. Conversely, this connection indirectly implies the best-known upper bounds on ρ_k for all $k \geq 3$, given by $N^{-(1-o(1))/\lceil k/2 \rceil}$ [FLW12, BG18]. However, a conjecture of Frantzikinakis, Lesigne and Wierdl [FLW16] states that over $\mathbb{Z}_N$ we have $\rho_k \ll_k N^{-1} \log N$ for all k, which would be best-possible. Truth of this conjecture would imply that over this group, Szemerédi's theorem with random differences cannot give LDCs better than the Hadamard code. For finite fields, Altman [Alt20] showed that this conjecture is false. In particular, over $\mathbb{F}_p^n$ for p odd, he proved that $\rho_3 \geq \Omega(p^{-n} n^2)$; generally, $\rho_k \geq \Omega(p^{-n} n^{k-1})$ holds when $p \geq k+1$ [Bri20]. In turn, these bounds are conjectured to be optimal for the finite-field setting, which would imply that over finite fields, Szemerédi's theorem with random differences cannot give LDCs better than Reed-Muller codes.

These conjectures appear to be motivated mainly by the possibility of an L_∞-version of Proposition 5.1.1 (and analogous variants over $\mathbb{Z}_N$) with dual functions based on 3-term progressions. Theorem 5.1.2 falls short of obstructing this route to obtaining optimal bounds in the finite-field setting for two reasons. First, our examples do not include "arithmetic-progression dual functions," those with $\mathbf{i} = (0, 1, \ldots, (k-1))$; in fact in the Appendix we show that our current framework cannot give such examples. Second, even if we had such examples, they do not appear to imply any new lower bounds on ρ_k. Nevertheless, we do not expect arithmetic progressions to be exceptional patterns for which there are no such examples.

5.1.3. Remark. Ideas behind Theorem 5.1.2 recently inspired similar examples in the integer setting for 3-term progressions [BG20].

5.2 Preliminaries

We will identify the set of maps $G \to \mathbb{C}$ with $\mathbb{C}^G$. If X and Y are quantities depending on some underlying variable $n \in \mathbb{N}$ and $\alpha_1, \ldots, \alpha_k$ are parameters, we then write $X = O_{\alpha_1,\ldots,\alpha_k}(Y)$ if $X \leq C_{\alpha_1,\ldots,\alpha_k} Y$ for all n large enough, where $C_{\alpha_1,\ldots,\alpha_k}$ is a constant depending only on the parameters $\alpha_1, \ldots, \alpha_k$. Similarly, $X = \Omega_{\alpha_1,\ldots,\alpha_k}(Y)$ means that $X \geq C_{\alpha_1,\ldots,\alpha_k} Y$ for all n large enough, where $C_{\alpha_1,\ldots,\alpha_k}$ is a constant depending only on the parameters $\alpha_1, \ldots, \alpha_k$.

For a polynomial $P(x) = \sum_{\iota=0}^{t} c_\iota x^\iota$, define its *support* $\mathbf{i}(P)$ to be the sequence of degrees $\iota \in \mathbb{Z}_{\geq 0}$ such that $c_\iota \neq 0$, arranged in increasing order. The *support size* is the length of $\mathbf{i}(P)$. We will use some basic facts from the theory of finite fields, for which we refer to [LN97]. The Minkowski sum of two sets $A, B \subseteq \mathbb{C}^n$ is the set given by

$$A + B = \{a + b : a \in A, b \in B\}.$$

We will use the following slight generalization of the notion of the convex hull, where we allow for complex coefficients. For a compact set $A \subseteq \mathbb{C}^n$, define

$$\mathrm{Conv}_{\mathbb{C}}(A) = \Big\{ \sum_{a \in A} \alpha_a a : \alpha_a \in \mathbb{C} \quad \forall a \in A, \quad \sum_{a \in A} |\alpha_a| \leq 1 \Big\}.$$

For a finite set $A \subseteq \mathbb{D}^n$ and $\varepsilon, M \in (0, \infty)$, define $\mathcal{N}(A, \varepsilon, M)$ to be the smallest size of a finite set $B \subseteq M\mathbb{D}^n$ such that

$$A \subseteq \mathrm{Conv}_{\mathbb{C}}(B) + \varepsilon \mathbb{D}^n.$$

Then, for any $a \in A$, there is a $b \in \mathrm{Conv}_{\mathbb{C}}(B)$ such that $\|a - b\|_{\ell_\infty} \leq \varepsilon$ and so $\mathcal{N}(A, \varepsilon, M)$ is a restricted form of the covering number of A relative to the ℓ_∞ distance. Note that for $I \subseteq [n]$, the projection of A to the set of coordinates I, given by $A_I = \{(a_i)_{i \in I} : a \in A\}$, is contained in $\mathrm{Conv}_{\mathbb{C}}(B_I) + \varepsilon \mathbb{D}^I$. Since $|B| \geq |B_I|$, it follows that

$$\mathcal{N}(A, \varepsilon, M) \geq \mathcal{N}(A_I, \varepsilon, M). \tag{5.3}$$

5.3 Covering numbers from hypercubes

We will use the following lemma, which shows that containment of a high-dimensional hypercube implies a large restricted covering number.

5.3.1. Lemma. *Let $c > 0$, $z \in \mathbb{C}$ be a complex number such that $\Re(z) \leq 0$ and let $S \subseteq \mathbb{C}^k$ be a finite set such that $\{c, z\}^k \subseteq S$. Then, for any $\varepsilon \in (0, \frac{c}{2})$ and $M > 0$, we have that*

$$\log_2 \big(\mathcal{N}(S, \varepsilon, M) \big) \geq \Omega_{c, \varepsilon, M}(k).$$

Proof:
Let θ be a uniformly distributed $\{-1, 1\}^k$-valued random vector. For a compact set $A \subseteq \mathbb{C}^k$, define

$$w(A) = \mathbb{E} \max_{a \in A} |\langle a, \theta \rangle|.$$

We use the following basic properties:

1. If $A \subseteq B$, then $w(A) \leq w(B)$.
2. For a finite set $A \subseteq \mathbb{C}^k$, it holds that $w(\mathrm{Conv}_{\mathbb{C}}(A)) = w(A)$.
3. For $A, B \subseteq \mathbb{C}^k$ finite, it holds that $w(A + B) \leq w(A) + w(B)$.

It follows from the first property that

$$w(S) \geq w(\{c,z\}^k) \geq \frac{ck}{2}. \tag{5.4}$$

For the second inequality, observe that for fixed $\theta \in \{-1,1\}^k$, we have

$$\begin{aligned}\max_{a\in\{c,z\}^k} |\langle a,\theta\rangle| &\geq \Big|\sum_{i:\theta_i=1} c - \sum_{i:\theta_i=-1} z\Big| \\ &\geq \Big|\Re\Big(\sum_{i:\theta_i=1} c - \sum_{i:\theta_i=-1} z\Big)\Big| \\ &\geq c|\{i\in[k]:\,\theta_i=1\}|.\end{aligned}$$

Averaging over θ then gives the result.

Let $B \subseteq M\mathbb{D}^k$ be a finite set such that $S \subseteq \mathrm{Conv}_{\mathbb{C}}(B) + \varepsilon\mathbb{D}^k$. Let $l = |B|$ and $p = \log_2 l$. By the second property of w, Jensen's inequality and the Khintchine inequality [MS86, Chapter 5],

$$\begin{aligned}w\big(\mathrm{Conv}_{\mathbb{C}}(B)\big) &= \mathbb{E}\max_{b\in B}|\langle b,\theta\rangle| \\ &\leq \mathbb{E}\Big(\sum_{b\in B}|\langle b,\theta\rangle|^p\Big)^{\frac{1}{p}} \\ &\leq \Big(\sum_{b\in B}\mathbb{E}|\langle b,\theta\rangle|^p\Big)^{\frac{1}{p}} \\ &\leq C\sqrt{p}\Big(|B|\max\{\|b\|_{\ell_2}^p : b\in B\}\Big)^{\frac{1}{p}} \\ &\leq C'M\sqrt{k\log l}.\end{aligned}$$

For some constants C, C'. We also have $w(\varepsilon\mathbb{D}^k) = \varepsilon k$. Since $S \subseteq \mathrm{Conv}_{\mathbb{C}}(B) + \varepsilon\mathbb{D}^k$, the second and third properties of w and (5.4) then give

$$\frac{ck}{2} \leq w(S) \leq w(\mathrm{Conv}_{\mathbb{C}}(B)) + \varepsilon\mathbb{D}^k) \leq O(M\sqrt{k\log_2 l} + \varepsilon k).$$

Rearranging the left- and right-hand sides now gives the claim. □

5.4 Locating high-dimensional hypercubes

Here we show that for certain primes p and some integer vectors $\mathbf{i}$, the dual functions in $\Delta_{\mathbf{i}}$ over $\mathbb{F}_p^n$ contain high-dimensional hypercubes.

5.4.1. PROPOSITION. *Let p, r be distinct primes, let $t = \mathrm{ord}_p(r)$ and let $G = \mathbb{F}_p^n$. Suppose there exists a polynomial $P(x) \in \mathbb{F}_r[x]$ that has a root in $\mathbb{F}_{r^t}^*$ of order p and such that $P(1) \neq 0$. Then, there exists a $z \in \mathbb{C}$ with $\Re(z) \leq 0$ and a set $D \subseteq G$ of size $|D| \geq \Omega_p(n^t)$ such that*

$$\{z, 1\}^D \subseteq \Delta^D_{\mathbf{i}(P)}.$$

The proof of this proposition relies on the following result due to Yekhanin, which is implicit in [Yek08] (and shown explicitly in [Rag07]). We include a proof for completeness.

5.4.2. THEOREM (Yekhanin). *Let p, r be distinct primes and $t := \mathrm{ord}_p(r)$. For integer $m > p - 1$, let*

$$k = \binom{m}{p-1} \qquad \text{and} \qquad n = \binom{m + \frac{p-1}{t} - 1}{\frac{p-1}{t}}.$$

Let

$$P(x) = \sum_{\iota=0}^{s} c_\iota x^\iota \in \mathbb{F}_r[x]$$

be a polynomial with a root $\gamma \in \mathbb{F}_{r^t}^*$ of order p. Then, for each $i \in [k]$ there exists a function $f_i \colon \mathbb{F}_p^n \to \mathbb{F}_{r^t}$ and vectors $d_i, w_i \in \mathbb{F}_p^n$ such that for every $x \in \mathbb{F}_p^n$, we have

$$\sum_{\iota=0}^{s} c_\iota f_i(x + \iota d_j) = \begin{cases} \gamma^{\langle x, w_i \rangle} P(1) & \text{if } i = j \\ 0 & \text{otherwise.} \end{cases}$$

Proof:
For a $(p-1)$-element subset $S \subseteq [m]$, define the vectors $u_S = 1_S$ and $v_S = 1_{[m]} - u_S$ in $\mathbb{F}_p^m$. Then, $\langle u_S, v_T \rangle = 0$ if and only if $S = T$. Let $l = \frac{p-1}{t}$. Then, for $a \in \mathbb{F}_p^*$, we have $a^l \in \{r^q : q = 0, 1, \ldots, p-1\}$.

Consider the expansion of the polynomial $Q(x) \in \mathbb{F}_p[x_1, \ldots, x_m]$ given by

$$Q(x) = (x_1 + \cdots + x_m)^l = \sum_{\beta \in \mathcal{M}_l} c_\beta x^\beta,$$

where $\mathcal{M}_l := \{\beta \in \mathbb{Z}^m_{\geq 0} : \sum_{i=1}^m \beta_i = l\}$ and $x^\beta := \prod_{i=1}^m x_i^{\beta_i}$. For each subset $S \subseteq [m]$ of size $p - 1$, define the vectors $w_S = (u_S^\beta)_{\beta \in \mathcal{M}_l}$ and $d_S = (c_\beta v_S^\beta)_{\beta \in \mathcal{M}_l}$. Since $x^\beta y^\beta = (x \circ y)^\beta$, where $\circ$ denotes the coordinate-wise product, we have that

$$\langle w_S, d_T \rangle = Q(u_S \circ v_T) = \langle u_S, v_T \rangle^l.$$

By the above, this equals zero if $S = T$ and a power of r otherwise. Moreover, the vectors w_S and d_S have dimension $|\mathcal{M}_l| = \binom{m+l-1}{l}$.

Define $f_S : \mathbb{F}_p^n \to \mathbb{F}_{r^t}^*$ by

$$f_S(x) = \gamma^{\langle x, w_S \rangle}.$$

Note that this a homomorphism, because γ has order p. Then,

$$\begin{aligned}\sum_{\iota=0}^{s} c_\iota f_S(x + \iota d_S) &= \gamma^{\langle x, w_S \rangle} \sum_{\iota=0}^{s} c_\iota \gamma^{\iota \langle d_S, w_S \rangle} \\ &= \gamma^{\langle x, w_S \rangle} \sum_{\iota=0}^{s} c_\iota \\ &= \gamma^{\langle x, w_S \rangle} P(1).\end{aligned}$$

If $S \neq T$, then $\langle d_T, w_S \rangle = r^q \bmod p$ for some integer q and therefore,

$$\begin{aligned}\sum_{\iota=0}^{s} c_\iota f_S(x + \iota d_T) &= \gamma^{\langle x, w_S \rangle} \sum_{\iota=0}^{s} c_\iota \gamma^{\iota \langle d_T, w_S \rangle} \\ &= \gamma^{\langle x, w_S \rangle} \sum_{\iota=0}^{s} c_\iota \gamma^{\iota r^q} \\ &= \gamma^{\langle x, w_S \rangle} P(\gamma)^{r^q} \\ &= 0.\end{aligned}$$

This completes the proof. □

Proof of Proposition 5.4.1:
Let $P(x) \in \mathbb{F}_r[x]$ be as in Proposition 5.4.1 and let $\gamma \in \mathbb{F}_{r^t}^*$ be a p-th root of unity such that $P(\gamma) = 0$. Let $f_i : \mathbb{F}_p^n \to \mathbb{F}_{r^t}^*$ and $d_i, w_i \in \mathbb{F}_p^n$ be as in Theorem 5.4.2. Let $\chi : \mathbb{F}_{r^t} \to \mathbb{C}$ be a nontrivial additive character such that the complex number

$$z := \mathbb{E}_{c \in \mathbb{F}_p} \chi\big(\gamma^c P(1)\big)$$

satisfies $\Re(z) \leq 0$. To see that such a character exists, observe that by orthogonality of the characters,

$$\mathbb{E}_{\chi \in \widehat{\mathbb{F}_{r^t}}} \mathbb{E}_{c \in \mathbb{F}_p} \chi\big(\gamma^c P(1)\big) = \mathbb{E}_{c \in \mathbb{F}_p} \Big(\mathbb{E}_{\chi \in \widehat{\mathbb{F}_{r^t}}} \chi\big(\gamma^c P(1)\big) \Big) = 0.$$

The existence of the desired character then follows by averaging. For each $a \in \{0,1\}^k$ and $\iota \in \mathbf{i}(P)$, define $F_a^\iota : \mathbb{F}_p^n \to \mathbb{C}$ by

$$F_a^\iota(x) = \chi\Big(c_\iota \sum_{j=1}^{k} a_j f_j(x) \Big). \tag{5.5}$$

Based on these functions, we define the dual function $\phi_a : \mathbb{F}_p^n \to \mathbb{D}$ by

$$\phi_a(y) = \mathbb{E}_{x \in \mathbb{F}_p^n} \prod_{\iota \in \mathbf{i}(P)} F_a^\iota(x + \iota y). \tag{5.6}$$

Then,

$$\begin{aligned}\phi_a(d_i) &= \mathbb{E}_{x\in\mathbb{F}_p^n}\chi\Big(\sum_{j=1}^{k} a_j \sum_{\iota\in\mathbf{i}(P)} c_\iota f_j(x+\iota d_i)\Big)\\ &= \mathbb{E}_{x\in\mathbb{F}_p^n}\chi(a_i\gamma^{\langle x,w_i\rangle}P(1))\\ &= \mathbb{E}_{c\in\mathbb{F}_p}\chi(a_i\gamma^c P(1)).\end{aligned}$$

The last expectation equals 1 if $a_i = 0$ and z if $a_i = 1$ and therefore,

$$\{1,z\}^k \subseteq \{(\phi(d))_{d\in D} : \phi \in \Delta_{\mathbf{i}(P)}\}.$$

Since $k \geq (\frac{m}{p})^{p-1}$, $n \leq (\frac{2etm}{p})^{\frac{p-1}{t}}$ and $t \leq p-1$, we have $k \geq \Omega_p(n^t)$. □

5.5 Sparse polynomials over $\mathbb{F}_2$

The following lemma supplies infinitely many primes and polynomials that can be used in Proposition 5.4.1 .

5.5.1. Lemma. *For infinitely many primes p, there is an irreducible polynomial $P(x) \in \mathbb{F}_2[x]$ with support size at most $t = \mathrm{ord}_p(2)$ and a root in $\mathbb{F}_{2^t}^*$ of order p.*

To prove Lemma 5.5.1, we use some basic theory of cyclotomic polynomials (see for example [LN97, Chapter 2]). Let r be a prime and $n \in \mathbb{N}$ not divisible by r. Recall that a primitive n-th root of unity over $\mathbb{F}_r$ is a generator of the non-zero elements of the splitting field of the polynomial $x^n - 1$ over $\mathbb{F}_r$. Then, for any such root of unity ζ, the n-th cyclotomic polynomial is given by

$$\Phi_n(x) = \prod_{\gcd(s,n)=1} (x - \zeta^s),$$

where the product is over $s \in \{1,\ldots,n\}$ such that $\gcd(s,n) = 1$. The following lemma gives the properties of cyclotomic polynomials we need.

5.5.2. Lemma. *Let r be a prime, $n \in \mathbb{N}$ not divisible by r. Then, the coefficients of $\Phi_n(x)$ lie in $\mathbb{F}_r$. Moreover, if n is a prime, then $\Phi_n(x)$ factors into $(n-1)/\mathrm{ord}_n(r)$ distinct monic irreducible polynomials all of which have degree exactly $\mathrm{ord}_n(r)$.*

For an integer $k \geq 2$, denote by $p(k)$ the largest prime number that divides k. We will use the following result of Stewart [Ste13].

5.5.3. Lemma (Stewart). *For all n large enough, we have*

$$p(2^n - 1) > n \exp\left(\frac{\log n}{104 \log\log n}\right).$$

Proof of Lemma 5.5.1:
By Lemma 5.5.3, for $p = p(2^n - 1)$ and n sufficiently large, we have that $\mathrm{ord}_p(2) \leq n < (p-1)/2$. Hence, there are infinitely many primes p such that $t := \mathrm{ord}_p(2) \leq (p-1)/2$. For such a p, consider the p-th cyclotomic polynomial $\Phi_p(x)$ over $\mathbb{F}_2$. By Lemma 5.5.2, $\Phi_p(x)$ factors into $(p-1)/t$ distinct monic irreducible polynomials over $\mathbb{F}_2$ of degree exactly t. Since over $\mathbb{F}_2$, there is only one polynomial of degree t with support size $t+1$, there must be an irreducible factor with support of size at most t. Let $P(x)$ be such a factor. Then, since $P(x)|\Phi_p(x)$, its roots lie in the set of p-th roots of unity in $\mathbb{F}_{2^t}$. □

5.5.4. Remark. For Mersenne primes $p = 2^t - 1$, there are polynomials over $\mathbb{F}_2$ with support size 3 that meet the conditions of Proposition 5.4.1. Indeed, since in this case, any p-th root of unity ζ in $\mathbb{F}_{2^t}$ is a generator of $\mathbb{F}_{2^t}^*$ and since $1 + \zeta \neq 0$, there exists an s such that $P(x) = 1 + x + x^s$ satisfies $P(1) = 1$ and $P(\zeta) = 0$.

5.6 Proof of Theorem 5.1.2

Let $p, t, P(x)$ be as in Lemma 5.5.1, so that P has support size $k \leq t$. Let $\mathbf{i} = \mathbf{i}(P)$. Since P is irreducible, $P(1) \neq 0$ and so it satisfies the conditions of Proposition 5.4.1. Fix $\varepsilon \in (0, \frac{1}{2})$ and $M \in (0, \infty)$. Suppose that Proposition 5.1.1 held with $\|\tau\|_{L_\infty} \leq \varepsilon$, which is to say that

$$\Delta_{\mathbf{i}} \subseteq \mathrm{Conv}_{\mathbb{C}}\left(M \cdot \{\text{polynomial phases of degree} \leq k-1\}\right) + \varepsilon \mathbb{D}^{\mathbb{F}_p^n}.$$

Then, since there are at most $p^{O(n^{k-1})}$ polynomial phase functions of degree at most $k-1$ (one for each n-variate polynomial of degree at most $k-1$), this implies that

$$\log_2 \mathcal{N}(\Delta_{\mathbf{i}}, \varepsilon, M) \leq O_p(n^{k-1}) \leq O_p(n^{t-1}). \tag{5.7}$$

At the same time, Proposition 5.4.1, Lemma 5.3.1 and property (5.3) give

$$\log_2 \mathcal{N}(\Delta_{\mathbf{i}}, \varepsilon, M) \geq \Omega_{p,\varepsilon,M}(n^t).$$

This contradicts (5.7) for large n and finishes the proof of Theorem 5.1.2.

5.7 On the possible arithmetic patterns

Here we show that our construction cannot give examples for dual functions corresponding to arithmetic progressions. Let p, r be primes and $t = \mathrm{ord}_p(r)$. Suppose that for some $k, s \in \mathbb{N}$, there is a polynomial $P(x) \in \mathbb{F}_r[x]$ of the form

$$P(x) = \sum_{\iota=0}^{k-1} c_\iota x^{\iota s}$$

such that $P(1) \neq 0$ and $P(x)$ has a root in $\mathbb{F}_{r^t}^*$ of order p. Then, the functions defined as in (5.5) and (5.6) belong to the set of dual functions corresponding to the progression $\mathbf{i} = (0, s, 2s, \ldots, (k-1)s)$ and generate in a hypercube of dimension at least n^t. We show that $k \geq t + 1$, which means that this does not contradict an L_∞-version of Proposition 5.1.1.

First note that s cannot be a multiple of p, since for any $\gamma \in \mathbb{F}_{r^t}^*$ of order p we would have $\gamma^s = 1$, which implies that $P(\gamma) = P(1) \neq 0$. It follows that for any such γ, the element γ^s also has order p and does not equal 1. Define the polynomial

$$Q(x) = \sum_{\iota=0}^{k-1} c_\iota x^\iota \in \mathbb{F}_r[x].$$

Then, this polynomial has a root α in $\mathbb{F}_{r^t}^*$ of order p (where $\alpha = \gamma^s$), satisfies $Q(1) = P(1) \neq 0$ and has degree $k - 1$. We claim that $k - 1 \geq \mathrm{ord}_p(r)$. If Q is reducible, then it has a factor of degree strictly less than $k - 1$ that has the same properties. So assume that Q is irreducible. Let $K = \mathbb{F}_r(\alpha)$ be the simple algebraic extension of $\mathbb{F}_r$ obtained by adjoining α. Then K is isomorphic to $\mathbb{F}_{r^{k-1}}$. Since α lies in $\mathbb{F}_{r^{k-1}}$ and has order p, it follows that $p \mid r^{k-1} - 1$. But this implies that $k - 1 \geq \mathrm{ord}_p(r) = t$.

Bibliography

[AC88] N. Alon and F. R. Chung. Explicit construction of linear sized tolerant networks. *Discrete Mathematics*, 72(1-3):15–19, 1988.

[AGR82] A. Aspect, P. Grangier, and G. Roger. Experimental realization of Einstein-Podolsky-Rosen-Bohm gedankenexperiment: A new violation of Bell's inequalities. *Phys. Rev. Lett.*, 49:91–94, Jul 1982. doi: 10.1103/PhysRevLett.49.91.

[ALM+98] S. Arora, C. Lund, R. Motwani, M. Sudan, and M. Szegedy. Proof verification and the hardness of approximation problems. *J. ACM*, 45(3):501–555, 1998. doi: 10.1145/278298.278306.

[Alt20] D. Altman. On Szemerédi's theorem with differences from a random set. *Acta Arith*, 195:97–108, 2020. doi: 10.4064/aa190531-25-10.

[AS98] S. Arora and S. Safra. Probabilistic checking of proofs: a new characterization of NP. *J. ACM*, 45(1):70–122, 1998. doi: 10.1145/273865.273901.

[AS04a] N. Alon and J. H. Spencer. *The probabilistic method*. John Wiley & Sons, 2004.

[AS04b] A. Ambainis and A. Smith. Small pseudo-random families of matrices: Derandomizing approximate quantum encryption. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*, pages 249–260. Springer, 2004. doi: 10.1007/978-3-540-27821-4_23.

[Aub09] G. Aubrun. On almost randomizing channels with a short Kraus decomposition. *Comm. Math. Phys.*, 288(3):1103–1116, 2009. ISSN 0010-3616. doi: 10.1007/s00220-008-0695-y.

[BBB+19] T. Bannink, J. Briët, H. Buhrman, F. Labib, and T. Lee. Bounding quantum-classical separations for classes of nonlocal games. In *36th International Symposium on Theoretical Aspects of Computer Science (STACS 2019)*, 126, pages 12:1–12:11. March 2019. doi: 10.4230/LIPIcs.STACS.2019.12.

[BBC+01] R. Beals, H. Buhrman, R. Cleve, M. Mosca, and R. d. Wolf. Quantum lower bounds by polynomials. *Journal of the ACM*, 48(4):778–797, 2001. Earlier version in FOCS'98.

[BBC+19] S. Bravyi, D. Browne, P. Calpin, E. Campbell, D. Gosset, and M. Howard. Simulation of quantum circuits by low-rank stabilizer decompositions. *Quantum*, 3:181, 2019.

[BBLM20] T. Bannink, J. Briët, F. Labib, and H. Maassen. Quasirandom quantum channels. *Quantum*, 4:298, 2020. doi: 10.22331/q-2020-07-16-298.

[BBLV13] J. Briët, H. Buhrman, T. Lee, and T. Vidick. Multipartite entanglement in XOR games. *Quantum Information & Computation*, 13(3-4):334–360, 2013.

[BCL+06] C. Borgs, J. Chayes, L. Lovász, V. T. Sós, and K. Vesztergombi. Counting graph homomorphisms. In *Topics in discrete mathematics*, pages 315–371. Springer, 2006.

[BDG19] J. Briët, Z. Dvir, and S. Gopi. Outlaw distributions and locally decodable codes. *Theory Comput.*, 15(12):1–24, 2019. doi: 10.4086/toc.2019.v015a012. Preliminary version in ITCS'17.

[Bel64] J. Bell. On the Einstein Podolsky Rosen paradox. *Physics*, 1(3):195–200, 1964.

[BG16] S. Bravyi and D. Gosset. Improved classical simulation of quantum circuits dominated by Clifford gates. *Physical review letters*, 116(25):250501, 2016.

[BG18] J. Briët and S. Gopi. Gaussian width bounds with applications to arithmetic progressions in random settings. *Int. Math. Res. Not.*, page rny238, 2018. doi: 10.1093/imrn/rny238.

[BG20] J. Briët and B. Green. Multiple correlation sequences not approximable by nilsequences, 2020. Preprint. Available at arXiv:2010.14960.

[BK95] M. Blum and S. Kannan. Designing programs that check their work. *J. ACM*, 42(1):269–291, 1995. doi: 10.1145/200836.200880.

[BK05] S. Bravyi and A. Kitaev. Universal quantum computation with ideal Clifford gates and noisy ancillas. *Physical Review A*, 71(2):022316, 2005.

[BL96] V. Bergelson and A. Leibman. Polynomial extensions of van der Waerden's and Szemerédi's theorems. *J. Amer. Math. Soc.*, 9(3):725–753, 1996.

[BL21] J. Briët and F. Labib. High-entropy dual functions over finite fields and locally decodable codes. *Forum of Mathematics, Sigma*, 9:e19, 2021. doi: 10.1017/fms.2021.1.

[BMMN13] M. Braverman, K. Makarychev, Y. Makarychev, and A. Naor. The Grothendieck constant is strictly smaller than Krivine's bound. *Forum Math. Pi*, 1:453–462, 2013. doi: 10.1017/fmp.2013.4. Preliminary version in FOCS'11. arXiv: 1103.6161.

[BN04] B. Bollobás and V. Nikiforov. Hermitian matrices and graphs: singular values and discrepancy. *Discrete Math.*, 285(1-3):17–32, 2004. ISSN 0012-365X. doi: 10.1016/j.disc.2004.05.006.

[BNR12] J. Briët, A. Naor, and O. Regev. Locally decodable codes and the failure of cotype for projective tensor products. *Electron. Res. Announc. Math. Sci.*, 19:120–130, 2012. doi: 10.3934/era.2012.19.120.

[BOGKW88] M. Ben-Or, S. Goldwasser, J. Kilian, and A. Wigderson. Multi-prover interactive proofs: how to remove intractability assumptions. In *20th Annual ACM Symposium on Theory of Computing (STOC '88)*, pages 113–131. 1988.

[Bri11] J. Briët. *Grothendieck inequalities, nonlocal games and optimization*. Ph.D. thesis, Institute for Logic, Language and Computation, 2011.

[Bri20] J. Briët. Subspaces of tensors with high analytic rank. *Online J. Anal. Comb.*, 2020. To appear. Available at arXiv: 1908.04169.

[BSS16] S. Bravyi, G. Smith, and J. A. Smolin. Trading classical and quantum computational resources. *Physical Review X*, 6(2):021043, 2016.

[BST10] A. Ben-Aroya, O. Schwartz, and A. Ta-Shma. Quantum expanders: Motivation and construction. *Theory of Computing*, 6(1):47–79, 2010. doi: 10.4086/toc.2010.v006a003.

[BV13] J. Briët and T. Vidick. Explicit lower and upper bounds on the entangled value of multiplayer XOR games. *Communications in Mathematical Physics*, 321(1):181–207, 2013.

[CG02] F. Chung and R. Graham. Sparse quasi-random graphs. *Combinatorica*, 22(2):217–244, 2002. ISSN 0209-9683. doi: 10.1007/s004930200010. Special issue: Paul Erdős and his mathematics.

[CGKS98] B. Chor, O. Goldreich, E. Kushilevitz, and M. Sudan. Private information retrieval. *J. ACM*, 45(6):965–982, 1998. doi: 10.1145/293347.293350.

[CGW89] F. R. K. Chung, R. L. Graham, and R. M. Wilson. Quasi-random graphs. *Combinatorica*, 9(4):345–362, 1989. doi: 10.1007/BF02125347.

[Che70] J. Cheeger. A lower bound for the smallest eigenvalue of the laplacian. In *Problems in analysis*, pages 195–200. Princeton University Press, 1970.

[CHPS12] D. Conlon, H. Hàn, Y. Person, and M. Schacht. Weak quasi-randomness for uniform hypergraphs. *Random Structures & Algorithms*, 40(1):1–38, 2012.

[CHSH69] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt. Proposed experiment to test local hidden-variable theories. *Physical review letters*, 23(15):880, 1969.

[CHTW04] R. Cleve, P. Høyer, B. Toner, and J. Watrous. Consequences and limits of nonlocal strategies. In *19th Annual IEEE Conference on Computational Complexity (CCC 2004), 21-24 June 2004, Amherst, MA, USA*, pages 236–249. 2004. doi: 10.1109/CCC.2004.1313847.

[CJPPG15] T. Cooney, M. Junge, C. Palazuelos, and D. Pérez-García. Rank-one quantum games. *computational complexity*, 24(1):133–196, 2015. doi: 10.1007/s00037-014-0096-x.

[CL17] D. Conlon and J. Lee. Finite reflection groups and graph norms. *Advances in Mathematics*, 315:130–165, 2017.

[CSUU08] R. Cleve, W. Slofstra, F. Unger, and S. Upadhyay. Perfect parallel repetition theorem for quantum XOR proof systems. *Computational Complexity*, 17(2):282–299, 2008.

[CZ17] D. Conlon and Y. Zhao. Quasirandom Cayley graphs. *Discrete Anal.*, pages Paper No. 6, 14, 2017. ISSN 2397-3129. doi: 10.19086/da.1294.

[CZH+18] J. Chen, F. Zhang, C. Huang, M. Newman, and Y. Shi. Classical simulation of intermediate-size quantum circuits. *arXiv preprint arXiv:1805.01450*, 2018.

[Dav84] A. Davie. Lower bound for K_G, 1984. Unpublished.

[DDM03] J. Dehaene and B. De Moor. Clifford group, stabilizer states, and linear and quadratic operations over gf (2). *Physical Review A*, 68(4):042318, 2003.

[DHVY16] I. Dinur, P. Harsha, R. Venkat, and H. Yuen. Multiplayer parallel repetition for expander games. *arXiv preprint arXiv:1610.08349*, 2016.

[Dod84] J. Dodziuk. Difference equations, isoperimetric inequality and transience of certain random walks. *Transactions of the American Mathematical Society*, 284(2):787–794, 1984.

[dW19] R. de Wolf. Quantum computing: Lecture notes. *arXiv preprint arXiv:1907.09415*, 2019.

[Efr12] K. Efremenko. 3-Query locally decodable codes of subexponential length. *SIAM J. Comput.*, 41(6):1694–1703, 2012. doi: 10.1137/090772721. Preliminary version in STOC'09.

[FH13] W. Fulton and J. Harris. *Representation theory: a first course*, volume 129. Springer Science & Business Media, 2013. doi: 10.1007/978-1-4612-0979-9.

[FHH+14] Y. Filmus, H. Hatami, S. Heilman, E. Mossel, R. O'Donnell, S. Sachdeva, A. Wan, and K. Wimmer. Real analysis in computer science: A collection of open problems. *Preprint available at https://simons. berkeley. edu/sites/default/files/openprobsmerged. pdf*, 2014.

[FK91] H. Furstenberg and Y. Katznelson. A density version of the Hales-Jewett theorem. *J. Anal. Math.*, 57(1):64–119, 1991.

[FLW12] N. Frantzikinakis, E. Lesigne, and M. Wierdl. Random sequences and pointwise convergence of multiple ergodic averages. *Indiana Univ. Math. J.*, pages 585–617, 2012.

[FLW16] N. Frantzikinakis, E. Lesigne, and M. Wierdl. Random differences in Szemerédi's theorem and related results. *J. Anal. Math.*, 130:91–133, 2016. doi: 10.1007/s11854-016-0030-z.

[Fra16] N. Frantzikinakis. Some open problems on multiple ergodic averages. *Bull. Hellenic Math. Soc.*, 60:41–90, 2016. doi: 10.1109/tac.2015.2392613.

[GHSY12] P. Gopalan, C. Huang, H. Simitci, and S. Yekhanin. On the locality of codeword symbols. *IEEE Trans. Inform. Theory*, 58(11):6925–6934, 2012. doi: 10.1109/TIT.2012.2208937.

[GIM] GIMPS. Great Internet Mersenne Prime Search. `https://www.mersenne.org/`. Last accessed: February 2020.

[Gop18] S. Gopi. *Locality in coding theory*. Ph.D. thesis, Princeton University, 2018.

[Got97] D. Gottesman. *Stabilizer codes and quantum error correction*. California Institute of Technology, 1997.

[Got98] D. Gottesman. The heisenberg representation of quantum computers. *arXiv preprint quant-ph/9807006*, 1998.

[Gow98] W. T. Gowers. A new proof of Szemerédi's theorem for arithmetic progressions of length four. *Geometric & Functional Analysis GAFA*, 8(3):529–551, 1998.

[Gow01] W. T. Gowers. A new proof of Szemerédi's theorem. *Geom. Funct. Anal.*, 11(3):465–588, 2001. ISSN 1016-443X.

[Gow07] W. T. Gowers. Hypergraph regularity and the multidimensional Szemerédi theorem. *Annals of Mathematics*, pages 897–946, 2007.

[Gow10] W. T. Gowers. Decompositions, approximate structure, transference, and the Hahn–Banach theorem. *Bull. Lon. Math. Soc.*, 42(4):573–606, 2010.

[Gow21] W. T. Gowers. A quasirandomness implication. `https://gowers.wordpress.com/2018/11/10/a-quasirandomness-implication/`, 2021. Last accessed: August 2021.

[Gro53] A. Grothendieck. Résumé de la théorie métrique des produits tensoriels topologiques. *Bol. Soc. Mat. São Paulo*, 8:1–79, 1953.

[Gro06] D. Gross. Hudson's theorem for finite-dimensional quantum systems. *Journal of mathematical physics*, 47(12):122107, 2006.

[GTKL+20] T. Giurgica-Tiron, I. Kerenidis, F. Labib, A. Prakash, and W. Zeng. Low depth algorithms for quantum amplitude estimation. *arXiv:2012.03348*, 2020.

[GTZ12] B. Green, T. Tao, and T. Ziegler. An inverse theorem for the Gowers $U^{s+1}[N]$-norm. *Ann. of Math.*, pages 1231–1372, 2012.

[Haa85] U. Haagerup. The Grothendieck inequality for bilinear forms on C^*-algebras. *Adv. in Math.*, 56(2):93–116, 1985. ISSN 0001-8708. doi: 10.1016/0001-8708(85)90026-X.

[Haa87] U. Haagerup. A new upper bound for the complex Grothendieck constant. *Israel J. Math.*, 60(2):199–224, 1987. ISSN 0021-2172. doi: 10.1007/BF02790792.

[Har08] A. W. Harrow. Quantum expanders from any classical cayley graph expander. *Quantum Information & Computation*, 8(8):715–721, 2008.

[Has07] M. B. Hastings. Random unitaries give quantum expanders. *Phys. Rev. A (3)*, 76(3):032315, 11, 2007. ISSN 1050-2947. doi: 10.1103/PhysRevA.76.032315.

[Has09] M. B. Hastings. Superadditivity of communication capacity using entangled inputs. *Nature Physics*, 5(4):255, 2009. doi: 10.1038/nphys1224.

[Hat09] H. Hatami. *On generalizations of Gowers norms.* Ph.D. thesis, University of Toronto, 2009.

[Hat10] H. Hatami. Graph norms and Sidorenko's conjecture. *Israel Journal of Mathematics*, 175(1):125–150, 2010.

[HBD+15] B. Hensen, H. Bernien, A. E. Dréau, A. Reiserer, N. Kalb, M. S. Blok, J. Ruitenberg, R. F. Vermeulen, R. N. Schouten, C. Abellán, et al. Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres. *Nature*, 526(7575):682–686, 2015.

[HDDM05] E. Hostens, J. Dehaene, and B. De Moor. Stabilizer states and clifford operations for systems of arbitrary dimensions and modular arithmetic. *Physical Review A*, 71(4):042315, 2005.

[HH09] M. B. Hastings and A. W. Harrow. Classical and quantum tensor product expanders. *Quantum Information & Computation*, 9(3):336–360, 2009.

[HHL19] H. Hatami, P. Hatami, and S. Lovett. Higher-order fourier analysis and applications. *Foundations and Trends® in Theoretical Computer Science*, 13(4):247–448, 2019.

[HI95] U. Haagerup and T. Itoh. Grothendieck type norms for bilinear forms on C^*-algebras. *J. Operator Theory*, 34(2):263–283, 1995. ISSN 0379-4024.

[HL09] A. W. Harrow and R. A. Low. Efficient quantum tensor product expanders and k-designs. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*, pages 548–561. Springer, 2009.

[HLW06] S. Hoory, N. Linial, and A. Wigderson. Expander graphs and their applications. *Bull. Amer. Math. Soc.*, 43:439–561, 2006. doi: 10.1090/S0273-0979-06-01126-8.

[Hol02] A. S. Holevo. Remarks on the classical capacity of quantum channel. *arXiv preprint quant-ph/0212025*, 2002.

[Hol05] A. Holevo. Additivity conjecture and covariant channels. *International Journal of Quantum Information*, 3(01):41–47, 2005.

[Hol06] A. S. Holevo. The additivity problem in quantum information theory. In *International Congress of Mathematicians. Vol. III*, pages 999–1018. Eur. Math. Soc., Zürich, 2006.

[HS17] T. Häner and D. S. Steiger. 5 petabyte simulation of a 45-qubit quantum circuit. In *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*, pages 1–10. 2017.

[HV12] M. Howard and J. Vala. Qudit versions of the qubit $\pi/8$ gate. *Physical Review A*, 86(2):022316, 2012.

[IKP+08] T. Ito, H. Kobayashi, D. Preda, X. Sun, and A. C. C. Yao. Generalized Tsirelson inequalities, commuting-operator provers, and multi-prover interactive proof systems. In *Proceedings of the 2008 IEEE 23rd Annual Conference on Computational Complexity*, CCC '08, pages 187–198. IEEE Computer Society, 2008.

[KdW04] I. Kerenidis and R. de Wolf. Exponential lower bound for 2-query locally decodable codes via a quantum argument. *J. Comput. System Sci.*, 69(3):395–420, 2004. doi: 10.1016/j.jcss.2004.04.007. Earlier version in STOC'03.

[Kho02] S. Khot. On the power of unique 2-prover 1-round games. In *Proceedings of the Thiry-fourth Annual ACM Symposium on Theory of Computing*, STOC '02, pages 767–775. ACM, New York, NY, USA, 2002. ISBN 1-58113-495-9. doi: 10.1145/509907.510017.

[Kri77] J.-L. Krivine. Sur la constante de Grothendieck. *C. R. Acad. Sci. Paris Sér. A-B*, 284(8):A445–A446, 1977.

[KRS16] Y. Kohayakawa, V. Rödl, and M. Schacht. Discrepancy and eigenvalues of Cayley graphs. *Czechoslovak Math. J.*, 66(141)(3):941–954, 2016. ISSN 0011-4642. doi: 10.1007/s10587-016-0302-x.

[KS06] M. Krivelevich and B. Sudakov. Pseudo-random graphs. In *More sets, graphs and numbers*, volume 15 of *Bolyai Soc. Math. Stud.*, pages 199–262. Springer, Berlin, 2006. doi: 10.1007/978-3-540-32439-3_10.

[KT00] J. Katz and L. Trevisan. On the efficiency of local decoding procedures for error-correcting codes. In *Proc. 32nd STOC*, pages 80–86. ACM Press, 2000. doi: 10.1145/335305.335315.

[Lab21] F. Labib. Stabilizer rank and higher-order Fourier analysis. *arXiv preprint arXiv:2107.10551*, 2021.

[LMS08] S. Lovett, R. Meshulam, and A. Samorodnitsky. Inverse conjecture for the Gowers norm is false. In *Proceedings of the fortieth annual ACM symposium on Theory of computing*, pages 547–556. 2008.

[LN97] R. Lidl and H. Niederreiter. *Finite fields*, volume 20. Cambridge university press, 1997.

[LPS88] A. Lubotzky, R. Phillips, and P. Sarnak. Ramanujan graphs. *Combinatorica*, 8(3):261–277, 1988. doi: 10.1007/BF02126799.

[LS09] N. Linial and A. Shraibman. Lower bounds in communication complexity based on factorization norms. *Random Structures and Algorithms*, 34:368–394, 2009.

[LW20] Z.-W. Liu and A. Winter. Many-body quantum magic. *arXiv preprint arXiv:2010.13817*, 2020.

[Mar88] G. A. Margulis. Explicit group-theoretical constructions of combinatorial schemes and their application to the design of expanders and concentrators. *Problems of Information Transmission*, 24(1):39–46, 1988.

[Mer90] N. D. Mermin. Simple unified form for the major no-hidden-variables theorems. *Physical review letters*, 65(27):3373, 1990.

[Mer93] N. D. Mermin. Hidden variables and the two theorems of John Bell. *Reviews of Modern Physics*, 65(3):803, 1993.

[Mes95] R. Meshulam. On subsets of finite abelian groups with no 3-term arithmetic progressions. *J. Combin. Theory Ser. A*, 71(1):168–172, 1995.

[MS86] V. D. Milman and G. Schechtman. *Asymptotic theory of finite-dimensional normed spaces*, volume 1200 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1986. ISBN 3-540-16769-2. With an appendix by M. Gromov.

[NC02] M. A. Nielsen and I. Chuang. Quantum computation and quantum information, 2002.

[NRV14] A. Naor, O. Regev, and T. Vidick. Efficient rounding for the noncommutative Grothendieck inequality. *Theory Comput.*, 10(11):257–295, 2014. doi: 10.1145/2488608.2488618. Earlier version in STOC'13.

[PGN+17] E. Pednault, J. A. Gunnels, G. Nannicini, L. Horesh, T. Magerlein, E. Solomonik, and R. Wisnieff. Breaking the 49-qubit barrier in the simulation of quantum circuits. *arXiv preprint arXiv:1710.05867*, 15, 2017.

[PGWP+08] D. Pérez-García, M. Wolf, C. Palazuelos, I. Villanueva, and M. Junge. Unbounded violation of tripartite Bell inequalities. *Communications in Mathematical Physics*, 279:455, 2008.

[Pis12] G. Pisier. Grothendieck's theorem, past and present. *Bull. Amer. Math. Soc. (N.S.)*, 49(2):237–323, 2012. ISSN 0273-0979. doi: 10.1090/S0273-0979-2011-01348-9.

[PSV21] S. Peleg, A. Shpilka, and B. L. Volk. Lower bounds on stabilizer rank. *arXiv preprint arXiv:2106.03214*, 2021.

[Rag07] P. Raghavendra. A note on Yekhanin's locally decodable codes. *Electron. Coll. Comput. Complex. (ECCC)*, 14(16), 2007.

[Raz99] R. Raz. Exponential separation of quantum and classical communication complexity. In *31st annual ACM symposium on theory of computing*, pages 358–367. 1999.

[Ree91] J. Reeds. A new lower bound on the real Grothendieck constant, 1991. Available at `http://www.dtc.umn.edu/~reedsj/bound2.dvi`.

[Rot53] K. F. Roth. On certain sets of integers. *Journal of the London Mathematical Society*, 1(1):104–109, 1953.

[RV15] O. Regev and T. Vidick. Quantum XOR games. *ACM Trans. Comput. Theory*, 7(4):Art. 15, 43, 2015. ISSN 1942-3454. doi: 10.1145/2799560.

[San19] S. Sanyal. Fourier sparsity and dimension. *Theory of Computing*, 15(1):1–13, 2019.

[Sha20] F. Shao. personal communication, 2020.

[Sim96] B. Simon. *Representations of finite and compact groups*. 10. American Mathematical Soc., 1996. doi: 10.1090/gsm/010.

[Slo19] W. Slofstra. The set of quantum correlations is not closed. In *Forum of Mathematics, Pi*, volume 7. Cambridge University Press, 2019.

[SSAG16] M. Smelyanskiy, N. P. Sawaya, and A. Aspuru-Guzik. qhipster: The quantum high performance software testing environment. *arXiv preprint arXiv:1601.07195*, 2016.

[Ste13] C. L. Stewart. On divisors of Lucas and Lehmer numbers. *Acta Math.*, 211(2):291–314, 2013.

[SZ08] Y. Shi and Y. Zhu. Tensor norms and the classical communication complexity of nonlocal quantum measurement. *SIAM J. Comput.*, 38(3):753–766, 2008.

[Sze75] E. Szemerédi. On sets of integers containing no k elements in arithmetic progression. *Acta Arith*, 27(199-245):2, 1975.

[Tao05] T. Tao. The dichotomy between structure and randomness, arithmetic progressions, and the primes. *arXiv preprint math/0512114*, 2005.

[Tao12] T. Tao. *Higher order Fourier analysis*, volume 142. American Mathematical Soc., 2012.

[TC80] B. S. (Tsirelson) Cirel'son. Quantum generalizations of Bell's inequality. *Letters in Mathematical Physics*, 4(2):93–100, 1980.

[Tho87a] A. Thomason. Pseudorandom graphs. In *Random graphs '85 (Poznań, 1985)*, volume 144 of *North-Holland Math. Stud.*, pages 307–331. North-Holland, Amsterdam, 1987.

[Tho87b] A. Thomason. Random graphs, strongly regular graphs and pseudorandom graphs. In *Surveys in combinatorics 1987 (New Cross, 1987)*, volume 123 of *London Math. Soc. Lecture Note Ser.*, pages 173–195. Cambridge Univ. Press, Cambridge, 1987.

[Tsi87] B. S. Tsirelson. Quantum analogues of the Bell inequalities. The case of two spatially separated domains. *J. Soviet Math.*, 36:557–570, 1987.

[Tsi06] B. Tsirelson. Bell inequalities and operator algebras. 2006.

[TV06] T. Tao and V. Vu. *Additive combinatorics*, volume 105 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2006. doi: 10.1017/CBO9780511755149.

[TZ10] T. Tao and T. Ziegler. The inverse conjecture for the Gowers norm over finite fields via the correspondence principle. *Anal. PDE*, 3(1):1–20, 2010.

[TZ12] T. Tao and T. Ziegler. The inverse conjecture for the Gowers norm over finite fields in low characteristic. *Annals of Combinatorics*, 16(1):121–188, 2012.

[WH20] A. B. Watts and J. W. Helton. 3XOR games with perfect commuting operator strategies have perfect tensor product strategies and are decidable in polynomial time. *arXiv preprint arXiv:2010.16290*, 2020.

[WHKN18] A. B. Watts, A. W. Harrow, G. Kanwar, and A. Natarajan. Algorithms, bounds, and strategies for entangled XOR games. *arXiv preprint arXiv:1801.00821*, 2018.

[Wil18] R. R. Williams. Limits on Representing Boolean Functions by Linear Combinations of Simple Functions: Thresholds, ReLUs, and Low-Degree Polynomials. In R. A. Servedio, editor, *33rd Computational Complexity Conference (CCC 2018)*, volume 102 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 6:1–6:24. Schloss Dagstuhl–Leibniz-Zentrum für Informatik, Dagstuhl, Germany, 2018. ISBN 978-3-95977-069-9. ISSN 1868-8969. doi: 10.4230/LIPIcs.CCC.2018.6.

[Woo07] D. Woodruff. New lower bounds for general locally decodable codes. *Electron. Coll. Comput. Complex. (ECCC)*, 14(6), 2007.

[WZ12] T. D. Wooley and T. D. Ziegler. Multiple recurrence and convergence along the primes. *American Journal of Mathematics*, 134(6):1705–1732, 2012.

[Yek08] S. Yekhanin. Towards 3-query locally decodable codes of subexponential length. *J. ACM*, 55(1):1:1–1:16, 2008. doi: 10.1145/1326554.1326555. Preliminary version in STOC'07.

[Yek12] S. Yekhanin. Locally decodable codes. *Found. Trends Theor. Comput. Sci.*, 6(3):139–255, 2012. doi: 10.1561/0400000030.

[Zuk93] M. Zukowski. Bell theorem involving all settings of measuring apparatus. *Phys. Lett. A*, 177(290), 1993.

Index

Samenvatting

In dit proefschrift bestuderen we *quasirandomness* in verschillende contexten. Maar wat is quasirandomness? Beschouw een eerlijke munt die we, zeg, 100 keer opgooien en de uitkomst van elke worp opslaan. Dit zal een rij zijn van K (Kop) en M (Munt), bijvoorbeeld $KKMKMKKMMMKMKKMK$... etc. De uitkomst van elke worp is natuurlijk willekeurig, dus zal deze rij ook willekeurig zijn. Een typische eigenschap van zo een rij dat getuigt dat deze rij willekeurig is, is dat in elke grote deelrij de letters K en T ongeveer even vaak voorkomen. Als we een rij van K's en T's construeren op een niet-willekeurige wijze en het heeft deze eigenschap, dan zeggen we dat het quasirandom is. Dit is niet het enige typische eigenschap dat zulke rijen hebben, maar het is een simpel voorbeeld dat het idee schetst. We zullen nu zien dat quasirandomness in veel verschillende contexten voorkomt.

In Hoofdstuk 2 bestuderen we een merkwaardige eigenschap van de natuur genaamd *verstrengeling* door middel van *nonlokale spellen*. In een nonlokale spel krijgen twee spelers, genaamd Alice en Bob, vragen toegestuurd van een scheidsrechter. Alice en Bob moeten antwoorden terugsturen vanuit een vooraf afgesproken verzameling van antwoorden en de scheidsrechter accepteert hun antwoorden aan de hand van een vooraf afgesproken regel. Alles van het spel wordt besproken met de spelers: de vragen die ze kunnen verwachten, wat voor antwoorden ze kunnen terugsturen en wat de regel is dat hun antwoorden wordt geaccepteerd. Alice en Bob mogen een strategie bedenken voordat het spel begint. Zodra ze de vragen hebben gekregen van de scheidsrechter, mogen ze niet meer communiceren. In het bijzonder, weten ze niet welke vraag de ander heeft gekregen. In een klassieke wereld kunnen ze alleen *deterministische strategiëen* gebruiken. Dit komt erop neer dat de spelers, voordat het spel begint, besluiten wat te antwoorden op een gegeven vraag. Maar in een quantum mechanische wereld kunnen Alice en Bob hun antwoorden baseren op *meetuitkomsten* van hun fysieke systemen. Zulke strategiëen worden *quantum strategiëen* genoemd en kunnen betere winkansen geven voor het nonlokale spel als hun fysieke systemen verstrengeld

zijn. Het is belangrijk om te onderzoeken hoeveel beter je een nonlokale spel kan spelen (hogere winkans) met quantum strategiëen tegenover deterministische strategiëen. In Hoofdstuk 2 laten we zien dat voor bepaalde natuurlijke klassen van nonlokale spellen, het voordeel van quantum strategiëen over deterministische strategiëen beperkt is. De technieken die we daarbij gebruiken zijn gebaseerd op quasirandomness van bepaalde functies die geassocieerd worden met de nonlokale spellen.

In grafentheorie zijn er procedures om willekeurige "reguliere" grafen te construeren. Dit zijn grafen waarbij elke knooppunt, zeg, precies drie buren heeft. Een typische eigenschap van zulke grafen is dat ze heel erg goed kunnen "mixen": als je op welke knooppunt dan ook begint en willekeurig een stap zet naar een buur en dit herhaalt, zal je je na een klein aantal stappen mogelijk overal op de graaf kunnen bevinden! Deze eigenschap wordt *expansie* genoemd. Een ander interessante eigenschap van zulke willekeurige reguliere grafen is dat ze *uniform* zijn: voor elke paar van deelverzamelingen van de knooppunten is het aantal kanten tussen deze twee deelverzamelingen ongeveer gelijk aan wat je zou verwachten van een willekeurige graaf met dezelfde kantdichtheid. Expansie en uniformiteit zijn daarom quasirandom eigenschappen, omdat willekeurige grafen deze eigenschappen hebben. Verrassend genoeg zijn deze twee eigenschappen equivalent voor grafen met "relatief veel" kanten en voor heel symmetrische grafen. In Hoofdstuk 3 generaliseren we deze verrassende equivalentie van quasirandom eigenschappen naar het quantum geval.

In het gebied van quantumcomputers is het bekend dat algoritmes die alleen een bepaald soort quantum circuit gebruiken, genaamd *stabilizer circuits*, efficient gesimuleerd kunnen worden met behulp van klassieke computers. De quantumtoestanden die zulke circuits produceren worden *stabilizer toestanden* genoemd. Als we naast zulke circuits kopiëen van de zogenaamde *magische quantumtoestand* tot onze beschikking hebben, dan zouden we de volledige kracht van quantumcomputers tot onze beschikking hebben en kunnen we, bijvoorbeeld, grote getallen factoriseren in priemfactoren met Shor's algoritme. Deze magische quantumtoestand is echter een duur middel en we willen daarom weten hoe we deze toestand kunnen verkrijgen uit de simpelere, en goedkopere, stabilizer toestanden. Het aantal stabilizer stabilizer toestanden die we hierbij nodig hebben wordt ook wel de *stabilizer rang* genoemd van de magische quantumtoestand. Voor een willekeurige quantumtoestand geldt er dat de stabilizer rang heel hoog is, wat wil zeggen dat zulke toestanden erg duur zijn. Dit is een typische eigenschap van willekeurige quantumtoestanden en als een expliciete quantumtoestand deze eigenschap heeft, dan zeggen we dat het quasirandom is. Er wordt vanuit gegaan dat de (expliciete) magische quantumtoestand quasirandom is in deze zin, maar dit is tot zover nog niet bewezen. In Hoofdstuk 4 bestuderen we dit open probleem vanuit een ander gezichtspunt, gebruikmakend van *hogere orde Fourier analyse*.

Hogere orde Fourier analyse is een generalisatie van "gewone" Fourier analyse dat is ontstaan vanuit Gowers's Fourier analytische bewijs van een bekend

resultaat van Szemerédi's in 1974. Dit resultaat zegt, grof gezegd, dat bepaalde soort patronen genaamd *arithmetische progressies* onvermijdbaar zijn in grote deelverzamelingen van $\{1, 2, \ldots, N\}$, waarbij N een groot geheel getal is. Een arithmetische progressie is een rij van getallen zodat het verschil tussen opeenvolgende getallen hetzelfde is. Bijvoorbeeld, voor getallen x en d is $x, x + d, x + 2d$ een arithemitische progressie van lengte drie. In Hoofdstuk 5 laten we zien dat een bepaalde hypothese in hogere orde Fourier analyse niet waar is: we laten zien dat *duale functies*, objecten die relevant zijn voor bepaalde verfijningen van Szemerédi's stelling, meer quasirandom zijn dan eerder gedacht.

Abstract

In this dissertation, we study *quasirandomness* in several contexts. But what is quasirandomness? Consider a fair coin that we throw, say, 100 times and record the outcome of each coin flip. It will be a list of H (Heads) and T's (Tails), for example $HHTHTHHTTTHTHHTH$... etc. Since each coin flip was random, this sequence will be random as well. A typical property of such a sequence "certifying" randomness would be that in any large subsequence, H and T will occur approximately the same number of times. When a sequence that we construct in a non-random way has this property, we say that it is quasirandom. This is not the only typical property that such sequences have but is a simple example demonstrating the idea. We will now see that quasirandomness appears in a wide variety of contexts.

In Chapter 2 we study a curious property of nature called *entanglement* through *nonlocal games*. In a nonlocal game, two players called Alice and Bob get questions from a referee. They then have to answer from a prescribed set of answers and the referee accepts or rejects their combined answer according to some known condition that is also known beforehand. Everything about the game is known, i.e. the set of questions and answers and the acceptance criterion of the referee. Alice and Bob are allowed to come up with a strategy before the game starts, but once they receive their questions, they are not allowed to communicate. In particular, they don't know what question the other player has received. In a classical world, they can only use a *deterministic strategy*. This is simply deciding, before the game starts, what to answer when a certain question is received. In a quantum mechanical world, Alice and Bob can base their answers on outcomes of a *measurement* on their private physical systems. Such a strategy is called a quantum strategy and it can give better chances of winning the game if their systems are entangled. It is important to understand how much better the chances of winning a nonlocal game are if we are allowed to use strategies based on entangled systems compared to deterministic strategies. In Chapter 2 we show that for certain natural classes of multiplayer nonlocal games, the advantage of

quantum strategies over deterministic strategies is bounded. The techniques used are based on the quasirandomness of certain functions that are associated with the nonlocal game.

In graph theory, there are procedures that can generate random "regular" graphs, i.e. graphs that where each vertex has exactly, say, three neighbors. A typical property of such graphs is that they "mix" very well: if you started at any vertex of the graph and randomly went to one of its neighbors at each step, then after a small number of steps you could be anywhere! This property is often referred to as *expansion* of the graph. So if a deterministic graph has this property, we say that it is quasirandom. Another interesting property is *uniformity*, in this case, the graph should have the property that for arbitrary pairs of subsets of vertices the number of edges is roughly equal to what one would expect from a random graph of the same density. Surprisingly, these two properties are equivalent for "dense" graphs and for a class of very symmetric graphs called *vertex-transitive* graphs. In Chapter 3 we generalize this surprising equivalence of quasirandom properties in the quantum realm.

In quantum computing, algorithms using only certain quantum circuits called *stabilizer circuits* can be easily simulated on classical computers. The quantum states that such circuits can produce are called *stabilizer states*. If besides such circuits we are allowed to use copies of the *magic state*, we obtain the full power of quantum computation and we can, for example, factor large numbers very efficiently using Shor's algorithm. This magic state is an expensive resource and therefore we would like to know how we can obtain this quantum state using the simpler, or cheaper, stabilizer states. We will refer to the number of such states needed as the *stabilizer rank*. Interestingly, a random quantum state has the property that it has a high stabilizer rank, meaning that such states are very expensive resources. This is a typical property of random quantum states and explicit quantum states having this property are then quasirandom in this sense. It is believed that the (explicit) magic state is quasirandom, but so far no one has been able to prove that it has a high stabilizer rank. In Chapter 4 we study this problem from a completely different viewpoint than the previous literature, using tools from *higher-order Fourier analysis*.

Higher-order Fourier analysis is a generalization of "ordinary" Fourier analysis that grew out of Gowers's Fourier-analytic proof of a famous result by Szemerédi from 1974. It says, intuitively, that certain patterns called *arithmetic progression* are unavoidable in large subsets of the integers $\{1, 2, \ldots, N\}$. More precisely, Szemerédi's theorem states that any "dense" subset of the integers $\{1, 2, \ldots, N\}$ contains an arithmetic progression of arbitrary length, provided that the integer N is large enough. An arithmetic progression is a sequence of numbers such that consecutive numbers have equal differences. For example, an arithmetic progression of length three has the form $x, x+d, x+2d$ for some integer x called the starting point and non-zero integer d called the common difference. In Chapter 5 we disprove a conjecture in higher-order Fourier analysis: we show that *dual*

functions, objects relevant for certain refinements of Szemerédi's theorem, are more quasirandom than thought before.

Titles in the ILLC Dissertation Series:

ILLC DS-2016-01: **Ivano A. Ciardelli**
Questions in Logic

ILLC DS-2016-02: **Zoé Christoff**
Dynamic Logics of Networks: Information Flow and the Spread of Opinion

ILLC DS-2016-03: **Fleur Leonie Bouwer**
What do we need to hear a beat? The influence of attention, musical abilities, and accents on the perception of metrical rhythm

ILLC DS-2016-04: **Johannes Marti**
Interpreting Linguistic Behavior with Possible World Models

ILLC DS-2016-05: **Phong Lê**
Learning Vector Representations for Sentences - The Recursive Deep Learning Approach

ILLC DS-2016-06: **Gideon Maillette de Buy Wenniger**
Aligning the Foundations of Hierarchical Statistical Machine Translation

ILLC DS-2016-07: **Andreas van Cranenburgh**
Rich Statistical Parsing and Literary Language

ILLC DS-2016-08: **Florian Speelman**
Position-based Quantum Cryptography and Catalytic Computation

ILLC DS-2016-09: **Teresa Piovesan**
Quantum entanglement: insights via graph parameters and conic optimization

ILLC DS-2016-10: **Paula Henk**
Nonstandard Provability for Peano Arithmetic. A Modal Perspective

ILLC DS-2017-01: **Paolo Galeazzi**
Play Without Regret

ILLC DS-2017-02: **Riccardo Pinosio**
The Logic of Kant's Temporal Continuum

ILLC DS-2017-03: **Matthijs Westera**
Exhaustivity and intonation: a unified theory

ILLC DS-2017-04: **Giovanni Cinà**
Categories for the working modal logician

ILLC DS-2017-05: **Shane Noah Steinert-Threlkeld**
Communication and Computation: New Questions About Compositionality

ILLC DS-2017-06: **Peter Hawke**
The Problem of Epistemic Relevance

ILLC DS-2017-07: **Aybüke Özgün**
Evidence in Epistemic Logic: A Topological Perspective

ILLC DS-2017-08: **Raquel Garrido Alhama**
Computational Modelling of Artificial Language Learning: Retention, Recognition & Recurrence

ILLC DS-2017-09: **Miloš Stanojević**
Permutation Forests for Modeling Word Order in Machine Translation

ILLC DS-2018-01: **Berit Janssen**
Retained or Lost in Transmission? Analyzing and Predicting Stability in Dutch Folk Songs

ILLC DS-2018-02: **Hugo Huurdeman**
Supporting the Complex Dynamics of the Information Seeking Process

ILLC DS-2018-03: **Corina Koolen**
Reading beyond the female: The relationship between perception of author gender and literary quality

ILLC DS-2018-04: **Jelle Bruineberg**
Anticipating Affordances: Intentionality in self-organizing brain-body-environment systems

ILLC DS-2018-05: **Joachim Daiber**
Typologically Robust Statistical Machine Translation: Understanding and Exploiting Differences and Similarities Between Languages in Machine Translation

ILLC DS-2018-06: **Thomas Brochhagen**
Signaling under Uncertainty

ILLC DS-2018-07: **Julian Schlöder**
Assertion and Rejection

ILLC DS-2018-08: **Srinivasan Arunachalam**
Quantum Algorithms and Learning Theory

ILLC DS-2018-09: **Hugo de Holanda Cunha Nobrega**
Games for functions: Baire classes, Weihrauch degrees, transfinite computations, and ranks

ILLC DS-2018-10: **Chenwei Shi**
Reason to Believe

ILLC DS-2018-11: **Malvin Gattinger**
New Directions in Model Checking Dynamic Epistemic Logic

ILLC DS-2018-12: **Julia Ilin**
Filtration Revisited: Lattices of Stable Non-Classical Logics

ILLC DS-2018-13: **Jeroen Zuiddam**
Algebraic complexity, asymptotic spectra and entanglement polytopes

ILLC DS-2019-01: **Carlos Vaquero**
What Makes A Performer Unique? Idiosyncrasies and commonalities in expressive music performance

ILLC DS-2019-02: **Jort Bergfeld**
Quantum logics for expressing and proving the correctness of quantum programs

ILLC DS-2019-03: **András Gilyén**
Quantum Singular Value Transformation & Its Algorithmic Applications

ILLC DS-2019-04: **Lorenzo Galeotti**
The theory of the generalised real numbers and other topics in logic

ILLC DS-2019-05: **Nadine Theiler**
Taking a unified perspective: Resolutions and highlighting in the semantics of attitudes and particles

ILLC DS-2019-06: **Peter T.S. van der Gulik**
Considerations in Evolutionary Biochemistry

ILLC DS-2019-07: **Frederik Möllerström Lauridsen**
Cuts and Completions: Algebraic aspects of structural proof theory

ILLC DS-2020-01: **Mostafa Dehghani**
Learning with Imperfect Supervision for Language Understanding

ILLC DS-2020-02: **Koen Groenland**
Quantum protocols for few-qubit devices

ILLC DS-2020-03: **Jouke Witteveen**
Parameterized Analysis of Complexity

ILLC DS-2020-04: **Joran van Apeldoorn**
A Quantum View on Convex Optimization

ILLC DS-2020-05: **Tom Bannink**
Quantum and stochastic processes

ILLC DS-2020-06: **Dieuwke Hupkes**
Hierarchy and interpretability in neural models of language processing

ILLC DS-2020-07: **Ana Lucia Vargas Sandoval**
On the Path to the Truth: Logical & Computational Aspects of Learning

ILLC DS-2020-08: **Philip Schulz**
Latent Variable Models for Machine Translation and How to Learn Them

ILLC DS-2020-09: **Jasmijn Bastings**
A Tale of Two Sequences: Interpretable and Linguistically-Informed Deep Learning for Natural Language Processing

ILLC DS-2020-10: **Arnold Kochari**
Perceiving and communicating magnitudes: Behavioral and electrophysiological studies

ILLC DS-2020-11: **Marco Del Tredici**
Linguistic Variation in Online Communities: A Computational Perspective

ILLC DS-2020-12: **Bastiaan van der Weij**
Experienced listeners: Modeling the influence of long-term musical exposure on rhythm perception

ILLC DS-2020-13: **Thom van Gessel**
Questions in Context

ILLC DS-2020-14: **Gianluca Grilletti**
Questions & Quantification: A study of first order inquisitive logic

ILLC DS-2020-15: **Tom Schoonen**
Tales of Similarity and Imagination. A modest epistemology of possibility

ILLC DS-2020-16: **Ilaria Canavotto**
Where Responsibility Takes You: Logics of Agency, Counterfactuals and Norms

ILLC DS-2020-17: **Francesca Zaffora Blando**
Patterns and Probabilities: A Study in Algorithmic Randomness and Computable Learning

ILLC DS-2021-01: **Yfke Dulek**
Delegated and Distributed Quantum Computation

ILLC DS-2021-02: **Elbert J. Booij**
The Things Before Us: On What it Is to Be an Object

ILLC DS-2021-03: **Seyyed Hadi Hashemi**
Modeling Users Interacting with Smart Devices

ILLC DS-2021-04: **Sophie Arnoult**
Adjunction in Hierarchical Phrase-Based Translation

ILLC DS-2021-05: **Cian Guilfoyle Chartier**
A Pragmatic Defense of Logical Pluralism

ILLC DS-2021-06: **Zoi Terzopoulou**
Collective Decisions with Incomplete Individual Opinions

ILLC DS-2021-07: **Anthia Solaki**
Logical Models for Bounded Reasoners

ILLC DS-2021-08: **Michael Sejr Schlichtkrull**
Incorporating Structure into Neural Models for Language Processing

ILLC DS-2021-09: **Taichi Uemura**
Abstract and Concrete Type Theories

ILLC DS-2021-10: **Levin Hornischer**
Dynamical Systems via Domains: Toward a Unified Foundation of Symbolic and Non-symbolic Computation

ILLC DS-2021-11: **Sirin Botan**
Strategyproof Social Choice for Restricted Domains

ILLC DS-2021-12: **Michael Cohen**
Dynamic Introspection

ILLC DS-2022-01: **Anna Bellomo**
Sums, Numbers and Infinity: Collections in Bolzano's Mathematics and Philosophy

ILLC DS-2022-02: **Jan Czajkowski**
Post-Quantum Security of Hash Functions