Efficiently stable presentations from error-correcting codes

Michael Chapman[1], Thomas Vidick[2], and Henry Yuen[3]

[1]*Courant Institute of Mathematical Sciences, New York University, USA*
[2] *Faculty of Mathematics and Computer Science, Weizmann Institute of Science, Israel*
[3]*Department of Computer Science, Columbia University, USA*

March 7, 2024

Abstract

We introduce a notion of *efficient stability* for finite presentations of groups. Informally, a finite presentation using generators S and relations R is *stable* if any map from S to unitaries that approximately satisfies the relations (in the tracial norm) is close to the restriction of a representation of G to the subset S. This notion and variants thereof have been extensively studied in recent years, in part motivated by connections to property testing in computer science. The novelty in our work is the focus on *efficiency*, which, informally, places an onus on small presentations — in the sense of encoding length. The goal in this setup is to achieve non-trivial tradeoffs between the presentation length and its modulus of stability.

With this goal in mind we analyze various natural examples of presentations. We provide a general method for constructing presentations of $\mathbb{Z}_2^k$ from linear error-correcting codes. We observe that the resulting presentation has a weak form of stability exactly when the code is *testable*. This raises the question of whether testable codes give rise to genuinely stable presentations using this method. While we cannot show that this is the case in general, we leverage recent results in the study of non-local games in quantum information theory (Ji et al., Discrete Analysis 2021) to show that a specific instantiation of our construction, based on the Reed-Muller family of codes, leads to a stable presentation of $\mathbb{Z}_2^k$ of size $\operatorname{poly}\log(k)$ only. As an application, we combine this result with recent work of de la Salle (arXiv:2204.07084) to re-derive the quantum low-degree test of Natarajan and Vidick (IEEE FOCS'18), which is a key building block in the recent refutation of Connes' Embedding Problem via complexity theory (Ji et al., arXiv:2001.04383).

1 Introduction

Motivation. A linear error-correcting code $\mathscr{C}$ is a k-dimensional subspace of the vector space $\mathbb{F}^n$ over a finite field $\mathbb{F}$ that has certain combinatorial properties. The foremost of these is the *minimal distance* d, which is defined as the smallest Hamming weight (number of nonzero coordinates) $|c|$ of a nonzero vector $c \in \mathscr{C}$. In general one would like to design families of codes of increasing length n, such that both k and d are bounded below by a positive linear function of n. Such codes are referred to as "good" codes.

A finer property which concerns us here is the *soundness* of the code, a parameter that is connected to the notion of *testability*. A code can be (non-uniquely) specified through a *parity-check matrix* $h \in \mathbb{F}^{m\times n}$ as $\mathscr{C} = \ker h$. The rows of h are thought of as constraints ("parity checks") that specify $\mathscr{C}$ as a subspace of $\mathbb{F}^n$. A code is called testable *with soundness* ρ if for every $x \in \mathbb{F}^n$, $\frac{1}{m}|hx| \geq \rho\,\frac{1}{n}\,d(x,\mathscr{C})$, where $d(x,\mathscr{C})$

denotes the minimum of $|x - c|$ over $c \in \mathscr{C}$.[1] Ideally one would like to design families of good codes such that in addition ρ is bounded below by a constant independent of n.[2] This is a challenging task, and the construction of families of good (locally) testable codes was only achieved very recently [**?**, **?**].

The terminology "testable" comes from an interpretation of $\frac{1}{m}|hx|$ as the probability of rejection of a natural "tester" for $\mathscr{C}$, i.e. an algorithm that on input x checks a randomly chosen row h_i of h and accepts if and only if $h_i \cdot x = \sum_j h_{ij} x_j = 0$. Thus a code is called testable if words that are far from the code have a high probability of being rejected according to this tester. This notion (when accompanied by the locality constraint) plays a central role in applications of codes to complexity theory [**?**, **?**], and continues to be actively studied. See e.g. [**?**, **?**] for a recent breakthrough on the topic.

We make an observation that connects the study of testable codes to questions of stability in group theory and motivates our work. Let $\mathscr{C} = \ker h$ be a linear code as above, and suppose for simplicity that $\mathbb{F} = \mathbb{F}_2$ is the binary field. Consider the finitely presented group

$$G(h) = \langle S : R \rangle = \Big\langle x_1, \ldots, x_n : \quad x_j^2 = e \quad \forall 1 \le j \le n \,, \\ \prod_{1 \le j \le n} x_j^{h_{ij}} = e \,, \quad [x_{j_1}, x_{j_2}]^{h_{ij_1} h_{ij_2}} = e \quad \forall 1 \le i \le m \, \forall 1 \le j_1, j_2 \le n \Big\rangle \,. \tag{1}$$

Here a commutation relation between two generators $[x_{j_1}, x_{j_2}] = e$ is imposed only when needed for the relations $\prod_{1 \le j \le n} x_j^{h_{ij}} = e$ to make sense, i.e. two generators are required to commute only if they both take part in the same equation, which is the case if and only if $h_{ij_1} h_{ij_2} = 1$ for some i. While one could add all pairwise commutation relations, forcing $G(h)$ to be abelian — and indeed we will do this in some cases later — we choose the specific presentation (**??**) for consistency with the literature on non-local games, and particularly so-called *linear constraint system games*, which we review in more detail below.

We observe that 1-dimensional representations, i.e. maps from $S = \{x_1, \ldots, x_n\}$ to $\{-1, 1\}$ that satisfy all relations R, are in one-to-one correspondence with elements of $\mathscr{C}$. Moreover, *approximate* 1-dimensional representations, i.e. maps from $S = \{x_1, \ldots, x_n\}$ to $\{-1, 1\}$ that satisfy a fraction $1 - \varepsilon$ of the matrix relations for small ε, can be identified with words $x \in \mathbb{F}_2^n$ such that $\frac{1}{m}|hx| \le \varepsilon$.[3] In particular, we notice that $\mathscr{C}$ is testable with soundness ρ if and only if ε-approximate 1-dimensional representations of $G(h)$ are $C\varepsilon/\rho$-close to genuine 1-dimensional representations, where C depends on the distribution we choose over the relations. E.g., if we check an involution relation with probability $1/3$, a commutation relation with probability $1/3$, and a matrix relation with probability $1/3$, then $C = 3$.

Group stability. This observation immediately raises many questions. The problem of relating approximate representations of a group to exact representations of it is termed *stability* in group theory, and has a long history. There are of course many flavors of the problem, depending on how one defines closeness (should it be the operator norm or the Hilbert-Schmidt norm? Should closeness hold for every relation, or every pair of group elements, or is it sufficient that it holds on average? Etc.) We will review some relevant results in this area below. For now, we mention Voiculescu's famous counter-example [**?**] about

[1] In the literature, the notion of *local* testability is emphasized, where in addition the parity-check matrix is required to have rows of low Hamming weight. This requirement is less important for us, and so we de-emphasize it.

[2] Note that in principle a code can have a small minimal distance d, and still be testable with soundness $\rho > 0$. So a family of codes can be "testable" without being "good."

[3] We conditioned only on the matrix relations since the commutation and incolution relations are automatically satisfied by out choice of range $\{\pm 1\}$. In principle we could allow the map to range over $U(\mathbb{C})$ instead of $\{-1, 1\}$, and in general we will allow this. But, for the purposes of this introduction, it is simpler to restrict to the $\{-1, 1\}$-valued case.

approximately commuting unitaries, which was motivated by a question of Halmos on pairs of approximately commuting Hermitian operators [**?**]. Using the terminology of stability, Voiculescu showed that the presentation $\mathbb{Z}^2 = \langle x, y : [x, y] = e \rangle$ is *not* stable with respect to the operator norm. However, much more recently Glebsky [**?**] showed that the same presentation *is* stable with respect to the normalized Hilbert-Schmidt norm.

This example and many others show that the notion of stability is, in general, highly sensitive to the notion of closeness considered. Returning to our main concern, so far we have argued that stability of approximate 1-*dimensional* representations of the presentation $G(h)$ is connected to local testability of the code $\ker h$. In the case of 1-dimensional representations of course the choice of norm does not matter; however, the choice of measuring the error on average over relations, as opposed to e.g. taking the maximum, is one in which we depart from most of the literature. We will motivate this choice further below; but before we can continue we must pause to introduce the key definitions that our work builds on. For the purposes of the introduction we focus the discussion on finite-dimensional representations; in the main paper we handle the general case of representations in a tracial von Neumann algebra. Let $\mathcal{U}(\mathbb{C}^d)$ denote the unitary operators on $\mathbb{C}^d$, and for a set S let $\mathcal{F}(S)$ denote the free group generated by the elements of S. For $X \in \mathbb{C}^{d\times d}$ let $\|X\|_{hs}^2 = \frac{1}{d}\mathrm{Tr}(X^*X)$ denote the (normalized) Hilbert-Schmidt norm.[4]

Definition 1.1 (Almost homomorphism)**.** Let $G = \langle S : R \rangle$ be a finitely presented group and μ_R a distribution on R. An (ε, μ_R)-almost homomorphism of G is a homomorphism $\phi : \mathcal{F}(S) \to \mathcal{U}(\mathbb{C}^d)$ for some $d \geq 1$ such that

$$\mathop{\mathbb{E}}_{r\sim\mu_R} \left\|\phi(r) - I\right\|_{hs}^2 \leq \varepsilon \,.$$

As already mentioned this definition makes two important choices: firstly, to measure closeness in the Hilbert-Schmidt norm, and secondly, to measure it on average over the choice of a relation. Next we give our definition for a finitely presented group to be stable; see Definition **??** for the general setting.

Definition 1.2 (Stability)**.** Let $G = \langle S : R \rangle$ be a finitely presented group, μ_S a distribution on S and μ_R a distribution on R. For $\delta : [0,1] \to [0,1]$ such that $\lim_{t\to 0} \delta(t) = 0$ and an integer $d \geq 1$ we say that the presentation $G = \langle S : R \rangle$ is $(\delta, \mu_S, \mu_R, d)$-stable if for every (ε, μ_R)-almost homomorphism $\phi : \mathcal{F}(S) \to \mathcal{U}(\mathbb{C}^d)$ there is a unitary representation $\psi : G \to \mathcal{U}(\mathbb{C}^d)$ such that

$$\mathop{\mathbb{E}}_{s\sim\mu_S} \left\|\phi(s) - \psi(s)\right\|_{hs}^2 \leq \delta(\varepsilon) \,.$$

We refer to any function δ satisfying the above as a "modulus of stability" of the presentation.

Remark 1.3 (The L^∞ analogue)**.** *As mentioned before, it is common to call a homomorphism $\phi\colon \mathcal{F}(S) \to \mathcal{U}(\mathbb{C}^d)$ an ε-almost homomorphism of $G = \langle S\colon R\rangle$, if $\max_{r\in R} \|\phi(r) - I\|_{hs}^2 \leq \varepsilon$. Furthermore, the distance between two homomorphisms $\phi, \psi\colon \mathcal{F}(S) \to \mathcal{U}(\mathbb{C}^d)$ is ususally taken to be $\max_{s\in S} \|\phi(s) - \psi(s)\|_{hs}^2$. The notion of stability induced by these definitions of 'almost' and 'close' is more commonly used. Note that when studying a fixed finitely presented group, and without caring about the exact modulus of stability, there is no difference between the two definitions. But, when one cares about the modulus of stability, which is the case when viewing stability as a property testing problem, our framework is the more natural one.*

[4] This norm is also commonly called the normalized *Frobenius* norm. Following [**?**], it became common in stability theory to call the normalized version Hilbert–Schmidt and the un-normalized version Frobenius. In any case, in the main part of this paper, we relate to it as the tracial norm because of the von Neumann algebraic framework we use.

In this paper we also consider a version of *flexible* stability, i.e. the representation ψ is allowed to range in $U_{d'}(\mathbb{C})$ for some $d' \neq d$; see Definition **??** for the general definition. This requires a more careful definition of closeness; for now we restrict our attention to the simpler definition.

We can now ask the question: is $G(h)$, the group presentation defined in (**??**), stable according to Definition **??**? One can verify that with μ_R which chooses with probability $1/3$ whether to check an involution, a commutation or a matrix row, and μ_S the uniform distribution, $G(h)$ is $(\delta, \mu_S, \mu_R, 1)$-stable for $\delta = 3\varepsilon/\rho$ if and only if $\mathscr{C} = \ker h$ is testable with soundness $\rho > 0$. But what about higher-dimensional approximate representations? Are these also stable, or does one need to make further requirements on $\mathscr{C}$ beyond local testability? We do not yet have a comprehensive answer to these questions. However, the answer cannot be straightforward: even deducing basic properties about the group $G(h)$ given its presentation – let alone its stability properties – appears to be a challenging task. For example, there are examples of parity check matrices h for which the group $G(h)$ is non-abelian or even non-amenable; see Remark **??**. In fact, every finitely generated group can be embedded into $G(h)$ for some h (see [**?**]).

Efficient stability. Faced with the apparent difficulty of studying the general question, it is time to refine our focus and formulate the question which we *do* address. To start, let us explicitly note that while stability has previously (for the most part) been studied as a question about a group, our formulation makes it a question about a *presentation* of the group. In particular it is known [**?**, **?**] that for finite groups G (which are the only groups we consider in this paper) the multiplication table presentation, which has $|G|$ generators, one for every group element, and $|G|^2$ relations, one for every pairwise product, is $\delta(\varepsilon) = C\varepsilon$-flexibly stable[5] for some constant C that is independent of the group. This holds even for approximate representations in arbitrary tracial von Neumann algebras.

To simplify the problem let us consider the following presentation for an obviously finite and abelian group

$$\widetilde{G(h)} = \langle S : R \rangle = \Big\langle x_1, \ldots, x_n : \quad x_j^2 = e \quad \forall 1 \leq j \leq n \,, \\ \prod_{1 \leq j \leq n} x_j^{h_{ij}} = e \,, \quad [x_{j_1}, x_{j_2}] = e \quad \forall 1 \leq i \leq m \, \forall 1 \leq j_1, j_2 \leq n \Big\rangle \,. \tag{2}$$

This is the same as $G(h)$, except that all pairwise commutations have been added. As we show formally later (see Lemma **??**), it is not hard to check that $\widetilde{G(h)}$ is a presentation of the group $\mathbb{Z}_2^k$, for $k = \dim \ker h$. Our most important results pertain to the stability of the presentation $\widetilde{G(h)}$; although we will later (Section **??**) consider some non-abelian extensions of it.

Importantly for us, the results of [**?**, **?**] do not imply the same quantitative stability bounds for $\widetilde{G(h)}$ with respect to its defining presentation (**??**). The reason to prefer the presentation (**??**) as opposed to the multiplication table presentation of $\widetilde{G(h)}$ is that (**??**) is much more succinct: if n, k and m are linearly related then it has $\mathrm{poly}(k)$ generators and relations, as opposed to 2^k and 2^{2k} respectively. Such a gain is essential when one recalls the interpretation of the presentation $\widetilde{G(h)}$ as a "tester" for G — the size of the presentation is then directly related to the amount of *randomness* required by the tester (to sample a random relation); and in computer science applications randomness is seen as an essential resource (we discuss this more below, in the context of quantum computing).

[5] Here we assume that μ_R, μ_S are uniform over the relations and generators of the multiplication table presentation, respectively. Furthermore, the results of [**?**, **?**] apply to a notion of *flexible* stability, where the nearby exact representation may act on a space of larger, but not too much larger, dimension. See Definition **??**.

Can such "efficient" presentations of $G(h)$ still be stable? The naive approach, of extending an ε-approximate homomorphism of $G(h) = \langle S : R \rangle$ into an ε'-approximate homomorphism of the multiplication table presentation of $\mathbb{Z}_2^k$, leads to $\varepsilon' = \Omega(k^2\varepsilon)$ and hence a logarithmic dependence of the modulus of stability on the group size. Is it possible to do better? Glebsky's result for the case of $\mathbb{Z}^2$, which is an infinite group, suggests that in some cases the modulus of stability can be independent of the group size. It is therefore not clear what if any of the known group parameters should play a role in it in general.

The main result of this paper is to exhibit presentations of $\mathbb{Z}_2^k$ (and of slightly more complex 2-groups built on them, e.g. the *Pauli*[6] group ubiquitous in quantum information theory) that are *efficiently stable*: the size of the presentation is quasi-polynomial in k (as opposed to exponential), yet the modulus of stability only depends poly-logarithmically on k. These presentations are constructed from specific error-correcting codes, namely the Reed–Muller polynomial codes, which are known to have good local testability properties [**?**]. Our main result on group stability can be stated as follows (see Theorem **??** for the precise statement).

Theorem 1 (Main, informal)**.** *For every integer $k \geq 1$ there is a presentation $\mathbb{Z}_2^k = \langle S_k : R_k \rangle$ such that $|S_k|, |R_k| = 2^{\mathrm{poly}\log(k)}$ and furthermore this presentation is $(\delta, \mu_{S_k}, \mu_{R_k}, d)$-stable for all $d \in \mathbb{N}$, where $\delta(\varepsilon) = \mathrm{poly}(\log k, \varepsilon)$, μ_{S_k} is uniform over the generators and μ_{R_k} is some distribution over the relations.*

Most of the technical legwork required to prove the theorem is due to prior work in the study of nonlocal games in quantum computing, and in particular [**?**] (we explain this connection below). Our contribution in the present work is to make an explicit connection with stability and show how the former results can be "imported" to obtain new stability results such as the one stated in our main theorem above.

We do not know of any other presentation of $\mathbb{Z}_2^k$, arguably one of the simplest groups one could think of, that is stable with similar parameters as the ones stated in the theorem. It would be very interesting to discover different such presentations, built from testable error-correcting codes or not, for this group or others.

Nonlocal games in quantum information theory. To motivate our focus on *efficient* stability we now sketch a connection between our results and problems in quantum complexity theory and in particular the theory of nonlocal games that motivate us. As a result we will recover a key technical result used in the proof of the complexity result $\mathsf{MIP}^* = \mathsf{RE}$ [**?**] and its corresponding resolution of the Connes' Embedding Problem.

A *nonlocal game* $\mathfrak{G}$ is specified by the following data: finite question and answer sets $\mathcal{X}$ and $\mathcal{A}$ respectively, a distribution μ on $\mathcal{X} \times \mathcal{X}$, and a decision predicate $D : \mathcal{X} \times \mathcal{X} \times \mathcal{A} \times \mathcal{A} \to \{0,1\}$ (see Section **??** for details). The interpretation of $\mathfrak{G} = (\mathcal{X}, \mu, \mathcal{A}, D)$ as a game is as follows. A "referee" is imagined to sample a pair of "questions" $(x, y) \sim \mu$. Each question is sent to a different player, who is tasked with responding with an answer $a, b \in \mathcal{A}$ respectively. Finally, the referee decides that the players win the game if and only if $D(x, y, a, b) = 1$. Interestingly, the maximum success probability of the players in this game, where the probability is over the referee's choice of questions and any randomness in the player's strategy, and the maximum is taken over all allowed strategies, depends on whether one relies on "classical" or "quantum" interpretations of the game to determine appropriate mathematical formalization of the set of strategies that the players may employ. While a classical viewpoint naturally models a strategy as a pair of functions $f, g : \mathcal{X} \to \mathcal{A}$, one for each player, quantum mechanics invites one to consider a broader set of strategies in which an additional form of coordination between the players is allowed in the form of *shared quantum entanglement*. Understanding when there is a gap between the resulting maxima, and how large this gap can be, is of great interest in the foundations of quantum mechanics. To study this question one is

[6] This group is commonly referred to as the (multi-dimensional) *Heisenberg* group over $\mathbb{F}_2$, or the *Weyl–Heisenberg* group.

drawn to investigate the structure of optimal quantum strategies in a game, and how to design games that enforce a specific structure — informally, forcing as much "non-classicality" in winning strategies as possible. Beyond their foundational appeal, the theory of nonlocal games has had a very large impact in quantum cryptography (such as the analysis of device-independent quantum key distribution protocols [**?**, **?**]) and quantum complexity, in particular the theory of multiprover interactive proof systems [**?**].

For concreteness let us focus on a class of games called *linear constraint system* (LCS) games. These games were introduced in [**?**] and their study plays a central role in the celebrated result by Slofstra showing non-closure of the set of quantum correlations [**?**]. A linear constraint system game is parametrized by a matrix $h \in \mathbb{F}^{m\times n}$. In the game $\mathfrak{G}_h$, the referee selects a pair $(i,j) \in \{1,\ldots,m\} \times \{1,\ldots,n\}$ by first sampling i uniformly at random, and then sampling j uniformly at random conditioned on $h_{ij} = 1$. They send i to the first player and j to the second. The first player returns values in $\mathbb{F}$ for *each* j' such that $h_{ij'} \neq 0$, whereas the second players returns a single value in $\mathbb{F}$. The players win if the first player's answers satisfy the parity constraint, and the players' answers are consistent (the first player's answer associated with index j matches the second player's answer).

Now we see that to each matrix h we have associated a group $G(h)$, and a game $\mathfrak{G}_h$.[7] Moreover, and quite interestingly, there is a one-to-one correspondence between representations of $G(h)$ and *perfect strategies* in $\mathfrak{G}_h$, i.e. strategies that have success probability 1 in the game.[8] This correspondence enables one to "embed" group representations into quantum strategies, thereby forcing them to demonstrate a high level of complexity; this is the approach at the heart of [**?**]. Going further, for applications one is often required to understand not only optimal but also near-optimal strategies, whose success probability is e.g. $1-\varepsilon$. The same correspondence associates to such strategies approximate homomorphisms of $G(h)$ [**?**]. Here once can see that measuring closeness on average over the choice of relation is a natural choice, which is all but forced by the definition of a game, where the questions are selected according to some pre-specified distribution.

To summarize, approximate stability results for $G(h)$ enable one to obtain structural results about near-optimal strategies in $\mathfrak{G}_h$ — such a result is known as a *rigidity* result in quantum information. A rigidity result about a game called the quantum low-degree test [**?**] is at the heart of the proof of $\mathsf{MIP}^* = \mathsf{RE}$. The analysis of this test requires an efficient stability result for the Pauli group. Informally, the reason that the stability result needs to be for an efficient presentation of the Pauli group, as opposed to e.g. the multiplication table presentation, is because to obtain the final result it is necessary that the "complexity" of the test, or game, is smaller than the "complexity" of the object, or group being tested. Of course here we are loose about what we mean by "complexity," and refer to [**?**, **?**] for high-level explanations. The quantum low-degree test is described and analyzed in Section **??**. We end by mentioning an open question: if one was able to obtain $|S_k|, |R_k| = \mathrm{poly}(k)$ while also having $\delta(\varepsilon) = \mathrm{poly}(\varepsilon)$ in the informal theorem stated above, then this result would likely, through the connection we just described, have consequences for the efficient verification of quantum computations in the framework of interactive proof systems—see e.g. [**?**, **?**] for a sample of known results in this direction.

Related works on stability. The general question of group stability was first formulated by Ulam [**?**], and later studied by Kazhdan for the case of the operator norm [**?**]. See the introduction of [**?**] for a thorough

[7] (A similar correspondence holds between $\widetilde{G(h)}$ and a natural associated game $\widetilde{\mathfrak{G}}_h$; our main results apply to the latter, but the present discussion is more general and applies to both.

[8] This correspondence was established for finite-dimensional strategies, and finite-dimensional representations, in [**?**]. Extensions to infinite-dimensional strategies and representations have appeared in [**?**]. See also e.g. [**?**] for generalizations to the broader class of *synchronous games*.

history of these problems through the lens of approximate commutation (cf. [**?**, **?**, **?**]). A major contribution was done by [**?**], in which they characterized the stable amenable groups according to properties of their space of characters.

Stability of finite groups, with respect to the multiplication table representation, is shown in [**?**]. One can also consider representations in permutations, see [**?**, **?**]. Specifically, the analogous problem of efficient stability in permutations is still open. See the open problems section of [**?**]. Both stability in permutations and in unitaries equipped with the Hilbert–Schmidt metric are closely related to the notions of sofic and hyperlinear groups, cf. [**?**, **?**].

Open questions. We leave many questions open. A natural direction is to determine if there are simple sufficient conditions on the matrix h that guarantee that the presentation $G(h)$ (or $\widetilde{G(h)}$) is stable. As we discussed, the condition that $\ker h$ is a testable code is equivalent to stability for 1-dimensional permutations. It thus seems likely that testability is not a sufficient condition in general; can testability be determined by a combinatorial parameter of h? More generally, finding other examples of efficient stability seems of intrinsic interest, besides potential applications to property testing and the construction of nonlocal games. In a different direction, as mentioned in the previous paragraph an analogue of Theorem **??** for the case of representations in permutations could have important implications towards showing the existence of non-sofic groups [**?**, **?**, **?**].

Outline. We start in Section **??** by giving a precise definition of stability that we work with, and give examples to motivate and illustrate the definition. In Section **??** we give a general method for constructing presentations from codes, and apply the method to the spacial case of the Reed-Muller code. This leads in Section **??** to the statement and proof of our main result, an efficiently stable presentation for $\mathbb{Z}_2^k$. Finally in Section **??** we elaborate on the connection with the theory of nonlocal games and detail our applications to this area.

Acknowledgments. We would like to thank John Wright for his remarks on an early draft of this paper. MC acknowledges with gratitude the Simons Society of Fellows and is supported by a grant from the Simons Foundation (N. 965535). TV is supported by a research grant from the Center for New Scientists at the Weizmann Institute of Science, a Simons Investigator award, AFOSR Grant No. FA9550-22-1-0391, and ERC Consolidator Grant VerNisQDevS (101086733). HY is supported by AFOSR award FA9550-21-1-0040, NSF CAREER award CCF-2144219, and the Sloan Foundation.

2 Efficient stability

In this section we give definitions associated with the notion of "efficient stability" used in the paper. We reformulate some previously known results in this framework, and give examples that will be used later on.

2.1 Algebra background and notation

Here we call *tracial von Neumann algebra* a pair $(\mathcal{M}, \tau)$ of a von Neumann algebra $\mathcal{M}$ together with a normal faithful tracial state τ on $\mathcal{M}$, which we often refer to as the *trace*. The main example of interest is $\mathcal{M} = M_n(\mathbb{C})$, the algebra of $n \times n$ complex matrices, with τ the dimension-normalized trace, which we denote $\mathrm{tr}(M) = \frac{1}{n}\mathrm{Tr}(M)$. We write $\|x\|_\tau = \tau(x^*x)^{1/2}$ to denote the 2-norm on $\mathcal{M}$ with respect to τ — which agrees in the case of $M_n(\mathbb{C})$ with the Hilbert–Schmidt norm $\|x\|_{hs}$ discussed in the introduction.

Let $B(\ell_2)$ be the von Neumann algebra of bounded operators on ℓ_2, the Hilbert space of convergent sequences in $\mathbb{C}^{\mathbb{Z}}$ equipped with the usual Euclidean norm (for which we let $(e_i)_{i\in\mathbb{Z}}$ denote the standard basis). We denote $\mathcal{M}_\infty = \mathcal{M}\overline{\otimes}B(\ell_2)$, where the overline denotes closure for the operator topology. $\mathcal{M}_\infty$ is a von Neumann algebra equipped with the (infinite) trace $\tau_\infty = \tau \otimes \mathrm{Tr}$, with $\mathrm{Tr}(X) = \sum_{i\in\mathbb{Z}} e_i^T X e_i$ the trace on $B(\ell_2)$. We generally identify $\mathcal{M}$ with the "corner" $\mathcal{M} \otimes I_1 \subset \mathcal{M}_\infty$, where I_1 is the projection on the 1^{st} coordinate in $\mathbb{C}^{\mathbb{Z}}$.

2.2 Efficiently stable presentations

Suppose we are given a finite presentation of a (possibly infinite) group G using generators S and relations R. Informally, we say that the presentation is *stable* if any map from S to unitaries that approximately respects the relations R is close, in an appropriate sense, to a representation of G. Furthermore, we will say that the presentation is *efficient* if it is stable and provides a good trade-off between its size (the number of relations used and their length) and how the closeness to a representation depends on the error in satisfying the relations R. All the notions referred to informally in the preceding sentences — "approximately," "close," good trade-off," etc., can be formalized in a variety of ways, leading to generally incomparable definitions. Here we present the formalization that is most natural to us, and is motivated by applications to quantum information and complexity.

Given a set S, we let $\mathcal{F}(S)$ denote the free group generated by S. We identify functions from S to H, where H is any group, with homomorphisms from $\mathcal{F}(S)$ to H. If R is a subset of $\mathcal{F}(S)$ then the quotient of $\mathcal{F}(S)$ by the normal subgroup generated by R is denoted $\langle S : R\rangle$.

We start with the notion of an almost-homomorphism, which formalizes what it means for a map defined on S to approximately satisfy the relations R. The notion we give is a small variant of the notion of ε-*almost homomorphism* from a finitely presented group to a unital tracial C^*-algebra $\mathcal{A}$ introduced in [**?**, Section 2]. We give a variant of their definition that quantifies the error in an average sense. Below, when μ is a distribution over a finite set $\mathcal{X}$ and $f : \mathcal{X} \to \mathbb{R}$ we write $\mathbb{E}_{x\sim\mu} f(x)$ for the expectation of f under μ.

Definition 2.1 (Almost homomorphism)**.** Let $G = \langle S : R\rangle$ be a finitely presented group, μ a distribution on R, and $(\mathcal{M}, \tau)$ a tracial von Neumann algebra. An (ε, μ)-almost homomorphism of G on $(\mathcal{M}, \tau)$ is a homomorphism $\phi : \mathcal{F}(S) \to \mathcal{U}(\mathcal{M})$ such that

$$\mathop{\mathbb{E}}_{r\sim\mu} \left\|\phi(r) - I\right\|_\tau^2 \leq \varepsilon\,.$$

We note that this notion depends on the presentation $\langle S : R\rangle$ of G, not only on the group itself. When the distribution μ is uniform over the set R, we simply write ε-homomorphism. The definition is consistent with the usual notion of a homomorphism that factors through G, which is recovered when $\varepsilon = 0$ as long as μ_R and μ_S are fully supported.

A stability result is a statement that ε-homomorphisms are close to homomorphisms. To measure the distance between homomorphisms into different algebras we make the following definition.

Definition 2.2 (Closeness for unitaries)**.** Let $\{U_i\} \subseteq \mathcal{M}$ and $\{V_i\} \subseteq \mathcal{N}$ be two families of unitaries on tracial algebras $(\mathcal{M}, \tau^{\mathcal{M}})$ and $(\mathcal{N}, \tau^{\mathcal{N}})$ respectively, indexed by the same set $\mathcal{I}$. For $\delta \geq 0$ and μ a measure on $\mathcal{I}$ we say that $\{U_i\}$ and $\{V_i\}$ are (δ, μ)-close if there exists a projection $P \in \mathcal{M}_\infty$ of finite trace such that $\mathcal{N} = P\mathcal{M}_\infty P$ and $\tau^{\mathcal{N}} = \tau_\infty / \tau_\infty(P)$, and a partial isometry $w \in P\mathcal{M}_\infty I_{\mathcal{M}}$ such that

$$\mathop{\mathbb{E}}_{i\sim\mu} \left\|U_i - w^* V_i w\right\|_\tau^2 \leq \delta,$$ [9]

[9] We denote the norm by $\|\cdot\|_\tau$ and not $\|\cdot\|_{\tau^{\mathcal{M}}}$ for clarity of the notations. The von Neumann algebra in which we take the

and

$$\max\left\{\tau^{\mathcal{M}}(I_{\mathcal{M}} - w^*w)\,,\ \tau^{\mathcal{N}}(P - ww^*)\right\} \ \leq\ \delta\ .$$

If $\phi : \mathcal{I} \to \mathcal{U}(\mathcal{M})$ and $\psi : \mathcal{I} \to \mathcal{U}(\mathcal{N})$ then we say that ϕ and ψ are (δ, μ) close if the families $\{\phi(i)\}$ and $\{\psi(i)\}$ are. If the measure μ is omitted then it is understood to be the uniform measure on $\mathcal{I}$.

We now give our definition of stability.

Definition 2.3 (Stability)**.** Let $G = \langle S : R\rangle$ be a finitely presented group. Let $\mathcal{C}$ be a class of tracial von Neumann algebras. Let μ_S be a distribution on S and μ_R a distribution on R. Let $\delta : [0,1] \to [0,1]$ be a function satisfying $\lim_{t\to 0}\delta(t) = 0$. The presentation $G = \langle S : R\rangle$ is $(\delta, \mu_S, \mu_R, \mathcal{C})$-stable if for every $(\mathcal{M}, \tau)$ in $\mathcal{C}$, every (ε, μ_R)-almost homomorphism of G is $(\delta(\varepsilon), \mu_S)$-close to a unitary representation of G on some $(\mathcal{N}, \tau^{\mathcal{N}}) \in \mathcal{C}$. We refer to the function δ as the *modulus of stability* of the presentation.[10]

Remark 2.4. *Fixing a distribution μ_R on R only, there is a natural induced distribution μ_S on S which is obtained by sampling $r \in R$ from μ_R, and then sampling an $s \in S$ according to the atomic measure induced by R (seen as a multiset). For example, if $r = aba$ then conditioned on having chosen r, a is given probability $2/3$ and b is given probability $1/3$. Throughout this text, μ_S is always chosen as this induced distribution and so we often omit it from the notation.*

Definition **??** specifies what it means for a presentation to be stable, but not when the presentation is *efficiently* stable. As mentioned in the introduction, the goal is to optimise the tradeoff between the encoding length of the presentation and the resulting modulus of stability. We now choose a complexity measurement for presentations:

Definition 2.5 (Length of a presentation)**.** Let $G = \langle S\colon R\rangle$ be a finite presentation. The *length* of the presentation will be

$$\ell(G) = |S| + \sum_{r\in R} |r|,$$

where $|r|$ is the length of the word $r \in \mathcal{F}(S)$ written in the basis S.[11]

The search for the shortest possible presentations of finite groups, and in particular of finite simple groups, contained many twists and turns. Clearly, the shortest presentation of some groups, e.g. $G = \mathbb{Z}_2^k$, needs to be of size at least $\operatorname{poly}\log|G|$. On the other hand, it turns out that there are finite simple groups with minimal presentation length $\operatorname{poly}\log\log|G|$, which was a big surprise (see [**?**] and the references therein). Also, all finite groups (without ${}^2G_2(q)$ composition factors) have a presentation of length $O(\log^3(|G|))$. Thus, we can say that a presentation is "efficient", if its length is not much larger than the shortest possible one. This can be phrased in various ways, and certain choices of parameters can be applied in different ways. Our exact choice of tradeoff parameters between length and modulus of stability, which we shall refer to as efficiently stable, is motivated by natural examples which we outline next.

---

norm should be understood from context, and will usually be $\mathcal{M}$.

[10] Every function that satisfies this condition is a modulus of stability for $G = \langle S\colon R\rangle$, but we occasionally refer that way to the best possible δ in the definition.

[11] Another popular notion of length is the bit length of the encoding of $\langle S\colon R\rangle$, where exponents in $r \in R$ can be written in binary. Up to a factor of $\log|S|$, our notion is stricter than that. See [**?**].

2.3 General results

As was shortly reviewed in the introduction, and specifically in the related works subsection, many results about various notions of stability are known. Here, we give two results that are most relevant to our work. First of all, for a finite group G we can always write $G = \langle S : R \rangle$ where $S = G$ and $R = \{g \cdot h \cdot (gh)^{-1} = e\}$. We refer to this presentation as the *multiplication table presentation*. If we let μ_S and μ_R be the uniform distribution on S and on R respectively then Definition **??** reduces to a widely used notion of *flexible (Hilbert-Schmidt) stability*. In particular, for finite groups the following result is known [**?**, **?**]. We adopt the formulation from [**?**, Theorem 1.4].

Theorem 2.6. *Let G be a finite group and $\mathcal{C}$ the class of all tracial von Neumann algebras. Let μ_S and μ_R be the uniform distribution on $S = G$ and $R = \{g \cdot h \cdot (gh)^{-1} = e\}$ respectively. Then $G = \langle S : R \rangle$ is $(c\varepsilon, \mu_S, \mu_R, \mathcal{C})$-stable, where $c > 0$ is a universal constant independent of G.*

Next we state for later use a result from [**?**] which allows us to combine stability results. The results in [**?**] are rather general and apply to direct products and certain central extensions of a class of finite groups. Here, we will only use the following specialization to the case of the central extension of $\mathbb{Z}_2^k \times \mathbb{Z}_2^k$ by $\{-1, 1\}$ given by $\gamma(a, b) = (-1)^{a \cdot b}$, with $a \cdot b$ the inner product modulo 2. For a measure μ on $\mathbb{Z}_2^k$, define its inverse spectral gap

$$\kappa = \max_{a \neq 0} \frac{1}{1 - \mathbb{E}_{b \sim \mu} (-1)^{a \cdot b}} .$$ [12]

Theorem 2.7 ([**?**] Corollary 2.6)**.** *Let μ be a measure on $\mathbb{Z}_2^k$ with inverse spectral gap κ. Let $\phi_X, \phi_Z : \mathbb{Z}_2^k \to \mathcal{U}(\mathcal{M})$ be two homomorphisms such that*

$$\mathop{\mathbb{E}}_{a,b \sim \mu} \left\| \phi_X(a)\phi_Z(b) - (-1)^{a \cdot b} \phi_Z(b)\phi_X(a) \right\|_\tau^2 \leq \varepsilon .$$

Then there is an $\mathcal{N} = P\mathcal{M}_\infty P$ and homomorphisms $U_X, U_Z : \mathbb{Z}_2^k \to \mathcal{U}(\mathcal{N})$ and $\delta = O(\kappa^2 \varepsilon)$ such that ϕ_X and U_X are (δ, μ)-close, ϕ_Z and U_Z are (δ, μ)-close, and moreover $U_X(a)U_Z(b) = (-1)^{a \cdot b} U_Z(b)U_X(a)$ for all $a, b \in \mathbb{Z}_2^k$.

Corollary 2.8. *Theorem **??** essentially tells us the following: if we are given an almost homomorphism of the Pauli group (**??**), then we can fix it to a homomorphism in two steps. First, fix its restriction to the X and Z observables independently. Then, apply the theorem to get a homomorphism from the whole Pauli group. This idea is spelled out in detail in the proof sketch of Corollary **??**.*

2.4 Measurements and orthonormalization

Before giving some examples, we introduce the notion of a *positive operator-valued measure* (POVM), or more simply a *measurement*. This is a notion that comes from quantum mechanics and will be useful to formulate some of our statements. If $\mathcal{M}$ is a tracial von Neumann algebra, a measurement on $\mathcal{M}$ with outcome set $\mathcal{A}$ is a finite collection of positive semidefinite operators $\{P_a\}_{a \in \mathcal{A}}$ such that $\sum_a P_a = I_{\mathcal{M}}$. A measurement is *projective* if for all a, P_a is a projection.

In our results we will make use of the following elementary but powerful result, which allows us to "pull back" projective measurements through an isometry. The result is an application of *orthonormalization*,

[12] Note that this is indeed the inverse of the spectral gap of the random walk operator on $\mathbb{Z}_2^k$ induced by μ. Equivalently, by viewing μ as an element of the group ring $\mathbb{C}[\mathbb{Z}_2^k]$, it acts on it from the left, and thus has a spectral decomposition and a spectral gap.

which transforms a nearly-orthogonal measurement to a nearby orthogonal measurement. See e.g. [**?**, **?**] or [**?**, Theorem 1.2] for the version that we use here.

Lemma 2.9. *Let $(\mathcal{M}, \tau^{\mathcal{M}})$ be a tracial von Neumann algebra, $P \in \mathcal{M}_\infty$ a projection of finite trace, $\mathcal{N} = P\mathcal{M}_\infty P$ and $\tau^{\mathcal{N}} = \tau_\infty / \tau_\infty(P)$, and $w \in P\mathcal{M}_\infty I_{\mathcal{M}}$ a partial isometry. Let*

$$\varepsilon = \max\left\{\tau^{\mathcal{M}}\left(I_{\mathcal{M}} - w^*w\right),\ \tau^{\mathcal{N}}\left(P - ww^*\right)\right\}.$$

Then for any projective measurement $\{T_a\}_{a\in\mathcal{A}}$ on $\mathcal{N}$, where $\mathcal{A}$ is a finite set, there is a projective measurement $\{Q_a\}_{a\in\mathcal{A}}$ on $\mathcal{M}$ such that

$$\sum_{a\in\mathcal{A}} \left\|Q_a - w^*T_aw\right\|_\tau^2 \ \leq\ 56\varepsilon\,. \tag{3}$$

Proof. If $\varepsilon \geq \frac{1}{2}$ the conclusion is straightforward. This is because, whatever projective measurement $\{Q_a\}_{a\in\mathcal{A}}$ one chooses, we have

$$\|Q_a - w^*T_aw\|_\tau^2 \leq 2\|Q_a\|_\tau^2 + 2\|w^*T_aw\|_\tau^2,$$

and

$$\sum_{a\in\mathcal{A}} \|Q_q\|_\tau^2,\ \sum_{a\in\mathcal{A}} \|w^*T_aw\|_\tau^2 \leq 1.$$

So, assume $\varepsilon < \frac{1}{2}$. Define

$$\tilde{Q}_a = w^*T_aw + \frac{1}{|\mathcal{A}|}\left(I_{\mathcal{M}} - w^*w\right) \in \mathcal{M}\,.$$

Then $\{\tilde{Q}_a\}$ is a POVM on $\mathcal{M}$. Moreover,

$$\begin{aligned}
\sum_a \tau^{\mathcal{M}}\left(\tilde{Q}_a^2\right) &\geq \sum_a \tau^{\mathcal{M}}\left(\left(w^*T_aw\right)^2\right)\\
&= \sum_a \tau^{\mathcal{M}}\left(w^*T_aww^*T_aw\right)\\
&= \sum_a \tau^{\mathcal{M}}\left(w^*T_aPT_aw\right) - \sum_a \tau^{\mathcal{M}}\left(w^*T_a(P - ww^*)T_aw\right)\\
&\geq 1 - \varepsilon - \sum_a \tau_\infty\left(w^*T_a(P - ww^*)T_aw\right)\\
&\geq 1 - \varepsilon - \tau_\infty\Big(\left(P - ww^*\right)\Big(\sum_a T_aww^*T_a\Big)\Big)\\
&\geq 1 - \varepsilon - \tau_\infty\left(P - ww^*\right),
\end{aligned}$$

where the third line uses that $T_aPT_a = T_a$, $\sum_a T_a = I_{\mathcal{N}}$ and the definition of ε for the first term, and for the second the fact that for $A \in \mathcal{M}$, $\tau^{\mathcal{M}}(A) = \tau_\infty(A)$ by definition of τ_∞ and the identification of $\mathcal{M}$ with a "corner" in $\mathcal{M}_\infty$, the fourth line uses cyclicity of the trace for the second, and the last uses $\|ww^*\|_\infty, \|\sum_a T_a\|_\infty \leq 1$.[13] By assumption,

$$\tau_\infty\left(P - ww^*\right) \ \leq\ \varepsilon\,\tau_\infty(P) \ \leq\ \frac{\varepsilon}{1-\varepsilon}\,, \tag{4}$$

[13] The notation $\|\cdot\|_\infty$ refers to the operator norm.

where the last inequality is because by definition, $\tau^N(P) = 1$, thus

$$1 - \varepsilon \;\le\; \tau^{\mathcal{N}}(ww^*) \;=\; \frac{\tau_\infty(ww^*)}{\tau_\infty(P)} \;=\; \frac{\tau_\infty(w^*w)}{\tau_\infty(P)} \;\le\; \frac{1}{\tau_\infty(P)}$$

since $\tau_\infty(w^*w) = \tau^{\mathcal{M}}(w^*w)$ and $w^*w \le I_{\mathcal{M}}$. Overall,

$$\sum_a \tau^{\mathcal{M}}\big(\tilde{Q}_a^2\big) \;\ge\; 1 - \varepsilon - \frac{\varepsilon}{1-\varepsilon} \;\ge\; 1 - 3\varepsilon\,. \tag{5}$$

To conclude we apply [**?**, Theorem 1.2] to obtain a projective measurement $\{Q_a\}$ on $\mathcal{M}$ such that

$$\sum_a \big\|Q_a - \tilde{Q}_a\big\|_\tau^2 \;=\; 27\varepsilon\,.$$

Finally,

$$\begin{aligned}
\sum_a \big\|Q_a - w^*T_a w\big\|_\tau^2 &= \sum_a \Big\|Q_a - \tilde{Q}_a + \frac{1}{|\mathcal{A}|}\big(I_{\mathcal{M}} - w^*w\big)\Big\|_2^2 \\
&\le \sum_a 2\big\|Q_a - \tilde{Q}_a\big\|_\tau^2 + 2\frac{1}{|\mathcal{A}|}\big\|I_{\mathcal{M}} - w^*w\big\|_\tau^2 \\
&\le 54\varepsilon + 2\tau^{\mathcal{M}}\big((I_{\mathcal{M}} - w^*w)^2\big) \\
&\le 54\varepsilon + 2\tau^{\mathcal{M}}(I_{\mathcal{M}} - w^*w) \\
&\le 56\varepsilon\,,
\end{aligned}$$

where the second line is by the triangle inequality, the fourth line is due to the fact that $I_{\mathcal{M}} - w^*w$ is positive and has operator norm at most 1, and the last line is by $\tau^{\mathcal{M}}(I_{\mathcal{M}} - w^*w) \le \varepsilon$. □

2.5 Examples

As a first example we spell out the application of Theorem **??** to the case of $G = \mathbb{Z}_2^k$. Below, when we write $\mathbb{E}_{i\in\mathcal{X}}$ where $\mathcal{X}$ is a finite set, we mean the expectation over i chosen uniformly at random from $\mathcal{X}$, i.e. $\frac{1}{|\mathcal{X}|}\sum_{i\in\mathcal{X}}\cdot$.

Corollary 2.10. *Let $(\mathcal{M},\tau)$ be a tracial von Neumann algebra and $\phi : \mathbb{Z}_2^k \to \mathcal{U}(\mathcal{M})$ such that*

$$\mathop{\mathbb{E}}_{x,y\in\mathbb{Z}_2^k} \big\|\phi(x)\phi(y) - \phi(x+y)\big\|_\tau^2 \;\le\; \varepsilon\,.$$

Then there is a projective measurement $\{P_u\}_{u\in\mathbb{Z}_2^k}$ on $\mathcal{M}$ such that

$$\mathop{\mathbb{E}}_{x\in\mathbb{Z}_2^k} \Big\|\phi(x) - \Big(\sum_u (-1)^{u\cdot x} P_u\Big)\Big\|_\tau^2 \;=\; O(\varepsilon)\,.$$

Proof. Any ϕ as in the corollary statement is an (ε, U_R)-almost homomorphism of $\mathbb{Z}_2^k$ into $(\mathcal{M},\tau)$ for the multiplication table presentation. Applying Theorem **??**, ϕ is $O(\varepsilon)$-close to a homomorphism from $\mathbb{Z}_2^k$ to some $(\mathcal{N},\tau^{\mathcal{N}})$. Because $\mathbb{Z}_2^k$ is Abelian, such a homomorphism is given by commuting unitaries $(U_x)_{x\in\mathbb{Z}_2^k}$ on $\mathcal{N}$. Moreover, since $\mathbb{Z}_2^k$ is a 2-group, each U_x satisfies $U_x^2 = I$, hence $U_x = U_x^*$.

For every $u \in \mathbb{Z}_2^k$ let $Q_u = \mathbb{E}_x(-1)^{u\cdot x}U_x$. Then each Q_u is a projection on $\mathcal{N}$ such that $\sum_u Q_u = I$, and $U_x = \sum_u (-1)^{u\cdot x}Q_u$. Furthermore, by the conclusion of Theorem **??** it holds that

$$\mathop{\mathbb{E}}_{x\in\mathbb{Z}_2^k} \Big\| \phi(x) - w^* \Big(\sum_u (-1)^{u\cdot x} Q_u \Big) w \Big\|_\tau^2 \;=\; O(\varepsilon) \;, \tag{6}$$

for some partial isometry $w \in P\mathcal{M}_\infty I_{\mathcal{M}}$ as in Definition **??**. Using Lemma **??**, we find a projective measurement $\{P_u\}$ on $\mathcal{M}$ that satisfies

$$\sum_u \big\| P_u - w^* Q_u w \big\|_\tau^2 \;=\; O(\varepsilon) \;. \tag{7}$$

Thus

$$\begin{aligned} \mathop{\mathbb{E}}_{x\in\mathbb{Z}_2^k} \Big\| \sum_u (-1)^{u\cdot x} \big(P_u - w^* Q_u w \big) \Big\|_\tau^2 &= \mathop{\mathbb{E}}_{x\in\mathbb{Z}_2^k} \sum_{u,v} (-1)^{(u+v)\cdot x} \tau\big(\big(P_u - w^* Q_u w \big) \big(P_v - w^* Q_v w \big) \big) \\ &= \sum_u \big\| P_u - w^* Q_u w \big\|_\tau^2 \\ &= O(\varepsilon) \;, \end{aligned} \tag{8}$$

where the second line uses $\mathbb{E}_x(-1)^{w\cdot x} = 0$ if $w \neq 0$, and 1 otherwise, and the last line is by (**??**). Plugging back into (**??**) and using the triangle inequality shows the corollary. □

The multiplication table presentation of $\mathbb{Z}_2^k$ is quite long, in the sense of Definition **??**. It has as many generators as the group size, and quadratically as many relations, which gives a length of $O(2^{2k})$. There are much shorter presentations of $\mathbb{Z}_2^k$, for example the straightforward

$$\mathbb{Z}_2^k = \langle x_1, \ldots, x_k : [x_i, x_j] = e, x_i^2 = e \;\forall i \neq j \rangle \;, \tag{9}$$

where $[x_i, x_j] = x_i x_j x_i^{-1} x_j^{-1}$ is the group commutator. Its length is $O(k^2)$, which is polylogarithmic in the group size instead compared to the polynomial length of the multiplication table presentation. By [**?**], every presentation of a finite group has **some** modulus of stability. The following two lemmas show that, though the presentation (**??**) has a linear modulus of stability, it deteriorates by a constant factor k. In some sense, these two examples are on opposite sides of the efficient stability tradeoof: (**??**) is short but has a bad modulus of stability, while the multiplication table presentation is very long, while having an essentially optimal modulus of stability. As we keep recalling, the goal of this paper is to provide an example of a somewhat short presentation of $\mathbb{Z}_2^k$, which has a good enough modulus of stability to deduce the main technical results needed for MIP*=RE [**?**].

Lemma 2.11 (Lemma 3.8 in [**?**])**.** *Let $\mathcal{C}$ be the class of tracial von Neumann algebras. Let μ_R be the equal mixture of the uniform distribution on all words $[x_i, x_j]$ $(i \neq j)$ and the uniform distribution on all words x_i^2. Let μ_S be the uniform distribution on $\{x_1, \ldots, x_k\}$. Then for every k, there is a $\delta_k = O_k(\varepsilon)$ such that the presentation* (**??**) *is $(\delta_k, \mu_S, \mu_R, \mathcal{C})$-stable. Furthermore, the close representation can be taken on the same algebra.*

As one would expect, the dependence of δ_k on k depends on both the diameter of $\mathbb{Z}_2^k$ for the presentation (**??**), and certain values of its *Dehn function* — namely, the minimal volume of a Van Kampen diagram with perimeters of length 3. By expressing each element of $\mathbb{Z}_2^k$ as a product of generators in the natural way, and by applying Corollary **??**, it is possible to show that $\delta_k = O(k^2\varepsilon)$ in Lemma **??**. With more work one can get $\delta_k = O(k\varepsilon)$, see [**?**, Theorem 3.2]. This turns up to be tight, as the next lemma shows.

Lemma 2.12. *Let μ_R and μ_S be as in Lemma **??**. Then, for every $1 \leq c \leq \frac{k}{2}$, there is a $(c/\binom{k}{2}, \mu_R)$-almost homomorphism of $\langle S \colon R \rangle$ which is at least $(c/16k, \mu_S)$-far from any homomorphism (according to Definition **??**).*

Lemma **??** implies that the modulus of stability of (**??**) is $\Omega(k\varepsilon)$ whenever $\varepsilon \leq \frac{1}{k-1}$. This is because, for any $0 < \varepsilon < 1/k-1$, one can add trivial representations of $\mathbb{Z}_2^k$ to copies of the approximate representation described in Lemma **??** (choosing $c = k/2$), such that the resulting map is an (ε, μ_R)-approximate representation which is at least $\left((k-1)\varepsilon/32, \mu_S\right)$-away from any actual representation of $\mathbb{Z}_2^k$. We prove this Lemma in Appendix **??**, which also includes some discussion on the L^∞ variant of it.

In the next section we will obtain presentations of $\mathbb{Z}_2^k$ that have a much better length/modulus tradeoff than the one given in (**??**) or the multiplication table presentation. In the meantime, we give one last example. To formulate it we recall the definition of the Pauli matrices

$$\sigma^X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma^Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{10}$$

and more generally for $a, b \in \mathbb{Z}_2^k$ let

$$\sigma^X(a) = \bigotimes_{i=1}^{k} (\sigma^X)^{a_i} \qquad \text{and} \qquad \sigma^Z(b) = \bigotimes_{i=1}^{k} (\sigma^Z)^{b_i}. \tag{11}$$

These are self-adjoint unitary operators called Pauli observables. Each observable $\sigma^X(a)$ (resp. $\sigma^Z(b)$) corresponds to the *Pauli measurement* $\{\sigma_a^X\}_{a \in \mathbb{Z}_2^k}$ (resp. $\{\sigma_b^Z\}_{b \in \mathbb{Z}_2^k}$) where (in a slight abuse of notation)

$$\sigma_a^X = \mathop{\mathbb{E}}_{\alpha \in \mathbb{F}_2^k} (-1)^{a \cdot \alpha} \sigma^X(\alpha) \qquad \text{and} \qquad \sigma_b^Z = \mathop{\mathbb{E}}_{\beta \in \mathbb{F}_2^k} (-1)^{b \cdot \beta} \sigma^Z(\beta).$$

It is easy to verify that $\{\sigma_a^X\}_a$ and $\{\sigma_b^Z\}_b$ are projections summing to identity.

For an integer $k \geq 1$, the Pauli group P_k is the group generated by the Pauli matrices $\sigma^X(a)$, $\sigma^Z(b)$ introduced in (**??**). It can also be defined more abstractly as follows. Let $\gamma : \mathbb{Z}_2^k \times \mathbb{Z}_2^k \to \{-1, 1\}$ be given by $\gamma(a, b) = (-1)^{a \cdot b}$. Then P_k is the central extension of $\mathbb{Z}_2^k \times \mathbb{Z}_2^k$ by $\{-1, 1\}$ given by γ. This group is also known as the Heisenberg group

$$H_{2k+1} = \left\{ \begin{pmatrix} 1 & a & c \\ 0 & I_{k \times k} & b \\ 0 & 0 & 1 \end{pmatrix} \right\} \subseteq \mathrm{GL}_{k+2}(\mathbb{F}_2). \tag{12}$$

We consider the following presentation for the Pauli group:

$$\begin{aligned} \mathrm{P}_k \;=\; \big\langle \{J\} \cup \{(a,0),(0,b) : a, b \in \mathbb{Z}_2^k\} : {} & (a,0)^2 = (0,b)^2 = J^2 = e,\ [(a,0),J] = [(0,b),J] = e, \\ & (a,0)(a',0) = (a+a',0), (0,b)(0,b') = (0,b+b') \\ & (a,0)(0,b) = J^{a \cdot b}(0,b)(a,0) \quad \forall a, b, a', b' \in \mathbb{Z}_2^k \big\rangle. \end{aligned} \tag{13}$$

This presentation is not quite the multiplication table presentation (it has $2^{k+1} + 1$ generators, whereas $|\mathrm{P}_k| = 2^{2k+1}$), but it is not far from it. Applying Theorem **??** we obtain the following consequence.

Corollary 2.13 (Pauli braiding test)**.** *Let $\phi_X, \phi_Z : \mathbb{Z}_2^k \to \mathcal{U}(\mathcal{M})$ be maps such that for all $W \in \{X, Z\}$,*

$$\mathop{\mathbb{E}}_{a,b\in\mathbb{Z}_2^k} \left\|\phi_W(a)\phi_W(b) - \phi_W(a+b)\right\|_\tau^2 \leq \varepsilon \,,$$

and

$$\mathop{\mathbb{E}}_{a,b\in\mathbb{Z}_2^k} \left\|\phi_X(a)\phi_Z(b) - (-1)^{a\cdot b}\phi_Z(b)\phi_X(a)\right\|_\tau^2 \leq \varepsilon \,.$$

Then there is a projection $P \in \mathcal{M}_\infty$ such that if $\mathcal{N} = P\mathcal{M}_\infty P$ then $\mathcal{N} \simeq (M_2(\mathbb{C}))^{\otimes k} \otimes \mathcal{N}'$ for some $\mathcal{N}'$, and a partial isometry $w \in P\mathcal{M}_\infty I_\mathcal{M}$ such that for all $W \in \{X, Z\}$,

$$\mathop{\mathbb{E}}_{W\in\mathbb{Z}_2^k} \left\|\phi_W(a) - w^*\big(\sigma_W(a) \otimes I_{\mathcal{N}'}\big)w\right\|_\tau^2 = O(\varepsilon) \,.$$

Since this statement already appears in [**?**] (restricted to approximate homomorphisms into the finite-dimensional unitaries, and with slightly worse dependence on ε), and we will not need it here, we only sketch the proof.

Proof sketch. The idea is the following: The maps ϕ_X, ϕ_Z induce an ε-approximate representation of the Pauli group with respect to the presentation (**??**). From that, we can use the canonical form of elements in P_k to create an $O(\varepsilon)$-approximate representation of the multiplication table presentation of P_k — define $\phi : \mathrm{P}_k \to \mathcal{U}(\mathcal{M})$ by $\phi((-1)^c\sigma_X(a)\sigma_Z(b)) = (-1)^c\phi_X(a)\phi_Z(b)$, where $c \in \{\pm 1\}$ and $a, b \in \mathbb{Z}_2^k$. Applying Gowers–Hatami (Theorem **??**), we deduce the corollary. □

Similarly to Lemma **??** we can state a short version of the preceding corollary, with a bad modulus of stability, which applies to the presentation

$$\begin{aligned}\mathrm{P}_k = \big\langle x_1, \ldots, x_k, z_1, \ldots, z_k : x_i^2 = z_i^2 = J^2 = e,\ [x_i, J] = [z_i, J] = e,\\ [x_i, x_j] = [z_i, z_j] = [x_i, z_j] = e, [x_i, z_i] = J \quad \forall i \neq j\big\rangle \,. \end{aligned} \tag{14}$$

Lemma 2.14. *Let $\mathcal{C}$ be the class of tracial von Neumann algebras. Let μ_R be the following sampling procedure: With equal probability do one of the following*

- *Choose $J^2 = e$.*
- *Sample $i \in [k]$ uniformly and choose $x_i^2 = e$ (respectively, $z_i^2 = e$).*
- *Sample $i \in [k]$ uniformly and choose $[x_i, z_i] = J$.*
- *Sample $i \neq j \in [k]$ uniformly at random and choose $[x_i, x_j] = e$ (respectively, $[z_i, z_j] = e$ or $[x_i, z_j] = e$).*

*Let μ_S be the marginal of μ_R, as described in Remark **??**. Then the presentation (**??**) is $(O(k\varepsilon), \mu_S, \mu_R, \mathcal{C})$-stable.*

Proof. This follows from Lemma **??** and Theorem **??**, in a similar manner to our proof sketch of Corollary **??**. □

3 Presentations from codes

In this section we lay the groundwork for our main result, an "efficient" stable presentation of $\mathbb{Z}_2^k$ that is presented in the next section. Most of the technical work required is done in [**?**]. In this section we introduce the language required to reformulate their result in the framework of this paper.

As a first step, we introduce a general method for translating any binary linear error-correcting code into a presentation of $\mathbb{Z}_2^k$. Later we apply this method to the specific case of the Reed-Muller code (composed with the Hadamard code to obtain a binary code). However, the general method may be of independent interest.

3.1 The general construction

For q a prime power we let $\mathbb{F}_q$ denote the finite field with q elements. For n, k, d integer, an $[n, k, d]_q$ linear code $\mathscr{C}$ is a k-dimensional subspace of $\mathbb{F}_q^n$ such that for all $x \in \mathscr{C}$ such that $x \neq 0$, the Hamming weight $|x|$ (i.e. the number of nonzero coordinates of x) is at least d. The parameter n is called the *length* of the code, k its *dimension* and d its *distance*. A code can be specified by a *parity check matrix* $h \in \mathbb{F}_q^{m\times n}$ such that $\mathscr{C} = \ker h$.

For the remainder of this section we specialize the discussion to the case where $q = 2$. We make the simple but key observation that a parity check matrix for a code of dimension k implies a finite group presentation in the following way. Introduce n generators $S = \{x_1, \ldots, x_n\}$. Each of the generators is required to be an involution: $x_i^2 = e$. For each row $i \in \{1, \ldots, m\}$ of the parity check matrix h, introduce a relation

$$R_i \,:\, \prod_{1\leq j\leq n} x_j^{h_{ij}} = e \,,$$

that "verifies" the parity check associated with the i-th row of h. Finally, to guarantee that R_i is independent of the order in which the x_j are multiplied, whenever $j \neq j'$ are such that h_{ij} and $h_{ij'}$ are both nonzero we require that x_j and $x_{j'}$ commute. This can be written succinctly using a relation

$$R'_{ijj'} \,:\, [x_j, x_{j'}]^{h_{ij}h_{ij'}} = e \,.$$

The presentation obtained in this way defines a group $G = G(h)$, already introduced in (**??**) and which with the present notation reads

$$G(h) \,=\, \left\langle x_1, \ldots, x_n \,:\, x_j^2 = e \,,\; R_i \,,\; R'_{ijj'} \,, \quad \forall 1 \leq i \leq m,\, 1 \leq j < j' \leq n \right\rangle . \tag{15}$$

Note that we made the dependence of $G(h)$ on h explicit. This is because in general, the group defined in this way may depend on h, and not only on $\mathscr{C}$. If however we further impose *all* pairwise commutation relations, as in (**??**), then we obtain the following.

Lemma 3.1. *Let $R''_{jj'}$ be the commutation relation $[x_j, x_{j'}] = e$. Then*

$$\mathbb{Z}_2^k \,=\, \left\langle x_1, \ldots, x_n \,:\, x_j^2 = e \,,\; R_i \,,\; R''_{jj'} \,, \quad \forall 1 \leq i \leq m,\, 1 \leq j < j' \leq n \right\rangle .$$

Proof. The group defined by the right-hand side is obviously abelian and a 2-group, so it is of the form $\mathbb{Z}_2^{k'}$ for some k'. In fact, it is equal to the quotient of $\mathbb{Z}_2^n$ by the subgroup generated by the $\prod_{1\leq j\leq n} x_j^{h_{ij}}$. So it is $\mathbb{Z}_2^k$ where $k = n - \dim \operatorname{im} h = \dim \ker h = \dim \mathscr{C}$. □

Remark 3.2. *There exists matrices h such that $G(h)$ is not $\mathbb{Z}_2^k$, and in fact is not Abelian. For an example, see [**?**, Example 2.16]. For that example, $n = 12$, $k = 7$, and the parity check matrix h can be described explicitly as follows: the 7 rows of h are indexed by the vertices of the complete bipartite graph $K_{3,4}$, the 12 columns are indexed by the 12 edges of $K_{3,4}$, and the entry (i, j) of h is 1 if and only if the edge j is incident on vertex i. As shown in [**?**], $G(h)$ is not Abelian. By considering $K_{3,6}$ instead of $K_{3,4}$, one in addition obtains a non-Abelian* infinite *group. Using similar arguments it is possible to construct h such that $G(h)$ is not amenable, etc.; see the discussion in [**?**, Section 6].*

For readability it is convenient to reformulate the parity check matrix as a *tester* for the code. This allows us to give a more succinct, "algorithmic" definition of a parity check matrix for a given code. Informally, the tester takes as input a word $w \in \mathbb{F}_2^n$ and determines if $w \in \mathscr{C}$ by selecting a parity check at random and evaluating it. Specifically we give the following definition. (For the sake of later use, we state the definition for the case of a general prime power q.)

Definition 3.3 (r-local linear tester)**.** Let $\mathscr{C}$ be an $[n, k, d]_q$ linear code and $r \in \mathbb{N}$. An *r-local linear tester for $\mathscr{C}$* is a pair $M = (h, \nu)$ where $h \in \mathbb{F}^{m \times n}$ is a parity check matrix for $\mathscr{C}$, whose every row has Hamming weight at most r, and ν is a distribution over $\{1, \ldots, m\}$.

An r-local linear tester $M = (h, \nu)$ for $\mathscr{C}$ induces a pair of distributions (ν_R, ν_S) on the relations and generators of the presentation $G(h)$ (**??**) in a natural way: For the generators, we let ν_S be induced from ν by first sampling $j \sim \nu$ and then a uniformly random i such that $h_{ji} \neq 0$. For the relations, we let ν_R be the uniform mixture of the distribution ν_S on relations $x_j^2 = e$, the distribution ν on relations R_i, and the distribution $\nu \times \nu_S \times \nu_S$ on relations $R'_{jii'}$.

We end this section with an example, the *Hadamard code*. This code can be defined for any $t \geq 1$ and it is a $[T, t, T/2]_2$ linear code, where $T = 2^t$. For simplicity we write $\mathscr{C}_{\text{HAD}}$ to denote this code, omitting t. The Hadamard code is the subspace of linear functionals from $\mathbb{F}_2^t$ to $\mathbb{F}_2$ out of all such functions. As a linear space, $\mathscr{C}_{\text{HAD}}$ can be described as $(a \cdot b)_{a \in \mathbb{F}_2^t} \in \mathbb{F}_2^{\mathbb{F}_2^t} \cong \mathbb{F}_2^T$, for all $b \in \mathbb{F}_2^t$, where $a \cdot b = \sum_{i=1}^t a_i b_i$ is again the dot product modulo 2.

A parity check matrix for $\mathscr{C}_{\text{HAD}}$ is the matrix $h_{\text{HAD}} \in \mathbb{F}_2^{T^2 \times T}$ defined as follows. Identify the rows of h_{HAD} with pairs $(x, y) \in \mathbb{F}_2^t \times \mathbb{F}_2^t$, and the columns of h_{HAD} with $\mathbb{F}_2^t$. Then the (x, y)-th row of h_{HAD} has nonzero entries at positions x, y and $x + y$ only. The corresponding 3-local linear tester is $M_{\text{HAD}} = (h_{\text{HAD}}, \nu)$ where ν is the uniform distribution over $\mathbb{F}_2^t \times \mathbb{F}_2^t$. This tester can be described algorithmically, see Figure **??**.

Given access to some $g \in \mathbb{F}_2^T$, where $T = 2^t$, identify g with a function $g : \mathbb{F}_2^t \to \mathbb{F}_2$. Perform the following.

1. Select $(x, y) \in \mathbb{F}_2^t \times \mathbb{F}_2^t$ uniformly at random.
2. Accept if and only if $g(x) + g(y) + g(x + y) = 0$.

Figure 1: A 3-local linear tester for $\mathscr{C}_{\text{HAD}}$

Remark 3.4. *Since each pair of coordinates (x, y) appears together in at least one parity check, we can apply Lemma **??** to deduce that $G(h_{\text{HAD}}) = \mathbb{Z}_2^t$.*

We state our first stability result for a code-based presentation, the presentation $G(h_{\mathrm{HAD}})$ defined as (**??**) where h_{HAD} is defined above. To state the result we need to specify distributions μ_S and μ_R. We let μ_S be uniform over the 2^t generators, and μ_R the uniform distribution over the relations R_i. (Here, there is no need to place any weight on the relations $x_i^2 = e$, or on the commutation relations $R'_{iji'}$, because it can be seen that they follow from the other relations.)

Lemma 3.5. *Let $\mathcal{C}$ be the class of all tracial von Neumann algebras. The presentation $\mathbb{Z}_2^t = G(h_{\mathrm{HAD}})$, together with the distributions μ_S and μ_R defined above, is $(\delta, \mu_S, \mu_R, \mathcal{C})$ stable with $\delta(\varepsilon) = O(\varepsilon)$.*

Proof. This an immediate consequence of Theorem **??**, because $G(h_{\mathrm{HAD}})$ is the multiplication table presentation for $\mathbb{Z}_2^t$. □

3.2 The Reed-Muller code over $\mathbb{F}_q$

We introduce a family of codes that will lead to interesting presentations $G(h)$, whose stability we are able to analyze. Fix integers $m, t \in \mathbb{N}$ and let $q = 2^t$ and $M = 2^m$. Let $\mathcal{P}(q, m, d)$ be the vector space over $\mathbb{F}_q$ that consists of all m-variate polynomials f over $\mathbb{F}_q$ of individual degree at most d, that is all functions of the form

$$f(x_1, \ldots, x_m) = \sum_{\alpha \in \{0,1,\ldots,d\}^m} c_\alpha \, x_1^{\alpha_1} \cdots x_m^{\alpha_m} \, ,$$

where $\{c_\alpha\}$ is a collection of coefficients in $\mathbb{F}_q$. It is easy to verify that $\mathcal{P}(q, m, d)$ has dimension $k = (d+1)^m$ over $\mathbb{F}_q$. It follows that the linear span of all $(f(x))_{x \in \mathbb{F}_q^m}$, when ranging over all possible $\{c_\alpha\}$, defines a $[q^m, (d+1)^m, D]_q$ linear code over $\mathbb{F}_q$, where $D \geq (1 - md/q)q^m$ follows from the Schwartz-Zippel lemma.

Lemma 3.6 (Schwartz-Zippel lemma [**?**, **?**])**.** *Let $f, g : \mathbb{F}_q^m \to \mathbb{F}_q$ be two unequal polynomials with total degree at most d. Then*

$$\Pr_{x \sim \mathbb{F}_q^m} \big(f(x) = g(x)\big) \leq \frac{d}{q} \, .$$

The resulting code is called the *Reed-Muller code* $\mathscr{C}_{\mathrm{RM}}$ with parameters q, m, d. For $m = 1$, one obtains the *Reed-Solomon code* $\mathscr{C}_{\mathrm{RS}}$. A useful feature is that $\mathscr{C}_{\mathrm{RM}}$ can be seen as the m-fold tensor product of $\mathscr{C}_{\mathrm{RS}}$, i.e. $\mathscr{C}_{\mathrm{RM}} = \mathscr{C}_{\mathrm{RS}}^{\otimes m}$ as vector spaces over $\mathbb{F}_q$.

We define a local linear tester M_{RM} for the code $\mathscr{C}_{\mathrm{RM}}$ over $\mathbb{F}_q$. The tester is described as an algorithmic procedure in Figure **??**. The description makes use of interpolation coefficients, which are defined as follows. Fix $d+1$ distinct values $t_0, \ldots, t_d \in \mathbb{F}_q$. Then for all $u, v \in \mathbb{F}_q$ and $i \in \{0, \ldots, d\}$ define the interpolation coefficients

$$\alpha_{u,v,i} = \prod_{\substack{i'=0 \\ i' \neq i}}^{d} \frac{v - (u + t_{i'})}{t_i - t_{i'}} \, . \tag{16}$$

These are defined so that any polynomial $f : \mathbb{F}_q \to \mathbb{F}_q$ of degree at most d satisfies that for all $v \in \mathbb{F}_q$,

$$f(v) \; = \; \sum_{i=0}^{d} \alpha_{u,v,i} \, f(u + t_i) \, .$$

The tester verifies this relation along a randomly chosen *axis-aligned direction*. For all points $u \in \mathbb{F}_q^m$ and $j \in \{1, \ldots, m\}$, we say that the line through u parallel to the j-th axis is the set of points $\{u + te_j : t \in \mathbb{F}_q\}$ where $e_j = (0, \ldots, 0, 1, 0, \ldots, 0) \in \mathbb{F}_q^m$, where the unique 1 is in the j-th position.

Given access to some $g \in \mathbb{F}_q^n$, where $n = q^m$, identify g with a function $g : \mathbb{F}_q^m \to \mathbb{F}_q$. Perform the following.

1. Sample $u \in \mathbb{F}_q^m$ and $j \in \{1, \ldots, m\}$ uniformly at random. Let v be a uniformly random point on the line through u parallel to the j-th axis.
2. Accept if and only if $g(v) = \sum_{i=0}^{d} \alpha_{u,v,i} g(u + t_i e_j)$.

Figure 2: A local test for $\mathscr{C}_{\mathrm{RM}}$

A parity check matrix $h_{\mathrm{RM}} \in \mathbb{F}^{S \times q^m}$ for $\mathscr{C}_{\mathrm{RM}}$, where $S = q^m \times m \times q$, is as follows. Identify the rows with triples $(u, j, t) \in \mathbb{F}_q^m \times \{1, \ldots, m\} \times \mathbb{F}_q$. The (u, j, t)-th row of h_{RM} is the vector in $\mathbb{F}_q^m$ that for $i \in \{0, \ldots, d\}$ has the value $\alpha_{u,v,i}$ for $v = u + te_j$ in the coordinate indexed by $u + t_i e_j$, the value -1 in the coordinate indexed by $u + te_j$, and the value 0 everywhere else.

Rubinfeld and Sudan [**?**] (building on the work of Babai, Fortnow, and Lund [**?**]) showed that h_{RM} is indeed a parity check matrix for $\mathscr{C}_{\mathrm{RM}}$. Furthermore, they showed that the parity check matrix gives rise to a $(d+2)$-local tester for $\mathscr{C}_{\mathrm{RM}}$, whose soundness ρ (as defined in the introduction) is at least $\frac{1}{6}$.

3.3 Code composition

The Reed-Muller code from the previous section is defined over $\mathbb{F}_q$, for q a prime power. We can transform any q-ary code, for $q = 2^t$, into a binary code using the idea of *code composition* which we now describe.

Let $\kappa : \mathbb{F}_q \to \mathbb{F}_2^t$ denote an invertible linear map such that $\kappa(a)$ is the $\mathbb{F}_2$-representation of $a \in \mathbb{F}_q$ over some (implicitly specified) self-dual basis of $\mathbb{F}_q$ over $\mathbb{F}_2$. We extend κ and its inverse κ^{-1} to vectors over $\mathbb{F}_q$ coordinate-wise. We let $\mathrm{tr}(\cdot) : \mathbb{F}_q \to \mathbb{F}_2$ denote the trace over $\mathbb{F}_2$. Because we chose a self-dual basis for the binary representation, the trace satisfies $\mathrm{tr}(ab) = \kappa(a) \cdot \kappa(b)$ where the right-hand side means the $\mathbb{F}_2$-inner product of the vectors $\kappa(a), \kappa(b)$. For more details on the map κ and its properties, see [**?**, Section 3.3].

Let $q = 2^t$ and $\mathscr{C}$ be an $[n, k, d]_q$ linear code. Let $\mathscr{C}_{\mathrm{HAD}}$ be the Hadamard code over $\mathbb{F}_2^t$ (introduced at the end of Section **??**). Let $\mathscr{C}'$ be the $[qn, tk, d']$ linear code over $\mathbb{F}_2$ defined as follows. Given $a \in (\mathbb{F}_2^t)^k$, first map $a \mapsto a' = \kappa^{-1}(a) \in \mathbb{F}_q^k$. Then encode a' to $b' = \mathscr{C}(a') \in \mathbb{F}_q^n$. Finally, return $b = \mathscr{C}_{\mathrm{HAD}}(\kappa(b')) \in (\mathbb{F}_2^q)^n$, where $\mathscr{C}_{\mathrm{HAD}}$ is applied component-wise. Using that $\mathscr{C}_{\mathrm{HAD}}$ has distance $q/2$, it is easy to verify that this code has distance $d' \geq dq/2$.

Given an r-local tester $M = (h, \nu)$ for $\mathscr{C}$, there is a natural $\max(r, 3)$-local tester M' for $\mathscr{C}'$ which can be described as follows. Index coordinates of $\mathscr{C}'$ by pairs $(i, \alpha) \in [n] \times \mathbb{F}_2^t$, fixing a bijection between $[qn]$ and $[n] \times \mathbb{F}_2^t$. We describe an $\max(r, 3)$-local tester $M' = (h', \nu')$ for $\mathscr{C}'$. Informally, h' contains two type of checks. First, the A-checks consist of the repetition of n copies of the checks for the Hadamard code, one for each Hadamard-code encoding of an $\mathbb{F}_q$-symbol from $\mathscr{C}$. Second, the B-checks implement the checks of $\mathscr{C}$ specified by M, directly on the Hadamard encoding. More precisely, define h' to be the block matrix $h' = \begin{pmatrix} A \\ B \end{pmatrix}$ where

- $A \in \mathbb{F}^{nq^2 \times nq}$ is itself a block-diagonal matrix where the diagonal blocks are the $q^2 \times q$ parity check matrix for the Hadamard code. In other words, A can be viewed as $I_{n \times n} \otimes h_{\mathrm{HAD}}$ where $h_{\mathrm{HAD}} \in \mathbb{F}_2^{q^2 \times q}$

is the parity check matrix for the Hadamard code.

- $B \in \mathbb{F}^{\ell q \times nq}$ is viewed as having rows indexed by pairs $(p,\gamma) \in \{1,\dots,\ell\} \times \mathbb{F}_q$ and columns indexed by pairs $(i,x) \in \{1,\dots,n\} \times \mathbb{F}_2^t$. The entry in row (p,γ) and column (i,x) is 1 if and only if $h_{pi} \neq 0$ and $x = \kappa(\gamma h_{pi})$.

Define the distribution ν' as the uniform mixture of the uniform distribution on the rows of the A block matrix and the uniform distribution on the rows of the B block matrix.

Claim 3.7. *h' is a parity check matrix for $\mathscr{C}'$.*

Proof. Let $x \in \mathbb{F}_2^{qn}$ be such that $h'x = 0$. Since the A-checks enforce that each block of k symbols contains the Hadamard-code encoding of an $\mathbb{F}_q$ symbol, x can be decoded to $x' \in \mathbb{F}_q^n$ such that for each $(i,x) \in \{1,\dots,n\} \times \mathbb{F}_2^t$, $x_{(i,x)} = \kappa(x'_i) \cdot x$. If the p-th row of h enforces the check $v_p \cdot x' = 0$, where $v_p \in \mathbb{F}_q^n$, then the (p,γ)-th row of B enforces the check

$$\begin{aligned} 0 &= \sum_{j=1}^{n} \kappa(\gamma(v_p)_j) \cdot \kappa(x'_j) \\ &= \sum_{j=1}^{n} \mathrm{tr}(\gamma(v_p)_j x'_j) \\ &= \mathrm{tr}(\gamma(v_p \cdot x')) \,. \end{aligned}$$

Therefore, $h'x = 0$ is equivalent to $\mathrm{tr}(\gamma(v_p \cdot x')) = 0$ for all γ, which is equivalent to $v_p \cdot x' = 0$. Thus $\mathscr{C}' = \ker h'$, as desired. □

4 An efficient presentation for $\mathbb{Z}_2^k$

Fix integers $m,t,d \in \mathbb{N}$ and let $q = 2^t$. Let $\mathscr{C}_{\mathrm{RM2}}$ be the $[q^{m+1}, t(d+1)^m, D']$ code obtained by applying the composition procedure from Section **??** to the $[q^m,(d+1)^m,D]_q$ Reed-Muller code $\mathscr{C}_{\mathrm{RM}}$ from Section **??**. Let $N = q^{m+1}$ and $h_{\mathrm{RM2}} \in \mathbb{F}_2^{M \times N}$ the parity check matrix for $\mathscr{C}_{\mathrm{RM2}}$ obtained from the composition of the $(d+2)$-local tester for $\mathscr{C}_{\mathrm{RM}}$ with the 3-local tester for $\mathscr{C}_{\mathrm{HAD}}$. Then h_{RM2} has $M = q^{m+2}(1+m)$ rows, such that each row has at most $(d+2)$ nonzero entries. **??** presents an algorithmic interpretation of the local tester corresponding to the parity check matrix h_{RM2}.

Let $G_{\mathrm{RM2}} = G(h_{\mathrm{RM2}})$ be the group that is presented from h_{RM2} (recall from **??** that codes give rise to group presentations through the general construction (**??**)). We do not know if $G_{\mathrm{RM2}} = \mathbb{Z}_2^K$, with $K = t(d+1)^m$. Instead we modify the presentation $G(h_{\mathrm{RM2}})$ by adding pairwise commutation relations, as in (**??**). Let

$$G(h_{\mathrm{RM2}}) = \left\langle x_1,\dots,x_N \,:\, \{R_k^{\mathrm{SQ}}\},\ \{R_k^{\mathrm{LD}}\},\ \{R_k^{\mathrm{HAD}}\} \right\rangle ,$$

where R_k^{SQ} ranges over all relations of the form $x_i^2 = e$ for $i \in \{1,\dots,N\}$, R_k^{LD} ranges over all relations implied by the "low-degree test" in Figure **??** and R_k^{HAD} ranges over all the relations implied by the "Hadamard test." For $k = (i,j) \in \{1,\dots,N\}^2$ such that $i < j$ let R_k^{COM} be the relation $[x_i,x_j] = e$. Then we define

$$\tilde{G} := \widetilde{G(h_{\mathrm{RM2}})} = \left\langle x_1,\dots,x_N \,:\, \{R_k^{\mathrm{SQ}}\},\ \{R_k^{\mathrm{LD}}\},\ \{R_k^{\mathrm{HAD}}\},\ \{R_k^{\mathrm{COM}}\} \right\rangle . \tag{17}$$

Given access to some $g \in \mathbb{F}_2^N$, where $n = q^{m+1}$, identify g with a function $g : (\mathbb{F}_2^t)^m \times \mathbb{F}_2^t \to \mathbb{F}_2$. Perform one of the following tests with probability $\frac{1}{2}$ each.

1. **Low-degree test:** Let $u \in \mathbb{F}_q^m$ be a uniformly random point and $j \in \{1, \ldots, m\}$ chosen uniformly at random. Let ℓ be the line through u in the j-th direction. Let $e_j = (0, \ldots, 0, 1, 0, \ldots, 0) \in \mathbb{F}_q^m$, where the unique 1 is in the j-th position. Choose a uniformly random $v \in \ell$ and $\gamma \in \mathbb{F}_q$ and check that

$$\sum_{i=0}^{d} g(u + t_i e_j, \kappa(\gamma \alpha_{u,v,i})) \;=\; g(v, \kappa(\gamma)) \;,$$

   where the interpolation points $t_0, \ldots, t_d \in \mathbb{F}_q$ and the $\alpha_{u,v,i}$ are defined in (**??**).

2. **Hadamard test:** Let $u \sim \mathbb{F}_q^m$ be chosen uniformly at random and $\alpha, \beta \in \mathbb{F}_2^t$ chosen uniformly at random. Check that

$$g(u, \alpha) + g(u, \beta) \;=\; g(u, \alpha + \beta) \;.$$

Figure 3: A local test for $\mathscr{C}_{\mathrm{RM2}}$

From Lemma **??** it follows that $\tilde{G} = \mathbb{Z}_2^K$. Our main result is an efficient stability result for this presentation. To state this we need to introduce distributions μ_S and μ_R on the generators and relations of $\tilde{G}$. The distribution μ_R is obtained as follows. With probability $1/4$ each, a relation from $\{R_k^{\mathrm{SQ}}\}$, $\{R_k^{\mathrm{LD}}\}$ or $\{R_k^{\mathrm{HAD}}\}$ is chosen uniformly at random. With probability $1/4$, a random commutation relation from R_k^{COM} is chosen according to the uniform mixture of the following two distributions:

1. For the first distribution, we select $u \in \mathbb{F}_q^m$ uniformly at random, $j \in \{1, \ldots, m\}$ uniformly at random, and $i \neq i' \in \{0, \ldots, d\}$ uniformly at random. Then select $\alpha, \beta \in \mathbb{F}_2^t$ uniformly at random and check commutation between $x_{u+t_i e_j, \alpha}$ and $x_{u+t_{i'} e_j, \beta}$, where we interpret the subscripts $u + t_i e_j, \alpha$ and $u + t_{i'} e_j, \beta$ as corresponding to some integer in $[N] = (\mathbb{F}_2^t)^m \times \mathbb{F}_2^t$ (as sets!).

2. For the second distribution, we first select $j \in \{1, \ldots, m\}$ and $u_{m-j+2}, \ldots, u_m \in \mathbb{F}_2^t$ uniformly at random. Then select $v, v' \in (\mathbb{F}_2^t)^m$ uniformly at random, conditioned on the last $(j-1)$ coordinates of each vector matching $u_{m-j+2}, \ldots, u_m$. Finally, select $\alpha, \beta \in \mathbb{F}_2^t$ uniformly at random and check commutation between $x_{v,\alpha}$ and $x_{v',\beta}$.

Having defined μ_R, we define μ_S as in Remark **??**. It is easy to check that in this way we obtain that μ_S is the uniform distribution over $[N]$. The following is our main technical result.

Theorem 4.1. *Let $\mathcal{C}$ be the class of all tracial von Neumann algebras. The presentation of $\mathbb{Z}_2^K$ given in (**??**), together with the distributions μ_S and μ_R defined above, is $(\delta, \mathcal{C})$ stable with $\delta(\varepsilon) = \mathrm{poly}(m, d, t) \cdot \mathrm{poly}(\varepsilon, q^{-1})$.*

We briefly explain a possible setting of parameters in Theorem **??**. For any integer $t \geq 1$, fix $q = 2^t$ and let $d = m = ct^c$ for some constant $c > 0$. Then $K = t(d+1)^m = 2^{\Theta((\log q)^c \log\log q)}$ and $N = q^{m+1} = 2^{\Theta((\log q)^{c+1} \log\log q)}$. Moreover, the number of relations in (**??**) is $O(N^2)$, which scakes as $2^{\mathrm{poly}\log K}$; and

the maximum length of a relation is $d = O(\log K)$. Finally, with this choice of parameters the function $\delta(\varepsilon)$ scales as $\mathrm{poly}(\log K) \cdot \mathrm{poly}(\varepsilon)$. We have thus obtained a presentation of $\mathbb{Z}_2^K$ of size only mildly superpolynomial in K, and with soundness that depends polylogarithmically on K.

Proof. Let $(\mathcal{M}, \tau)$ be a tracial von Neumann algebra and ϕ be an ε-homomorphism of $\langle S : R \rangle$ on $(\mathcal{M}, \tau)$. Here, $S = \{s_{u,a} : u \in (\mathbb{F}_2^t)^m, a \in \mathbb{F}_2^t\}$. We sometimes enumerate the items of S as $S = \{x_i : i \in \{1, \ldots, N\}\}$, where $N = 2^{tm+t}$ and we fixed an arbitrary bijection between $(\mathbb{F}_2^t)^m \times \mathbb{F}_2^t$ and $\{1, \ldots, N\}$. Let R be the set of all relations in (**??**), i.e. $R = \{R_k^{\mathrm{SQ}}\} \cup \{R_k^{\mathrm{LD}}\} \cup \{R_k^{\mathrm{HAD}}\} \cup \{R_k^{\mathrm{COM}}\}$. Let $N = |S|$.

The proof strategy is to perform a reduction to [**?**, Theorem 4.1]. Towards this, the main technical work in the proof consists in using the unitaries $\phi(s_{u,a})$ in order to define a synchronous strategy in the tensor code test from [**?**], where the underlying code is the Reed-Solomon code with degree d over $\mathbb{F}_q$. To define the synchronous strategy, we need "points," "lines," and "pair" measurements (see [**?**] for the terminology). Each of these is a family of projective measurements that obey certain constraints.

The proof consists of a sequence of claims, which examine the constraints imposed on the $\phi(s_{u,a})$ by each of the four collections of relations in (**??**) in turn, each time using a set of relations to evidence a particular structure among these unitaries. Eventually, this will allow us to define the required families of projective measurements from them.

We first exploit the relations R_k^{SQ} to show that we may assume without loss of generality that ϕ sends each element of S to a Hermitian involution.

Claim 4.2. *There is an $\varepsilon^{(1)} = O(d^2\varepsilon)$ and an $\varepsilon^{(1)}$-homomorphism $\phi^{(1)}$ of $\langle S : R \rangle$ on $(\mathcal{M}, \tau)$ such that $\phi^{(1)}(x)$ is a Hermitian involution for all $x \in S$, and furthermore*

$$\mathop{\mathbb{E}}_{x \sim \mu_S} \left\| \phi(x) - \phi^{(1)}(x) \right\|_\tau^2 \leq \varepsilon^{(1)} . \tag{18}$$

Proof. Using elementary calculations (see e.g. [**?**, Lemma 3.6]) we see that for any complex α,

$$\left| \operatorname{sgn} \Re \alpha - \alpha \right| \leq \left(1 + \frac{1}{\sqrt{2}}\right) \left| \alpha^2 - 1 \right| .$$

For any $i \in \{1, \ldots, N\}$ let $\phi^{(1)}(x_i) = \operatorname{sgn} \Re(\phi(x_i))$. Then the claim follows since

$$\mathop{\mathbb{E}}_{i \in \{1, \ldots, N\}} \left\| \phi(x_i)^2 - I \right\|_\tau^2 \leq 4\varepsilon ,$$

by assumption and the definition of μ_R, which places weight $\frac{1}{4}$ on the relations in R_k^{SQ}. Since μ_S is uniform on S, (**??**) follows. Furthermore, because all relations in R have length at most $O(d)$, and μ_S is defined from μ_R as in Remark **??**, a simple averaging argument shows that $\phi^{(1)}$ is an $O(d^2\varepsilon)$-homomorphism of $\langle S : R \rangle$ (here the factor d^2, as opposed to d, is because of the way that the error is measured in Definition **??**). □

For ease of notation we relabel $\phi^{(1)}$ and $\varepsilon^{(1)}$ as ϕ and ε respectively. Next we exploit the relations $\{R_k^{\mathrm{HAD}}\}$ to show the following. Recall that $q = 2^t$.

Claim 4.3. *For every $u \in \mathbb{F}_q^m$ there is a projective measurement $\{P_\beta^u\}_{\beta \in \mathbb{F}_q}$ on $\mathcal{M}$ such that*

$$\mathop{\mathbb{E}}_{u \in \mathbb{F}_q^m} \mathop{\mathbb{E}}_{a \in \mathbb{F}_2^t} \left\| \phi(s_{u,a}) - \sum_{\beta \in \mathbb{F}_q} (-1)^{a \cdot \kappa(\beta)} P_\beta^u \right\|_\tau^2 = O(\varepsilon) . \tag{19}$$

Proof. Since μ_R places weight $1/4$ on relations $\{R_k^{\mathrm{HAD}}\}$ we deduce that

$$\mathop{\mathbb{E}}_{u\in\mathbb{F}_q^m}\mathop{\mathbb{E}}_{a,b\in\mathbb{F}_2^t}\left\|\phi(s_{u,a})\phi(s_{u,b})\phi(s_{u,a+b})-I\right\|_\tau^2 \le 4\varepsilon\,. \tag{20}$$

Fix an $u\in\mathbb{F}_q^m$ and apply Lemma **??** for that u. This gives a partial isometry $w\in P\mathcal{M}_\infty I_{\mathcal{M}}$ and a unitary representation $\{U_a\}$ of $\mathbb{Z}_2^t$ on $\mathcal{N}=P\mathcal{M}_\infty P$ (both depending on u) such that

$$\mathop{\mathbb{E}}_{a\in\mathbb{F}_2^t}\left\|\phi(s_{u,a})-w^*U_aw\right\|_\tau^2 = O(\varepsilon_u)\,, \tag{21}$$

where $\varepsilon_u\ge 0$ is such that $\mathbb{E}_{u\in\mathbb{F}_q^m}\varepsilon_u=4\varepsilon$. Because $\{U_a\}$ are a representation of the abelian group $\mathbb{Z}_2^t$, there is a projective measurement $\{Q_b^u\}_{b\in\mathbb{F}_2^t}$ on $\mathcal{N}$ such that $U_a=\sum_b(-1)^{a\cdot b}Q_b^u$. ($\{Q_b^u\}$ can be found explicitly by applying the Fourier transform, i.e. $Q_b^u=\sum_\alpha(-1)^{a\cdot b}U_a$.) Applying Lemma **??**, we deduce a projective measurement $\{P_\beta^u\}_{\beta\in\mathbb{F}_q}$ on $\mathcal{M}$ such that (by the triangle inequality)

$$\begin{aligned}
\mathop{\mathbb{E}}_{a\in\mathbb{F}_2^t}\Big\|\phi(s_{u,a})-\sum_{\beta\in\mathbb{F}_q}(-1)^{a\cdot\kappa(\beta)}P_\beta^u\Big\|_\tau^2 &\le 2\mathop{\mathbb{E}}_{a\in\mathbb{F}_2^t}\Big(\Big\|\phi(s_{u,a})-w^*\Big(\sum_{b\in\mathbb{F}_2^t}(-1)^{\alpha\cdot b}Q_b^u\Big)w\Big\|_\tau^2 \\
&\qquad +\Big\|\sum_{b\in\mathbb{F}_2^t}(-1)^{a\cdot b}w^*Q_b^uw-\sum_{\beta\in\mathbb{F}_q}(-1)^{a\cdot\kappa(\beta)}P_\beta^u\Big\|_\tau^2\Big) \\
&\le O(\varepsilon_u)+\sum_{\beta\in\mathbb{F}_q}\Big\|w^*Q_{\kappa(\beta)}^uw-P_\beta^u\Big\|_\tau^2 \\
&\le O(\varepsilon_u)\,,
\end{aligned}$$

where the first inequality follows from the triangle inequality, the second is by (**??**) for the first term and Parseval's identity (as in the derivation (**??**)) for the second, and the last by the guarantees obtained from Lemma **??**. Averaging over u gives the desired result. □

For $u\in\mathbb{F}_q^m$ let $\{P_\beta^u\}_{\beta\in\mathbb{F}_q}$ be the projective measurement obtained from Claim **??**. For $\alpha\in\mathbb{F}_q$, let

$$U_{u,\alpha}=\sum_{\beta\in\mathbb{F}_q}(-1)^{\mathrm{tr}(\alpha\beta)}P_\beta^u\,.$$

Then $U_{u,\alpha}\in\mathcal{U}(\mathcal{M})$.

The next claim uses the relations $\{R_k^{\mathrm{COM}}\}$.

Claim 4.4. *For $u\in\mathbb{F}_q^m$, $\alpha\in\mathbb{F}_q$, $j\in\{1,\dots,m\}$ and $i\in\{0,\dots,d\}$ let $U_{i,\alpha}=U_{u+t_ie_j,\alpha}$ and, for all $\beta\in\mathbb{F}_q$, $P_\beta^{(i)}=P_\beta^{u+t_ie_j}$ (leaving the dependence on u and j implicit). Then*

$$\mathop{\mathbb{E}}_{u\in\mathbb{F}_q^m}\mathop{\mathbb{E}}_{\substack{j\in\{1,\dots,m\}\\ i\neq i'\in\{0,\dots,d\}}}\mathop{\mathbb{E}}_{\alpha,\alpha'\in\mathbb{F}_q}\left\|\left[U_{i,\alpha},U_{i',\alpha'}\right]-I\right\|_\tau^2 = O\big(\sqrt{\varepsilon}\big)\,, \tag{22}$$

and

$$\mathop{\mathbb{E}}_{u\in\mathbb{F}_q^m}\mathop{\mathbb{E}}_{j\in\{1,\dots,m\}}\mathop{\mathbb{E}}_{\substack{v\in\mathbb{F}_q^m\\ v_{m-j+2}=u_{m-j+2},\dots,v_m=u_m}}\mathop{\mathbb{E}}_{\alpha,\alpha'\in\mathbb{F}_q}\left\|\left[U_{u,\alpha},U_{v,\alpha'}\right]-I\right\|_\tau^2 = O\big(\sqrt{\varepsilon}\big)\,, \tag{23}$$

where the expectation is over a uniformly random u and j, and a uniformly random v conditioned on its last $(j-1)$ coordinates matching those of u, and $[U,V]=UVU^*V^*$ is the group commutator.

Proof. Due to the test of the relations $\{R_k^{\mathrm{COM}}\}$, it holds that

$$\mathop{\mathbb{E}}_{u\in\mathbb{F}_q^m}\ \mathop{\mathbb{E}}_{\substack{j\in\{1,\ldots,m\}\\ i\neq i'\in\{0,\ldots,d\}}}\ \mathop{\mathbb{E}}_{a,b\in\mathbb{F}_2^t}\left\|[\phi(s_{u+t_ie_j,a}),\phi(s_{u+t_{i'}e_j,b})]-I\right\|_\tau^2\le 8\varepsilon\ , \tag{24}$$

$$\mathop{\mathbb{E}}_{u\in\mathbb{F}_q^m}\ \mathop{\mathbb{E}}_{j\in\{1,\ldots,m\}}\ \mathop{\mathbb{E}}_{\substack{v\in\mathbb{F}_q^m\\ v_{m-j+2}=u_{m-j+2},\ldots,v_m=u_m}}\ \mathop{\mathbb{E}}_{a,b\in\mathbb{F}_2^t}\left\|[\phi(s_{u,a}),\phi(s_{v,b})]-I\right\|_\tau^2\le 8\varepsilon\ . \tag{25}$$

Now, for any $u,v\in\mathbb{F}_q^m$

$$\begin{aligned}
&\mathop{\mathbb{E}}_{\alpha,\alpha'\in\mathbb{F}_q}\tau\big(U_{u,\alpha}U_{v,\alpha'}U_{u,-\alpha}U_{v,-\alpha'}\big)\\
&=\sum_{\beta,\beta',\gamma,\gamma'\in\mathbb{F}_q}\mathop{\mathbb{E}}_{\alpha\in\mathbb{F}_q}(-1)^{\mathrm{tr}(\alpha(\beta-\beta'))}\mathop{\mathbb{E}}_{\alpha'\in\mathbb{F}_q}(-1)^{\mathrm{tr}(\alpha'(\gamma-\gamma'))}\tau\big(P_\beta^uP_\gamma^vP_{\beta'}^uP_{\gamma'}^v\big)\\
&=\sum_{\beta,\beta',\gamma,\gamma'\in\mathbb{F}_q}\mathop{\mathbb{E}}_{a\in\mathbb{F}_2^t}(-1)^{a\cdot\kappa(\beta-\beta')}\mathop{\mathbb{E}}_{a'\in\mathbb{F}_2^t}(-1)^{a'\cdot\kappa(\gamma-\gamma')}\tau\big(P_\beta^uP_\gamma^vP_{\beta'}^uP_{\gamma'}^v\big)\\
&=\mathop{\mathbb{E}}_{a,a'\in\mathbb{F}_2^t}\tau\Big(\Big(\sum_{\beta\in\mathbb{F}_q}(-1)^{a\cdot\kappa(\beta)}P_\beta^u\Big)\Big(\sum_{\gamma\in\mathbb{F}_q}(-1)^{a'\cdot\kappa(\gamma)}P_\gamma^v\Big)\Big(\sum_{\beta'\in\mathbb{F}_q}(-1)^{a\cdot\kappa(\beta')}P_{\beta'}^u\Big)\Big(\sum_{\gamma'\in\mathbb{F}_q}(-1)^{a'\cdot\kappa(\gamma')}P_{\gamma'}^v\Big)\Big)\ .
\end{aligned}$$

It follows that, for any distribution on (u,v) such that both marginals are uniform over $\mathbb{F}_q^m$,

$$\begin{aligned}
\mathop{\mathbb{E}}_{u,v}\mathop{\mathbb{E}}_{\alpha,\alpha'\in\mathbb{F}_q}\left\|[U_{u,\alpha},U_{v,\alpha'}]-I\right\|_\tau^2&=\mathop{\mathbb{E}}_{u,v}\mathop{\mathbb{E}}_{a,a'\in\mathbb{F}_2^t}\left\|\Big[\sum_{\beta\in\mathbb{F}_q}(-1)^{a\cdot\kappa(\beta)}P_\beta^u,\sum_{\beta'\in\mathbb{F}_q}(-1)^{a'\cdot\kappa(\beta')}P_{\beta'}^v\Big]-I\right\|_\tau^2\\
&=\mathop{\mathbb{E}}_{u,v}\mathop{\mathbb{E}}_{a,a'\in\mathbb{F}_2^t}\left\|\big[\phi(s_{u,a}),\phi(s_{v,a'})\big]-I\right\|_\tau^2+O\big(\sqrt{\varepsilon}\big)\ ,
\end{aligned} \tag{26}$$

where the first line is by expanding the square and using (**??**) and the second line follows from Claim **??** and the triangle inequality. Thus (**??**) follows from (**??**), and (**??**) follows from (**??**). □

Now we show the following, which essentially follows from the previous claim.

Claim 4.5. *For every $u\in\mathbb{F}_q^m$ and $j\in\{1,\ldots,m\}$, there is a projective measurement $\{R_\alpha^{u,j}\}_{\alpha\in\mathbb{F}_q^{d+1}}$ on $\mathcal{M}$ such that*

$$\mathop{\mathbb{E}}_{u\in\mathbb{F}_q^m}\ \mathop{\mathbb{E}}_{j\in\{1,\ldots,m\}}\ \mathop{\mathbb{E}}_{z_0,\ldots,z_d\in\mathbb{F}_q}\left\|U_{0,z_0}\cdots U_{d,z_d}-\sum_\alpha(-1)^{\mathrm{tr}(\sum z_i\alpha_i)}R_\alpha^{u,j}\right\|_\tau^2\ =\ \mathrm{poly}(d,\varepsilon)\ .$$

Similarly, for any $u,v\in\mathbb{F}_q^m$ there is a projective measurement $\{R_{\alpha,\beta}^{u,v}\}_{(\alpha,\beta)\in\mathbb{F}_q^2}$ on $\mathcal{M}$ such that

$$\mathop{\mathbb{E}}_{u\in\mathbb{F}_q^m}\ \mathop{\mathbb{E}}_{j\in\{1,\ldots,m\}}\ \mathop{\mathbb{E}}_{\substack{v\in\mathbb{F}_q^m\\ v_{m-j+2}=u_{m-j+2},\ldots,v_m=u_m}}\ \mathop{\mathbb{E}}_{y,z\in\mathbb{F}_q}\left\|U_{u,z}U_{v,y}-\sum_\alpha(-1)^{\mathrm{tr}(\alpha z+\beta y)}R_{\alpha,\beta}^{u,v}\right\|_\tau^2\ =\ \mathrm{poly}(\varepsilon)\ .$$

Proof. We show the first part only, as the second part is analogous. Fix an $u\in\mathbb{F}_q^m$ and a direction $j\in\{1,\ldots,m\}$. Define $\psi:(\mathbb{Z}_2^t)^{d+1}\to\mathcal{U}(\mathcal{M})$ by

$$\psi(z_0,\ldots,z_d)\ =\ U_{0,z_0}\cdots U_{d,z_d}\ ,$$

where for each $i \in \{0,\dots,d\}$ and $z_i \in \mathbb{F}_q$, U_{i,z_i} is the unitary defined in Claim **??** and we slightly abused notation to identify $z_i \in \mathbb{Z}_2^t$ with the unique $\tilde{z}_i \in \mathbb{F}_q$ such that $\kappa(\tilde{z}_i) = z_i$. Using Claim **??** and the triangle inequality we verify that the map ψ is a $\delta = O(d^4\sqrt{\varepsilon})$-approximate homomorphism of $\mathbb{Z}_2^{t(d+1)}$ on $\mathcal{U}(\mathcal{A})$, with respect to the multiplication table presentation. To verify this first note that for any unitaries V, W, and any i, i' we have that on average over u and j,

$$\mathop{\mathbb{E}}_{\alpha,\alpha' \in \mathbb{F}_q} \left\| V U_{i,\alpha} U_{i',\alpha'} W^* - V U_{i',\alpha'} U_{i,\alpha} W^* \right\|_\tau^2 = O\big(d^2\sqrt{\varepsilon}\big) ,$$

by Claim **??**, where the factor d^2 is because we require the relation to hold for all i, i'. Moreover, $U_{i,\alpha} U_{i,\alpha'} = U_{i,\alpha+\alpha'}$ by definition. Applying these relation $O(d^2)$ times gives

$$\mathop{\mathbb{E}}_{z_0,\dots,z_d \in \mathbb{F}_2^t} \mathop{\mathbb{E}}_{y_0,\dots,y_d \in \mathbb{F}_2^t} \left\| \psi(z_0,\dots,z_d)\psi(y_0,\dots,y_d) - \psi(z_0+y_0,\dots,z_d+y_d) \right\|_\tau^2 = O\big(d^4\sqrt{\varepsilon}\big) ,$$

as desired. Thus we may apply Theorem **??** to obtain a partial isometry $w \in P\mathcal{M}_\infty I_{\mathcal{M}}$ and a family of commuting unitaries $\{V_{i,k}\}$ of order 2 each on $\mathcal{N} = P\mathcal{M}_\infty P$ such that, on average over u and j, defining $V_i^z = \prod_{k=1}^t V_{i,k}^{z_k}$,

$$\mathop{\mathbb{E}}_{z_0,\dots,z_d \in \mathbb{F}_2^t} \left\| \psi(z_0,\dots,z_{d+1}) - w^* V_0^{z_0} \cdots V_d^{z_d} w \right\|_\tau^2 = \mathrm{poly}(d,\varepsilon) . \tag{27}$$

Because $\{V_{i,z}\}$ are a representation of $\mathbb{Z}_2^{t(d+1)}$ (which is Abelian), there is a projective measurement $\{Q_\alpha\}_{\alpha \in \mathbb{Z}_2^{t(d+1)}}$ on $\mathcal{N}$ such that for each $i \in \{0,\dots,d\}$,

$$V_i^z = \sum_{\alpha \in (\mathbb{Z}_2^t)^{d+1}} (-1)^{\alpha_i \cdot z} Q_\alpha ,$$

and

$$V_0^{z_0} \cdots V_d^{z_d} = \sum_{\alpha \in (\mathbb{Z}_2^t)^{d+1}} (-1)^{\sum \alpha_i \cdot z_i} Q_\alpha .$$

Identifying α with an element of $\mathbb{F}_q^{d+1}$, and interpreting z_i as an element of $\mathbb{F}_q$ as well, the preceding equation can be rewritten as

$$V_0^{z_0} \cdots V_d^{z_d} = \sum_{\alpha \in \mathbb{F}_q^{d+1}} (-1)^{\mathrm{tr}(\sum \alpha_i z_i)} Q_\alpha .$$

Applying Lemma **??**, we deduce a projective measurement $\{R_\alpha^{u,j}\}_{\alpha \in \mathbb{F}_q^{d+1}}$ on $\mathcal{M}$ such that by (**??**), on average over u and j,

$$\mathop{\mathbb{E}}_{z_0,\dots,z_d \in \mathbb{F}_q} \left\| \psi(z_0,\dots,z_d) - \sum_{\alpha \in \mathbb{F}_q^{d+1}} (-1)^{\mathrm{tr}(\sum z_i \alpha_i)} R_\alpha^{u,j} \right\|_\tau^2 = \mathrm{poly}(d,\varepsilon) .$$

□

Finally we exploit the relations $\{R_k^{\mathrm{LD}}\}$ to obtain the following.

Claim 4.6. *Use the same notation as in Claim* **??**. *Let $u \in \mathbb{F}_q^m$, ℓ an axis-parallel line through u and $v \in \ell$. Let $(\alpha_{u,v,i})_{i=0,\ldots,d}$ be the interpolation coefficients defined in* (**??**). *Then*

$$\mathop{\mathbb{E}}_{u\in\mathbb{F}_q^m}\mathop{\mathbb{E}}_{\ell:u\in\ell}\mathop{\mathbb{E}}_{v\in\ell}\mathop{\mathbb{E}}_{\gamma\in\mathbb{F}_q}\left\|U_{0,\gamma\alpha_{u,v,0}}\cdots U_{d,\gamma\alpha_{u,v,d}}-U_{v,\gamma}\right\|_\tau^2 \;=\; O\big(d\sqrt{\varepsilon}\big)\;, \tag{28}$$

where the expectation is over a uniformly random axis-parallel line that contains u, and a uniformly random $v \in \ell$.

Proof. Write α_i for $\alpha_{u,v,i}$. We observe that

$$\begin{aligned}
\mathop{\mathbb{E}}_{\gamma\in\mathbb{F}_q}\tau\big(U_{0,\gamma\alpha_0}\cdots U_{d,\gamma\alpha_d}U_{v,-\gamma}\big) &= \sum_{\beta_0,\ldots,\beta_d\in\mathbb{F}_q}\sum_{\beta\in\mathbb{F}_q}\mathop{\mathbb{E}}_{\gamma\in\mathbb{F}_q}(-1)^{\mathrm{tr}(\sum\gamma\alpha_i\beta_i)}(-1)^{\mathrm{tr}(-\gamma\beta)}\tau\big(P^{(0)}_{\beta_0}\cdots P^{(d)}_{\beta_d}P^v_\beta\big)\\
&= \sum_{\beta_0,\ldots,\beta_d\in\mathbb{F}_q}\sum_{\beta\in\mathbb{F}_q}\mathop{\mathbb{E}}_{\gamma\in\mathbb{F}_2^t}(-1)^{\gamma\cdot\kappa(\sum\alpha_i\beta_i-\beta)}\tau\big(P^{(0)}_{\beta_0}\cdots P^{(d)}_{\beta_d}P^v_\beta\big)\\
&= \mathop{\mathbb{E}}_{\gamma\in\mathbb{F}_2^t}\tau\Big(\Big(\sum_{\beta_0\in\mathbb{F}_q}(-1)^{\gamma\cdot\kappa(\alpha_0\beta_0)}P^{(0)}_{\beta_0}\Big)\cdots\Big(\sum_{\beta\in\mathbb{F}_q}(-1)^{-\gamma\cdot\kappa(\beta)}P^v_\beta\Big)\Big)\\
&= \mathop{\mathbb{E}}_{\gamma\in\mathbb{F}_q}\tau\Big(\Big(\sum_{\beta_0\in\mathbb{F}_q}(-1)^{\kappa(\gamma\alpha_0)\cdot\kappa(\beta_0)}P^{(0)}_{\beta_0}\Big)\cdots\Big(\sum_{\beta\in\mathbb{F}_q}(-1)^{\kappa(-\gamma)\cdot\kappa(\beta)}P^v_\beta\Big)\Big)\;.
\end{aligned}$$

By repeated application of Claim **??**, the last expression is within $O(d\sqrt{\varepsilon})$ of

$$\mathop{\mathbb{E}}_{\gamma\in\mathbb{F}_q}\tau\big(\phi(s_{u+t_ie_j,\kappa(\gamma\alpha_0)})\cdots\phi(s_{v,\kappa(-\gamma)})\big)\;.$$

The latter expression is a random relation in R^{LD}_k, and so, on average over u, j and v it is $1-O(\varepsilon)$. □

Claim 4.7. *For every axis-parallel line ℓ there is a projective measurement $\{Q^\ell_g\}$ with outcomes $g \in \mathcal{P}(q,1,d)$ that range over degree-d polynomials on ℓ such that*

$$\mathop{\mathbb{E}}_{\ell\subset\mathbb{F}_q^m}\mathop{\mathbb{E}}_{v\in\ell}\sum_g\tau\big(P^v_{g(v)}Q^\ell_g\big)\;\geq\;1-\mathrm{poly}(d,\varepsilon)\;,$$

where the expectation is over a uniformly random axis-parallel line ℓ and point $v \in \ell$.

Proof. First we show that the conclusion of Claim **??** can be strengthened to hold for *every* $z_0,\ldots,z_d$, instead of on average. To show this, note that for any $y_0,\ldots,y_d \in \mathbb{F}_q$ we can write

$$U_{0,z_0}\cdots U_{d,z_d}\;=\;U_{0,y_0}U_{0,z_0-y_0}\cdots U_{d,y_d}U_{d,z_d-y_d}\;.$$

By repeated application of Claim **??**, the term on the right-hand side satisfies

$$\big\|U_{0,y_0}U_{0,z_0-y_0}\cdots U_{d,y_d}U_{d,z_d-y_d}-U_{0,y_0}\cdots U_{d,y_d}U_{0,z_0-y_0}\cdots U_{d,z_d-y_d}\big\|_\tau^2\;=\;\mathrm{poly}(d,\varepsilon)\;.$$

To show this it suffices to verify that we only need to "commute" pairs of terms whose exponents are independent and uniformly random. Using Claim **??** twice, the right-hand side satisfies

$$\big\|U_{0,y_0}\cdots U_{d,y_d}U_{0,z_0-y_0}\cdots U_{d,z_d-y_d}-\Big(\sum_\alpha\omega^{\mathrm{tr}(\sum y_i\alpha_i)}R^{u,j}_\alpha\Big)\Big(\sum_\alpha(-1)^{\mathrm{tr}(\sum(z_i-y_i)\alpha_i)}R^{u,j}_\alpha\Big)\big\|_\tau^2\;=\;\mathrm{poly}(d,\varepsilon)\;.$$

Since $\{R_\alpha^{u,j}\}$ is a projective measurement, we get the desired conclusion: on average over u and j, for any $z_0, \ldots, z_d \in \mathbb{F}_q$, it holds that

$$\left\| U_{0,z_0} \cdots U_{d,z_d} - \sum_\alpha (-1)^{\mathrm{tr}(\sum z_i \alpha_i)} R_\alpha^{u,j} \right\|_\tau^2 = \mathrm{poly}(d,\varepsilon) \ . \tag{29}$$

For any line $\ell \subset \mathbb{F}_q^m$, let $u_\ell \in \ell$ be chosen such that, for a uniformly random ℓ and conditioned on that u_ℓ, (**??**) and (**??**) both hold, with right-hand side multiplied by a factor at most 2. For any ℓ in the j-th direction and degree-d polynomial g, define the operator $Q_g^\ell = R^{u_\ell,j}_{g(u_\ell+t_0e_j),\ldots,g(u_\ell+t_de_j)}$. Combining the two equations we deduce

$$\mathop{\mathbb{E}}_{\gamma\in\mathbb{F}_q} \left\| U_{v,\gamma} - \sum_g (-1)^{\mathrm{tr}(\gamma \sum \alpha_{u_\ell,v,i} g(u_\ell+t_ie_j))} Q_g^\ell \right\|_\tau^2 = \mathrm{poly}(d,\varepsilon) \ . \tag{30}$$

By definition, $\sum \alpha_{u_\ell,v,i} g(u_\ell + t_i e_j) = g(v)$. By Fourier transform, we obtain the desired conclusion. □

We are now in a position to apply [**?**, Theorem 4.1]. For this we need to define a synchronous strategy in the tensor code test $\mathscr{C}^{\otimes m}$, where $\mathscr{C}$ is the Reed-Solomon code with degree d over $\mathbb{F}_q$, and thus $\mathscr{C}^{\otimes m}$ is the code $\mathscr{C}_{\mathrm{RM}}$ considered in Section **??** (see the remark right after Lemma **??**). For the "points measurement" A^u we choose P^u. For the "lines measurement" B^ℓ we choose Q^ℓ from Claim **??**. Finally, for the "pair measurement" $P^{u,v}$ we choose $R^{u,v}$ from Claim **??**. By Claim **??** this strategy succeeds with probability $1 - \mathrm{poly}(d,\varepsilon)$ in the "axis-parallel lines test", and by Claim **??** it succeeds with probability $1 - \mathrm{poly}(\varepsilon)$ in the "subcube commutation test." Applying [**?**, Theorem 4.1] we deduce the existence of a projective measurement $\{G_c\}_{c\in\mathscr{C}^{\otimes m}}$ on $\mathcal{A}$ such that for all integers $r \geq 12mt$,

$$\mathop{\mathbb{E}}_{u\in\mathbb{F}_q^m} \sum_{c\in\mathscr{C}^{\otimes m}} \tau\big(G_c P^u_{c(u)}\big) \geq 1 - \eta \ ,$$

where $\eta = \mathrm{poly}(m,d,r) \cdot \mathrm{poly}(\varepsilon, q^{-1}, e^{-\Omega(r/m^2)})$. Choosing $r = \Omega(m^2 t)$, since $q = 2^t$, the bound becomes $\mathrm{poly}(m,d,t) \cdot \mathrm{poly}(\varepsilon, q^{-1})$.

□

5 Testing entanglement

In this section we give an application of our stability results to the problem of entanglement testing in quantum information. We first introduce the language of nonlocal games. Then we associate a nonlocal game to any presentation of $\mathbb{Z}_2^k$. Finally, we show that, if the presentation is stable, then the game is a robust entanglement test.

5.1 Nonlocal games

We give standard definitions on nonlocal games. For background from a computer science point of view, see [**?**]; for the operator algebra perspective, see e.g. [**?**].

Definition 5.1 (Game)**.** A game is a tuple $(\mathcal{X}, \mu, \mathcal{A}, D)$ where $\mathcal{X}$ is a finite set, μ a distribution on $\mathcal{X} \times \mathcal{X}$, $\mathcal{A} = (\mathcal{A}(x))_{x\in\mathcal{X}}$ a collection of finite sets, and

$$D : \big\{(x,y,a,b) : (x,y) \in \mathrm{supp}(\mu), a \in \mathcal{A}(x), b \in \mathcal{A}(y)\big\} \ \to \ \{0,1\}$$

such that D is symmetric, i.e. $D(x,y,a,b) = D(y,x,b,a)$ whenever both terms are defined. We often abuse notation and write μ for the symmetrized marginal of μ, i.e.

$$\mu(x) := \sum_{x' \in \mathcal{X}} \frac{1}{2}\big(\mu(x,x') + \mu(x',x')\big) \ .$$

The interpretation of $G = (\mathcal{X}, \mu, \mathcal{A}, D)$ as a nonlocal game is the following. In the "game," a referee is imagined to sample a pair of "questions" $(x,y) \sim \mu$. The question x is sent to a first player, "Alice," and the question y is sent to a second player, "Bob." Each player is tasked with responding with an answer, $a \in \mathcal{A}(x)$ for Alice and $b \in \mathcal{A}(y)$ for Bob. The referee accepts the players' answers if and only if $D(x,y,a,b) = 1$.

Nonlocal games provide a framework to study different kinds of bipartite correlations: depending on the level of coordination allowed between Alice and Bob, they may have varying chances of success in the game. A "classical" strategy consists of functions $f_A : \mathcal{X} \to \mathcal{A}$ for Alice and $f_B : \mathcal{X} \to \mathcal{A}$ for Bob; any such pair of functions leads to a probability of success in the game which can be computed in the obvious manner.

An important motivation for studying nonlocal games is that in quantum mechanics, local strategies for the players (meaning strategies that do not require any communication between the players to determine their answer, given their question) are a larger set than the above-described classical strategies. Specifically, a *quantum local* (or *quantum* for short) strategy is specified by the following.

Definition 5.2 (Synchronous strategy)**.** If $G = (\mathcal{X}, \mu, \mathcal{A}, D)$ is a game and $(\mathcal{M}, \tau)$ a tracial von Neumann algebra, a *synchronous strategy $\mathscr{S}$ for G on $(\mathcal{M}, \tau)$* is, for every $x \in \mathcal{X}$, a projective measurement $(P_a^x)_{a \in \mathcal{A}(x)}$ on $\mathcal{M}$. The value of a strategy $\mathscr{S}$ in G is

$$\omega(G;\mathscr{S}) = \sum_{(x,y) \in \mathcal{X} \times \mathcal{X}} \frac{1}{2}\big(\mu(x,y) + \mu(y,x)\big) \sum_{(a,b) \in \mathcal{A}(x) \times \mathcal{A}(y)} D(x,y,a,b)\, \tau\big(P_a^x\, P_b^y\big) \ .^{14}$$

We say that $\mathscr{S}$ is *perfect* if $\omega(G;\mathscr{S}) = 1$.

The name *synchronous* stems from the fact that whenever an identical pair (x,x) is chosen, $\tau(P_a^x P_b^x) = 0$ for $a \neq b$ due to the requirement that $\{P_a^x\}_a$ is a projective measurement. Thus a synchronous strategy always returns the same answer to the same question. More general strategies, which allow different operators $\{P_a^x\}$ and $\{Q_b^y\}$, do not automatically enforce the synchronicity condition, but we do not consider such strategies here.

It will be convenient to have a measure of closeness for strategies. The following definition parallels Definition **??**.

Definition 5.3 (Closeness for strategies)**.** Let $\{A_a^i\} \subseteq \mathcal{M}$ and $\{B_a^i\} \subseteq \mathcal{N}$ be two families of projective measurements on tracial algebras $(\mathcal{M}, \tau^{\mathcal{M}})$ and $(\mathcal{N}, \tau^{\mathcal{N}})$ respectively, indexed by the same set $i \in \mathcal{I}$ and with the same sets of outcomes $a, b \in \mathcal{A}(i)$. For $\delta \geq 0$ and μ a measure on $\mathcal{I}$ we say that $\{A^i\}$ and $\{B^i\}$ are (δ,μ)-close if there exists a projection $P \in \mathcal{M}_\infty$ of finite trace such that $\mathcal{N} = P\mathcal{M}_\infty P$ and $\tau^{\mathcal{N}} = \tau_\infty / \tau_\infty(P)$, and a partial isometry $w \in P\mathcal{M}_\infty I_{\mathcal{M}}$ such that

$$\mathop{\mathbb{E}}_{i \sim \mu} \sum_{a \in \mathcal{A}(i)} \big\| A_a^i - w^* B_a^i w \big\|_{\tau^{\mathcal{M}}}^2 \leq \delta$$

[14] Note the symmetrization of μ. This is to avoid explicitly requiring μ to be permutation-invariant in the definition of a game. We will often abuse notation and write $\mathbb{E}_{(x,y)\sim\mu}$ when taking expectations under the symmetrized distribution.

and

$$\max\left\{\tau^{\mathcal{M}}(I_{\mathcal{M}} - w^*w)\,,\ \tau^{\mathcal{N}}(P - ww^*)\right\} \ \leq\ \delta\,.$$

If the measure μ is omitted then it is understood to be the uniform measure on $\mathcal{I}$.

In Definition **??** closeness is measured in the L_2 sense. The following lemma gives a consequence for distance measured in an L_1 sense.

Lemma 5.4. *Let $\{P_a^i\}$ and $\{Q_a^i\}$ be two families of projective measurements that are (ε,μ)-close. Then*

$$\mathop{\mathbb{E}}_{i\in\mathcal{I}} \sum_{a\in\mathcal{A}(i)} \tau\big(|P_a - w^*Q_a w|\big) \ \leq\ \varepsilon + 2\sqrt{\varepsilon}\,. \tag{31}$$

Proof. Using the triangle inequality,

$$\begin{aligned}\mathop{\mathbb{E}}_i \sum_a \tau\big(|P_a^i - w^*Q_a^i w|\big) \leq\ \mathop{\mathbb{E}}_i \sum_a \Big(&\tau\big(|(P_a^i - (P_a^i)^2)|\big) + \tau\big(|P_a(P_a^i - w^*Q_a^i w)|\big)\\ &+ \tau\big(|(P_a^i - w^*Q_a^i w)w^*Q_a^i w|\big)\\ &+ \tau\big(|(w^*Q_a^i w^*wQ_a^i w - w^*Q_a^i w)|\big)\Big)\,.\end{aligned} \tag{32}$$

The first term on the right-hand side is zero, because P^i is assumed projective. The terms in the middle are bounded using Hölder's inequality:

$$\begin{aligned}\mathop{\mathbb{E}}_i \sum_a \tau\big(|P_a^i(P_a^i - w^*Q_a^i w)|\big) &\leq \mathop{\mathbb{E}}_i \sum_a \|P_a^i\|_\tau\ \|P_a^i - w^*Q_a^i w\|_\tau\\ &\leq \Big(\mathop{\mathbb{E}}_i \sum_a \|P_a^i\|_2^\tau\Big)^{1/2}\Big(\mathop{\mathbb{E}}_i \sum_a \|P_a^i - w^*Q_a^i w\|_\tau^2\Big)^{1/2}\\ &\leq \sqrt{\varepsilon}\end{aligned}$$

by closeness. The third term of (**??**) is bounded in a similar fashion. Finally the last term of (**??**) can be bounded as

$$\begin{aligned}\mathop{\mathbb{E}}_i \sum_a \tau\big(|w^*Q_a^i w^*wQ_a^i w - w^*Q_a^i w|\big) &= \mathop{\mathbb{E}}_i \sum_a \tau\big((w^*Q_a^i(I - w^*w)Q_a^i w)\big)\\ &= \tau\Big((I - w^*w)\Big(\mathop{\mathbb{E}}_i \sum_a Q_a^i ww^*Q_a^i\Big)\Big)\\ &\leq \tau(I - w^*w)\ \Big\|\mathop{\mathbb{E}}_i \sum_a Q_a^i ww^*Q_a^i\Big\|\\ &\leq \varepsilon\,.\end{aligned}$$

□

We will make use of the following elementary lemma, which shows that two strategies for the same game G that are close according to Definition **??** have a close value.

Lemma 5.5. *Let $G = (\mathcal{X},\mu,\mathcal{A},D)$ be a game and $\mathscr{S} = \{P_a^x\}$ and $\mathscr{S}' = \{Q_a^x\}$ strategies on $\mathcal{M}$ and $\mathcal{N}$ respectively such that $\{P_a^x\}$ and $\{Q_a^x\}$ are (ε,μ)-close. Then*

$$\big|\omega(G;\mathscr{S}) - \omega(G;\mathscr{S}')\big| \ \leq\ 8\sqrt{\varepsilon}\,.$$

Proof. By definition, there exists a projection $P \in \mathcal{M}_\infty$ and a partial isometry $w \in P\mathcal{M}_\infty I_{\mathcal{M}}$ such that $\mathcal{N} = P\mathcal{M}_\infty P$ and

$$\begin{aligned}
\omega(G;\mathscr{S}') &= \mathop{\mathbb{E}}_{(x,y)\sim\mu} \sum_{a,b} D(a,b,x,y)\tau\big(Q_a^x\, Q_b^y\big) \\
&= \mathop{\mathbb{E}}_{(x,y)\sim\mu} \sum_{a,b} D(a,b,x,y)\tau\big(w^* Q_a^x w\, w^* Q_b^y w\big) \\
&\qquad + \mathop{\mathbb{E}}_{(x,y)\sim\mu} \sum_{a,b} D(a,b,x,y)\tau\big(w^* Q_a^x (P - w\,w^*) Q_b^y w\big) \\
&\qquad + \mathop{\mathbb{E}}_{(x,y)\sim\mu} \sum_{a,b} D(a,b,x,y)\tau\big(Q_a^x Q_b^y (P - ww^*)\big)\,. \qquad (33)
\end{aligned}$$

Here we used the fact that $Q_a^x \in \mathcal{N}$ so $PQ_a^xP = Q_a^x$ for all x, a. The second term on the right-hand side can be bounded as

$$\begin{aligned}
&\Big| \mathop{\mathbb{E}}_{(x,y)\sim\mu} \sum_{a,b} D(a,b,x,y)\tau\big(w^* Q_a^x (P - w\,w^*) Q_b^y w\big)\Big| \\
&\qquad \le \mathop{\mathbb{E}}_{(x,y)\sim\mu} \Big|\tau\big(\sum_{a,b} D(a,b,x,y) Q_b^y w w^* Q_a^x (P - w\,w^*)\big)\Big| \\
&\qquad \le \sqrt{\tau((P-ww^*)^2)} \cdot \sqrt{\mathop{\mathbb{E}}_{(x,y)\sim\mu} \tau\big(\big|\sum_{a,b} D(a,b,x,y) Q_a^x w w^* Q_b^y\big|^2\big)} \\
&\qquad \le \sqrt{\varepsilon} \sqrt{\mathop{\mathbb{E}}_{(x,y)\sim\mu} \tau\big(\sum_{a,b} D(a,b,x,y) Q_b^y w w^* Q_a^x w w^* Q_b^y\big)} \\
&\qquad \le \sqrt{\varepsilon} \sqrt{\mathop{\mathbb{E}}_{(x,y)\sim\mu} \sum_{a,b} \tau\big(Q_b^y w w^* Q_a^x w w^* Q_b^y\big)} \\
&\qquad = \sqrt{\varepsilon} \sqrt{\mathop{\mathbb{E}}_{(x,y)\sim\mu} \sum_{b} \tau\big(Q_b^y w w^*\big)} \\
&\qquad \le \sqrt{\varepsilon}\,. \qquad (34)
\end{aligned}$$

The third line is due to Cauchy-Schwarz and Jensen's inequality, where $|A|^2 = A^*A$. The fourth line uses that $P - ww^*$ is a positive operator with operator norm at most 1 for the first term, and that for fixed x, y, the measurements $\{Q_a^x\}_a$ and $\{Q_b^y\}_b$ are projective. The fifth line is due to $D(x,y,a,b) \in \{0,1\}$. The sixth line is due to $\sum_a Q_a^x = I_{\mathcal{N}}$. The seventh line is due to $\sum_a Q_b^y = I_{\mathcal{N}}$. The third time on the right-hand side of (**??**) is bounded in the same manner.

For $x \in \mathcal{X}$ such that $\mu(x) \neq 0$, denote by μ_x the (symmetrized) conditional distribution $\mu_x(y) = \frac{1}{2}(\mu(x,y) + \mu(y,x))/\mu(x)$. Then using (**??**) and the preceding bounds,

$$\begin{aligned}
\big|\omega(G;\mathscr{S}) - \omega(G;\mathscr{S}')\big| &\le 2\sqrt{\varepsilon} + \Big| \mathop{\mathbb{E}}_{(x,y)\sim\mu} \sum_{a,b} D(a,b,x,y)\big(\tau\big(P_a^x\, P_b^y\big) - \tau\big(w^* Q_a^x w\, w^* Q_b^y w\big)\big)\Big| \\
&\le 2\sqrt{\varepsilon} + \Big| \mathop{\mathbb{E}}_{x\sim\mu} \sum_a \tau\Big(\big(P_a^x - w^* Q_a^x w\big)\Big(\mathop{\mathbb{E}}_{y\sim\mu_x} \sum_b D(a,b,x,y) P_b^y\Big)\Big)\Big| \\
&\qquad + \Big| \mathop{\mathbb{E}}_{y\sim\mu} \sum_b \tau\Big(\Big(\mathop{\mathbb{E}}_{x\sim\mu_y} \sum_a D(a,b,x,y) w^* Q_a^x w\Big)\big(P_b^y - w^* Q_b^y w\big)\Big)\Big| \\
&\le 2\sqrt{\varepsilon} + 2 \mathop{\mathbb{E}}_{x\sim\mu} \sum_a \tau\big(\big|P_a^x - w^* Q_a^x w\big|\big) \\
&\le 2\sqrt{\varepsilon} + 2(\varepsilon + 2\sqrt{\varepsilon}) \le 8\sqrt{\varepsilon}\,,
\end{aligned}$$

where the third line uses $\tau(AB) \leq \tau(|A|)\|B\|$ and the last line is by Lemma **??**. □

5.2 Some simple games

We introduce games previously used in the literature, that will serve as building blocks. For details on the implementation of these games we refer to [**?**].

Commutation game. We call *commutation game* the game $G_{\text{com}} = (\mathcal{X}_{\text{com}}, \mu_{\text{com}}, \mathcal{A}_{\text{com}}, D_{\text{com}})$ defined in [**?**, Section 3.1]. For convenience we change the notation slightly and denote $x_{X,0}, x_{Z,0} \in \mathcal{X}_{\text{com}}$ the two special questions, $x_{\text{com},1}$ and $x_{\text{com},2}$ respectively.

Anti-commutation game. We call *anti-commutation game* the game $G_{\text{ac}} = (\mathcal{X}_{\text{ac}}, \mu_{\text{ac}}, \mathcal{A}_{\text{ac}}, D_{\text{ac}})$ defined in [**?**, Section 3.2]. For convenience we change the notation slightly and denote $x_{X,1}, x_{Z,1} \in \mathcal{X}_{\text{ac}}$ the two special questions, $x_{\text{ac},1}$ and $x_{\text{ac},2}$ respectively.

Braiding game. Finally we introduce a game, which we denote G_{dlS}, that is built from the previous two games and is a based on a more general class of games analyzed in [**?**, Section 3.4]. Using notation from [**?**], the game G_{dlS} is obtained by making the following choices. The game can be constructed from any $E \in \mathbb{F}_2^{k\times n}$. Let

$$S_X \;=\; S_Z \;=\; \{Ee_i : \, i \in \{1,\ldots,n\}\} \subseteq \mathbb{F}_2^k \,, \tag{35}$$

and let μ_{dlS} be the uniform distribution over $\Omega = S_X \times S_Z$. Let α, β be the coordinate projections on Ω. Let $\Omega_+ = \{(a,b) \in \Omega : a \cdot b = 0\}$ and $\Omega_- = \{(a,b) \in \Omega : a \cdot b = 1\}$. Then $G_{\text{dlS}} = (\mathcal{X}_{\text{dlS}}, \mu_{\text{dlS}}, \mathcal{A}_{\text{dlS}}, D_{\text{dlS}})$ has question set

$$\mathcal{X}_{\text{dlS}} = \{X, Z\} \cup (\mathcal{X}_{\text{com}} \times \Omega_+) \cup (\mathcal{X}_{\text{ac}} \times \Omega_-) \cup (\{X\} \times S_X) \cup (\{Z\} \times S_Z) \,.$$

The game itself is based on the game described in [**?**, Section 3.4], with a small modification (the addition of the consistency test). For clarity we recall the entire game in Figure **??**.

Let $S_X, S_Z \subseteq \mathbb{F}_2^k$. Sample $\omega = (\omega_X, \omega_Z) \in S_X \times S_Z$ uniformly at random. Let $\gamma = \omega_X \cdot \omega_Z \in \mathbb{F}_2$. Execute either of the following tests with probability $1/3$ each.

1. **(Anti-)commutation test:**
   (a) If $\gamma = 0$ then sample a pair of questions $(x_c, y_c) \sim \mu_{\text{com}}$ as in the commutation game. If $x_c = x_{W,0}$ (resp. $y_c = x_{W,0}$) for some $W \in \{X, Z\}$ then send (W, ω_W) to Alice (resp. (W, ω_W) to Bob). Otherwise, send (x_c, ω) to A (resp. (y_c, ω) to Bob). Accept if and only if their answers are accepted in the commutation game.
   (b) If $\gamma \neq 0$ then do the same but for the anti-commutation game.
2. **Consistency test:** Select $W \in \{X, Z\}$ uniformly at random. Send W to A and (W, ω_W) to B. Receive $a \in \mathbb{F}_2^k$ and $b \in \mathbb{F}_2$ respectively. Accept if and only if $a \cdot \omega_W = b$.

Figure 4: The game G_{dlS} checks (anti)commutation relations between two collections of observables.

We recall the following result from [**?**] about this game. (The result from [**?**] is more general, and applies to non-uniform measures on Ω. We only need the consequence stated here.) Before stating the result, we introduce the notion of a *qubit test*.

For a projective measurement $P = \{P_a\}_{a \in \mathbb{F}_2^k}$, we use $\widehat{P}(b)$ to denote the observable

$$\widehat{P}(b) \;=\; \sum_a (-1)^{a \cdot b} P_a \;.$$

For binary outcome measurements $P = \{P_0, P_1\}$, we write $\widehat{P}$ to denote $P_0 - P_1$. Recall the notation σ^W for the Pauli observables introduced in Section **??**.

Definition 5.6 (Qubit test)**.** Let $k \in \mathbb{N}$ and $\delta : [0,1] \to \mathbb{R}_+$. A $(k, \delta(\varepsilon))$*-qubit test* is a synchronous game $G = (\mathcal{X}, \mu, \mathcal{A}, D)$ such there are two sets $S_X, S_Z \subseteq \mathbb{F}_2^k$ that each span $\mathbb{F}_2^k$ and an injection $\phi : (\{X\} \times S_X) \cup (\{Z\} \times S_Z) \to \mathcal{X}$ such that $\mathcal{A}(\phi(X,a)) = \mathcal{A}(\phi(Z,b)) = \mathbb{F}_2$ for all $a \in S_X$, $b \in S_Z$ and such that the following holds:

- (Completeness:) There is a synchronous strategy $\mathscr{S} = \{P^x\}_{x \in \mathcal{X}}$ for G on $\mathcal{M} = M_{2^k}(\mathbb{C}) \otimes \mathcal{H}$, for some Hilbert space $\mathcal{H}$, that succeeds with probability 1 in G and is such that $\widehat{P}^{\phi(W,a)} = \sigma^W(a)$ for every $W \in \{X, Z\}$ and $a \in S_W$.
- (Soundness:) Let μ' denote the (renormalized) restriction of (the marginal of) μ to the image of ϕ in $\mathcal{X}$. Any synchronous strategy $\mathscr{S} = \{P^x\}_{x \in \mathcal{X}}$ on $(\mathcal{M}, \tau)$ for G that succeeds with probability $1 - \varepsilon$ for some $\varepsilon \geq 0$ is $(\delta(\varepsilon), \tilde{\mu})$-close to a strategy on some algebra $(M_{2^k}(\mathbb{C}) \otimes \mathcal{N}, \mathrm{tr} \otimes \tau')$ where $(\mathcal{N}, \tau')$ is a tracial sub-algebra of $\mathcal{M}_\infty$ and such that

$$\widehat{P}^{\phi(W,a)} = \sigma^W(a) \otimes I_{\mathcal{N}} \;.$$

The terminology "qubit test" is motivated by the notion of a *qubit* as introduced in [**?**]. Informally, a qubit is a copy of the space $\mathbb{C}^2$ together with a pair of anticommuting observables X, Z acting on it. A qubit test is then a test that forces any successful strategy for the players in it to "contain", as a subset of its measurement operators, a representation of (generators of) the algebra of k qubits.

Theorem 5.7 (Corollary 3.9 in [**?**])**.** *Suppose that $E \in \mathbb{F}_2^{k \times n}$ is such that the rows of E span an $[n,k,d]_2$ linear code. The game G_{dlS} is a $(k, O(\varepsilon))$-qubit test, where the sets S_X and S_Z are as in* (**??**) *and $\phi((X,a)) = (X,a)$ and $\phi((Z,b)) = (Z,b)$ and the $O(\varepsilon)$ hides a (quadratic) dependence on d/n.*

The goal in the remaining sections is to design a qubit test where the size of the question set is $\mathrm{polylog}(k)$, as opposed to $\Omega(k^2)$ here.

5.3 The code game

In this section we associate a game $G_{\mathscr{C},M}$ to any $[n,k,d]_2$ code $\mathscr{C}$ and r-local tester $M = (h, \nu)$ for it. In the game, one player is asked to provide an assignment to all generators in the support of a randomly chosen row of h, such that this assignment satisfies the check enforced by that row. The other player is asked to provide an assignment to a single of these variables, and checked for consistency with the first player. The formal definition follows.

Definition 5.8. Let $\mathscr{C}$ be an $[n,k,d]_q$ linear code and $M(h,\nu)$ an r-local tester for $\mathscr{C}$ such that $h \in \mathbb{F}_2^{m\times n}$. The game $G_{\mathscr{C},M}$ is defined as follows. We set

$$\mathcal{X} = \big\{\{\text{eq}\}\times\{1,\dots,m\} \sqcup \{\text{var}\}\times\{1,\dots,n\}\big\}\ ,$$

and

$$\mu((\text{var},i),(\text{eq},j)) = \mu((\text{eq},j),(\text{var},i)) = \frac{1}{m}\frac{1}{|h_{j\cdot}|}1_{h_{ji}=1}\ ,$$

where $|h_{j\cdot}|$ denotes the Hamming weight of the j-th row of h. For any j, $\mathcal{A}((\text{eq},j)) = \mathbb{F}_2^{|h_{j\cdot}|}$, and for any $i\in\{1,\dots,n\}$, $\mathcal{A}((\text{var},i)) = \mathbb{F}_2$. Finally $D((\text{eq},j),(\text{var},i),a,b) = 1_{h_{j\cdot}a=0}1_{a_i=b}$, where $h_{j\cdot}a$ is naturally computed as the sum, in $\mathbb{F}_2$, of all entries of a.

We show the following. For $\mathscr{C}$ an $[n,k,d]_2$ linear code and $M=(h,\nu)$ an r-local tester for $\mathscr{C}$, recall the presentation (**??**) of $G(h)$. We define a distribution μ on the relations defining that presentation as follows. First, sample a uniformly random $j\in\{1,\dots,m\}$ and uniformly random $i,i'\in\{1,\dots,n\}$, conditioned on $h_{ji}=h_{ji'}=1$ and $i\neq i'$. Then, with probability $1/3$ we return the relation $x_i^2=e$, with probability $1/3$ we return the relation R_i, and with probability $1/3$ we return $R'_{jii'}$.

Proposition 5.9. *Let $\mathscr{C}$ be an $[n,k,d]_2$ linear code, $M=(h,\nu)$ an r-local tester for $\mathscr{C}$, and μ the distribution defined above. Let $\mathscr{S}=\{P^x\}_{x\in\mathcal{X}}$ be a strategy for $G_{\mathscr{C},M}$ on $(\mathcal{N},\tau^{\mathcal{N}})$ such that $\omega^*(G_{\mathscr{C},M};\mathscr{S})\geq 1-\varepsilon$. Then $\phi: x_i\mapsto \widehat{P}^{(\text{var},i)}$, for $i\in\{1,\dots,n\}$, is an $(O(r\varepsilon),\mu)$-homomorphism of the presentation* (**??**) *of $G(h)$.*

Proof. Let $\mathscr{S}$ be a synchronous strategy for $G_{\mathscr{C},M}$ in $(\mathcal{M},\tau)$ that succeeds with probability at least $1-\varepsilon$. For $i\in\{1,\dots,n\}$ let

$$\phi(x_i) \;=\; \widehat{P}^{(\text{var},i)} \;=\; P_0^{(\text{var},i)} - P_1^{(\text{var},i)}\ .$$

Then by definition $\phi(x_i)^2 = I$ for all i. Recalling (**??**), it remains to verify the relations R_j and $R'_{jii'}$, for $1\leq i<i'\leq n$ and $1\leq j\leq m$. To show these relations, first express the assumption that $\mathscr{S}$ succeeds in $G_{\mathscr{C},M}$ as

$$1-\varepsilon \leq \mathop{\mathbb{E}}_{j\in\{1,\dots,m\}}\ \mathop{\mathbb{E}}_{i:h_{ji}=1}\ \sum_{a:h_{j\cdot}a=0}\tau\big(P_a^{(\text{eq},j)}P_{a_i}^{(\text{var},i)}\big) \tag{36}$$

For an equation j and variables i,i', let

$$R_{ii'}^{(\text{eq},j)} = \sum_a (-1)^{a_i+a_{i'}}P_a^{(\text{eq},j)} \qquad\text{and}\qquad R_i^{(\text{eq},j)} = \sum_a (-1)^{a_i}P_a^{(\text{eq},j)}\ ,$$

so that, using that $\{P_a^{(\text{eq},j)}\}$ is a projective measurement,

$$R_{ii'}^{(\text{eq},j)} \;=\; R_i^{(\text{eq},j)}R_{i'}^{(\text{eq},j)} \;=\; R_{i'}^{(\text{eq},j)}R_i^{(\text{eq},j)}\ .$$

Using this we compute

$$\begin{aligned}
\mathop{\mathbb{E}}_{(j,i,i')\sim\mu} \|R'_{jii'} - I\|_\tau^2 &= \mathop{\mathbb{E}}_{(j,i,i')} \big\| \widehat{P}^{(\mathrm{var},i)} \widehat{P}^{(\mathrm{var},i')} - \widehat{P}^{(\mathrm{var},i')} \widehat{P}^{(\mathrm{var},i)} \big\|_\tau^2 \\
&= \mathop{\mathbb{E}}_{(j,i,i')} \big\| \big(\widehat{P}^{(\mathrm{var},i)} - R_i^{(\mathrm{eq},j)}\big) \widehat{P}^{(\mathrm{var},i')} + R_i^{(\mathrm{eq},j)} \big(\widehat{P}^{(\mathrm{var},i')} - R_{i'}^{(\mathrm{eq},j)}\big) \\
&\qquad\qquad - R_{i'}^{(\mathrm{eq},j)} \big(\widehat{P}^{(\mathrm{var},i)} - R_i^{(\mathrm{eq},j)}\big) - \big(\widehat{P}^{(\mathrm{var},i')} - R_{i'}^{(\mathrm{eq},j)}\big) \widehat{P}^{(\mathrm{var},i)} \big\|_\tau^2 \\
&\leq 2\Big(\mathop{\mathbb{E}}_{(j,i,i')} \big\| \widehat{P}^{(\mathrm{var},i)} - R_i^{(\mathrm{eq},j)} \big\|_\tau^2 + \mathop{\mathbb{E}}_{(j,i,i')} \big\| \widehat{P}^{(\mathrm{var},i')} - R_{i'}^{(\mathrm{eq},j)} \big\|_\tau^2 \Big) \\
&= 2\Big(4 - 2\Big(\mathop{\mathbb{E}}_{(j,i,i')} \tau\big(\widehat{P}^{(\mathrm{var},i)} R_i^{(\mathrm{eq},j)}\big) + \tau\big(\widehat{P}^{(\mathrm{var},i')} R_{i'}^{(\mathrm{eq},j)}\big) \Big) \Big) \\
&= 2\Big(8 - 8 \mathop{\mathbb{E}}_{j\in\{1,\ldots,m\}} \mathop{\mathbb{E}}_{i:h_{ji}=1} \sum_a \tau\big(\widehat{P}_{a_i}^{(\mathrm{var},i)} P_a^{(\mathrm{eq},j)}\big) \Big) \,.
\end{aligned}$$

Here we abused notation slightly and denoted $\mathbb{E}_{(j,i,i')}$ the expectation for a relation $R'_{jii'}$ sampled according to μ, conditioned on such a relation being sampled. The third line is the Cauchy-Schwarz inequality and uses that $\widehat{P}^{(\mathrm{var},i)}$ and $R_i^{(\mathrm{eq},j)}$ are observables, hence square to identity; the fourth line expands the norms and also uses this fact; and the fifth line uses that for observables $A = A0 - A_1$ and $B = B_0 - B_1$, $AB = I - 2(A_0B_0 + A_1B_1)$. The last line also uses that the marginal of $(j,i,i') \sim \mu$ on (j,i) or (j,i') are identical and match the distribution indicated in the last line. Using (**??**) we deduce that

$$\mathop{\mathbb{E}}_{(j,i,i')\sim\mu} \|R'_{jii'} - I\|_\tau^2 \leq 16\varepsilon \,.$$

Now we consider the relations R_j. For $j \in \{1,\ldots,m\}$ we denote $i_1,\ldots,i_r$ the indices such that $h_{ji} = 1$ (assume for simplicity of notation that there are exactly r). Then

$$\begin{aligned}
\mathop{\mathbb{E}}_j \|R_j - I\|_\tau^2 &= \mathop{\mathbb{E}}_j \big\| \widehat{P}^{(\mathrm{var},i_1)} \cdots \widehat{P}^{(\mathrm{var},i_r)} - I \big\|_\tau^2 \\
&\leq (r+1)\Big(\mathop{\mathbb{E}}_j \big\| R_{i_1}^{(\mathrm{eq},j)} \cdots R_{i_r}^{(\mathrm{eq},j)} - I \big\|_\tau^2 + \sum_{t=1}^r \big\| R_{i_t}^{(\mathrm{eq},j)} - \widehat{P}^{(\mathrm{var},i_t)} \big\|_\tau^2 \Big) \\
&\leq (r+1)\big(2 - 2 \mathop{\mathbb{E}}_j \tau\big(R_{i_1}^{(\mathrm{eq},j)} \cdots R_{i_r}^{(\mathrm{eq},j)}\big)\big) + O(r\varepsilon) \\
&\leq O(r\varepsilon) \,,
\end{aligned}$$

where the second line follows from the triangle inequality and using a telescoping sum, the third line uses the definition of R and (**??**), and the last line again uses the definition of R and (**??**). This concludes the proof. □

5.4 Braiding the code test

Let $\mathscr{C}$ be an $[n,k,d]_2$ linear code and $M = (h,\nu)$ an r-local tester for $\mathscr{C}$. We let $E_{\mathscr{C}} \in \mathbb{F}_2^{k\times n}$ be a generating matrix for $\mathscr{C}$, i.e. $E_{\mathscr{C}}$ is such that its rows are linearly independent and span the codespace. The braiding test constructed from $\mathscr{C}$ and M is a synchronous game described in Figure **??**. The test combines two independent copies of the code game from Section **??** with appropriate commutation and anti-commutation sub-tests. The braiding test is designed to force any successful strategy in it to be close, in some sense, to a representation of the Pauli group generated by observables $\sigma^X(a)$ and $\sigma^Z(b)$, $a,b \in \mathbb{F}_2^k$ (recall the notation from (**??**)). This is shown in the following theorem.

Let $M = (h, \nu)$ be an r-local tester for the $[n, k, d]_2$ code $\mathscr{C}$. Execute either of the following tests with probability $1/3$ each.

1. **Code test:** Sample $W \in \{X, Z\}$ uniformly at random. Execute the code game $G_{\mathscr{C},M}$ with both players, prepending the symbol W to all questions (which now take the form (W, var, i) or (W, eq, j)).
2. **(Anti-)commutation test:** Sample $(i_X, i_Z) \in \{1, \ldots, n\}^2$ uniformly at random. Let $\omega = (E_{\mathscr{C}} e_{i_X}, E_{\mathscr{C}} e_{i_Z})$ and $\gamma = (E_{\mathscr{C}} e_{i_X}) \cdot (E_{\mathscr{C}} e_{i_Z}) \in \mathbb{F}_2$.
   (a) If $\gamma = 0$ then sample a pair of questions $(x_c, y_c) \sim \mu_{\mathrm{com}}$ as in the commutation game. Send (x_c, ω) to A and (y_c, ω) to Bob. Accept if and only if their answers are accepted in the commutation game.
   (b) If $\gamma \neq 0$ then do the same but for the anti-commutation game.
3. **Consistency test:** Sample $(i_X, i_Z) \in \{1, \ldots, n\}^2$ and $W \in \{X, Z\}$ uniformly at random. Let $\omega = (E_{\mathscr{C}} e_{i_X}, E_{\mathscr{C}} e_{i_Z})$ and $\gamma = (E_{\mathscr{C}} e_{i_X}) \cdot (E_{\mathscr{C}} e_{i_Z}) \in \mathbb{F}_2$. Send (W, var, i_W) to A and $(x_{W,\gamma}, \omega)$ to B, where $x_{W,\gamma}$ is a question from the (anti-)commutation game. Receive $a \in \mathbb{F}_2$ and $b \in \mathbb{F}_2$ respectively. Accept if and only if $a = b$.

Figure 5: The braiding test over $\mathscr{C}$ verifies that the players respond consistently with a uniformly random codeword from $\mathscr{C}$. $E_{\mathscr{C}} \in \mathbb{F}_2^{k \times n}$ is a generating matrix for $\mathscr{C}$, and for $i \in \{1, \ldots, n\}$ we let e_i be the i-th canonical basis vector of $\mathbb{F}_2^n$.

Theorem 5.10. *Let $\mathcal{C}$ be a class of tracial von Neumann algebras. Let $\mathscr{C}$ be an $[n,k,d]_2$ linear code and $M=(h,\nu)$ an r-local tester for $\mathscr{C}$. Suppose that the presentation $G(h)$ in (**??**) is such that $G(h)=\mathbb{Z}_2^k$ and furthermore this presentation is $(\delta,\nu_R,\nu_S,\mathcal{C})$-stable.*[15] *Then the braiding test over $\mathscr{C}$ is a (k,δ')-qubit test with sets $S_X=S_Z=\{E_{\mathscr{C}}e_i : i\in\{1,\dots,n\}\}\subseteq\mathbb{F}_2^k$, map $\phi(W,E_{\mathscr{C}}e_i)=(W,\mathrm{eq},i)$ and error function $\delta'=O(\delta^{1/2}(6\varepsilon))$.*[16]

We will make use of the following simple fact.

Lemma 5.11 (Data-processing)**.** *Let $\{P_a\}$ and $\{Q_a\}$ be two POVMs on $(\mathcal{M},\tau)$ with the same outcome set $\mathcal{A}$. Then for any function $f:\mathcal{A}\to\mathcal{B}$ for some finite set $\mathcal{B}$,*

$$\sum_{b\in\mathcal{B}}\Big\|\sum_{a\in f^{-1}(b)}(P_a-Q_a)\Big\|_2^2\leq\sum_{a\in\mathcal{A}}\|P_a-Q_a\|_2^2\,. \tag{37}$$

Proof. This follows by expanding the left-hand side and using $\tau(P_aQ_{a'})\geq 0$ for all $a\neq a'$. □

Proof of Theorem **??**. Completeness: We first verify completeness. For $W\in\{X,Z\}$, $i\in\{1,\dots,n\}$ and $b\in\mathbb{F}_2$ let $P_b^{(W,\mathrm{var},i)}=\frac{1}{2}(I+(-1)^b\sigma^W(E_{\mathscr{C}}e_i))$, and for $j\in\{1,\dots,m\}$, with m the number of rows of h, and $a\in\mathbb{F}_2^r$ let $P_a^{(W,\mathrm{eq},j)}=\prod_{i:h_{ji}=1}P_{a_i}^{(W,\mathrm{var},i)}$. Writing (f_0,f_1) for the canonical basis of $\mathbb{C}^2$, $P_a^{(W,\mathrm{eq},j)}$ is the projection on the span of all $\otimes_{i=1}^k f_{u_i}$ where $u=(u_1,\dots,u_k)$ is such that $(u^TE_{\mathscr{C}})_{|S_j}=a$, where S_j denotes the support of $h_{j\cdot}$. For $(i_X,i_Z)\in\{1,\dots,n\}^2$ let $\omega=(E_{\mathscr{C}}e_{i_X},E_{\mathscr{C}}e_{i_Z})$ and $\gamma=(E_{\mathscr{C}}e_{i_X})\cdot(E_{\mathscr{C}}e_{i_Z})$. We let $P^{x_{W,\gamma},\omega}=P^{(W,\mathrm{var},i_W)}$.

These choices already ensure that the strategy succeeds with probability 1 in the consistency test. We verify that it succeeds in the code test. Let $j\in\{1,\dots,m\}$. As observed above, for any $a\in\mathbb{F}_2^r$ such that $P_a^{W,\mathrm{eq},j}\neq 0$ there is an $u\in\mathbb{F}_2^k$ such that $(u^TE_{\mathscr{C}})_{|S_j}=a$, which means that a is the restriction of a valid element of $\mathscr{C}$. Using the completeness property of M it follows that M must accept any a such that $P_a^{W,\mathrm{eq},j}\neq 0$, which shows that the strategy succeeds in the code test with probability 1.

It remains to verify that the anti-commutation test is passed with probability 1. For this we observe that the binary observables

$$U=\widehat{P}^{x_{X,\gamma},\omega}\quad\text{and}\quad V=\widehat{P}^{x_{Z,\gamma},\omega}$$

commute in case $\gamma=0$ and anti-commute in case $\gamma=1$. This is because by construction $U=\sigma^X(E_{\mathscr{C}}i_X)$ and $V=\sigma^W(E_{\mathscr{C}}i_W)$, and because of the definition of γ. Hence the pair (U,V) can be completed to a perfect strategy for the commutation game (if $\gamma=0$) or anti-commutation game (if $\gamma=1$). This defines the measurements $P^{(x,\omega)}$ for $x\notin\{x_{W,\gamma}:W\in\{X,Z\},\gamma\in\{0,1\}\}$.

Soundness: Next we show soundness. Let $\mathscr{S}$ be a synchronous strategy for the braiding test in $(\mathcal{M},\tau)$ that succeeds with probability at least $1-\varepsilon$. For $W\in\{X,Z\}$ let $\mathscr{S}^W$ be the strategy in $G_{\mathscr{C},M}$ that is obtained by restricting $\mathscr{S}$ to the relevant measurements corresponding to the "Code test" part of the braiding test, i.e. the $P^{W,\mathrm{eq},j}$ and $P^{W,\mathrm{var},i}$. Then $\mathscr{S}^W$ succeeds with probability at least $1-6\varepsilon$ in $G_{\mathscr{C},M}$. Using Proposition **??** it follows that $\{\widehat{P}^{W,\mathrm{var},i}\}$ is an $O(\varepsilon)$-homomorphism of $G(h)$. Under the assumption that $G(h)=\mathbb{Z}_2^k$ is $(\delta,\nu_R,\nu_S,\mathcal{C})$-stable, we deduce that there is a $\delta_1=O(\delta(6\varepsilon))$ such that for each $W\in\{X,Z\}$,

[15] The distributions ν_R and ν_S are defined from ν as described right after Definition **??**.

[16] Assume $E_{\mathscr{C}}$ has no repeated columns.

$\mathscr{S}^W$ is (δ_1,ν)-close to a perfect strategy $\tilde{\mathscr{S}}^W$ on $(\mathcal{N}^W,\tau^W)$ for $G_{\mathscr{C},M}$, where ν is uniform on $\{1,\dots,n\}$. This strategy has measurement operators $\{\tilde{P}_a^{W,\mathrm{eq},j}\}$ and $\{\tilde{P}_b^{W,\mathrm{var},i}\}$ which satisfy

$$\mathop{\mathbb{E}}_{i\in\{1,\dots,n\}} \sum_{b\in\mathbb{F}_2} \left\| P_b^{W,\mathrm{var},i} - (w^W)^* \tilde{P}_b^{W,\mathrm{var},i} w^W \right\|_2^2 = O(\delta_1) \,. \tag{38}$$

Furthermore, using that $G(h)=\mathbb{Z}_2^k$ is Abelian, there is a POVM $\{\tilde{P}_u^W\}_{u\in\mathbb{F}_2^n}$ such that $\sum_{u\in\mathscr{C}}\tilde{P}_u^W = I$ and for each $i\in\{1,\dots,n\}$, $\tilde{P}_b^{W,i}=\sum_{u:u_i=b}\tilde{P}_u^W$.

Applying Lemma **??**, we obtain projective measurements $\{Q_u^W\}$ on $\mathcal{M}$ such that

$$\sum_u \left\| Q_u^W - (w^W)^* \tilde{P}_u^W w^W \right\|_2^2 = O(\delta_1) \,. \tag{39}$$

For any $i\in\{1,\dots,n\}$ let $Q_b^{W,i}=\sum_{u:\,u_i=b} Q_u^W$. Then by Lemma **??**,

$$\begin{aligned} \mathop{\mathbb{E}}_i \sum_b \tau\big(Q_b^{W,i}(w^W)^*\tilde{P}_b^{W,i}(w^W)\big) &\geq \mathop{\mathbb{E}}_i \sum_u \tau\big(Q_u^W(w^W)^*\tilde{P}_u^W(w^W)\big) \\ &\geq 1-O(\delta_1) \,. \end{aligned}$$

For $W\in\{X,Z\}$ and $b\in\mathbb{F}_2^k$ let $R_b^W = Q^W_{G_{\mathcal{C}}^T b}$, where by definition $G_{\mathcal{C}}^T b\in\mathcal{C}$.

We now define a strategy $\mathscr{S}'$ for the game G_{dlS}. On question $W\in\{X,Z\}$ the projective measurement is $\{R_b^W\}$. On question of the form (x_c,ω) for x_c a question in the commutation game, the projective measurement is $\{P^{x_c,\omega}\}$, i.e. the same projective measurement as used in $\mathscr{S}$. Similarly, on a question of the form (x_{ac},ω) for x_{ac} a question in the anti-commutation game, the projective measurement is $\{P^{x_{ac},\omega}\}$.

To conclude we show that this strategy succeeds in the game G_{dlS} with probability $1-O(\sqrt{\delta_1})$. Assuming that this has been shown, by Theorem **??** the strategy $\mathscr{S}'$ is $O(\sqrt{\delta_1})$-close to a strategy $\mathscr{S}''$ on an algebra of the form $(M_{2^k}(\mathbb{C})\otimes\mathcal{N},\mathrm{tr}\otimes\tau')$ such that $P_b^W=\sigma_b^W\otimes I_{\mathcal{N}}$. By definition of R,

$$Q_b^{W,i} = \sum_{a\in\mathbb{F}_2^n:\,a_i=b} Q_a^W = \sum_{c\in\mathbb{F}_2^k:(G_{\mathcal{C}}^T c)_i=b} R_c^W \,, \tag{40}$$

hence

$$\widehat{Q}^{W,i} = \sum_c (-1)^{c\cdot(G_{\mathcal{C}}e_i)} R_c^W = \widehat{R}^W(G_{\mathcal{C}}e_i) \,.$$

Using the definition of the game distribution, closeness of $\mathscr{S}'$ and $\mathscr{S}''$ thus implies that

$$\mathop{\mathbb{E}}_{i\in\{1,\dots,n\}} \left\| \widehat{Q^{W,i}} - (w'')^*\sigma^W(G_{\mathcal{C}}e_i)(w'') \right\|_2^2 = O(\sqrt{\delta_1}) \,.$$

Combining with (**??**) and (**??**), this shows the theorem.

It remains to verify that $\mathscr{S}'$ succeeds in the game G_{dlS} with probability $1-O(\sqrt{\delta_1})$. By definition $\mathscr{S}'$ succeeds in the (anti)-commutation test with probability $1-O(\varepsilon)$. It remains to check the W-consistency test, for $W\in\{X,Z\}$. Because $\mathscr{S}$ succeeds with probability $1-O(\varepsilon)$ in the consistency test,

$$\mathop{\mathbb{E}}_{i_X,i_Z\in\{1,\dots,n\}} \sum_b \tau\big(P_b^{W,\mathrm{var},i} P_b^{x_{W,\gamma},\omega}\big) \geq 1-O(\varepsilon) \,,$$

where ω and γ are defined as in Figure **??**. Using (**??**), (**??**) and Lemma **??** it follows that

$$\mathop{\mathbb{E}}_{i_X,i_Z\in\{1,\dots,n\}} \sum_b \tau\big(Q_b^{W,i} P_b^{x_{W,\gamma},\omega}\big) \geq 1-O(\sqrt{\delta_1}) \,,$$

Using (**??**), this can be rewritten as

$$\mathop{\mathbb{E}}_{i_X,i_Z\in\{1,\ldots,n\}} \sum_{b,c:(G_C^T c)_i=b} \tau\big(R_c^W P_b^{x_{W,\gamma},\omega}\big) \;\geq\; 1 - O(\sqrt{\delta_1})\;. \tag{41}$$

Since $\omega_W = G_C e_i$, $(G_C^T c)_i = c \cdot \omega_W$. Thus (**??**) shows that $\mathscr{S}'$ succeeds with probability $1 - O(\sqrt{\delta_1})$ in the W-consistency test, as desired. □

5.5 The quantum low-degree test

By instantiating $\mathscr{C}$ using the Reed-Muller code from Section **??** we obtain an efficient qubit test. Because we do not know if $G(h_{\mathrm{RM2}})$ is Abelian, we need to introduce an additional test for the relations $\{R_k^{\mathrm{COM}}\}$ in (**??**). The resulting test is described in Figure **??**. It is a variant of a test first introduced in [**?**] (with a flawed analysis). This paper (together with the work on which our analysis relies) corrects this. In the next section we detail an important application of this test in the work **??**.

Let $h_{\mathrm{RM2}} \in \mathbb{F}_2^{M\times N}$ be the parity check matrix for $\mathscr{C}_{\mathrm{RM2}}$ considered in Section **??**. Here, $N = q^{m+1}$ and $M = q^{m+2}(1+m)$. Let μ_R be the distribution on relations (**??**) described in Section **??**. Execute either of the following tests with probability $1/4$ each.

1. **Code test:** Identical to that in Figure **??**.
2. **(Anti-)commutation test:** Identical to that in Figure **??**.
3. **Consistency test:** Identical to that in Figure **??**.
4. **Pairwise commutation test:** Sample $(i,i') \in \{1,\ldots,N\}^2$ according to the distribution μ_R, conditioned on choosing a relation from R_k^{COM}, and $W \in \{X,Z\}$ uniformly at random. Let $\omega = (E_{\mathscr{C}} e_i, E_{\mathscr{C}} e_{i'})$. Sample a pair of questions (x_c, y_c) as in the commutation game. If either question is $x_{X,0}$ or $x_{Z,0}$, replace it with (W,var,i) or (W,var,i') respectively. Otherwise, send the original question together with ω, i.e. (x_c,ω) or (y_c,ω) respectively. Accept if and only if the players' answers are accepted in the commutation game.

Figure 6: The quantum low-degree test, obtained by adapting the braiding test from Figure **??** to the $[q^{m+1}, t(d+1)^m, D']_2$ code $\mathscr{C}_{\mathrm{RM2}}$.

Corollary 5.12 (Quantum low-degree test)**.** *Let $\mathscr{C}_{\mathrm{RM2}}$ be the $[q^{m+1}, t(d+1)^m, D']_2$ Reed-Muller from Section **??**, and $M = (h_{\mathrm{RM2}}, \nu_{\mathrm{RM2}})$ the $(d+2)$-local tester for $\mathscr{C}_{\mathrm{RM2}}$ described in Figure **??**. Then the associated braiding test (Figure **??**) is a (k,δ')-qubit test with error function $\delta' = \mathrm{poly}(m,d,t)\cdot\mathrm{poly}(\varepsilon, q^{-1})$.*

Proof. Let $\mathscr{S}$ be a synchronous strategy that succeeds in the braiding test with probability at least $1-\varepsilon$. Then in particular the strategy succeeds with probability at least $1 - 4\varepsilon/3$ in the braiding test over $\mathscr{C}_{\mathrm{RM2}}$. Since $(\mathscr{C}_{\mathrm{RM2}}, M)$ is not known to be abelian, we cannot apply Theorem **??** directly. However, we can follow its proof.

The completeness part of the proof follows in a straightforward manner, since the measurement operators $P^{(W,\mathrm{var},i)}$ and $P^{(W,\mathrm{var},i')}$ defined in the proof pairwise commute, for any pair $(i,i') \in \{1,\ldots,N\}^2$.

For the soundness part, we first define the same pair of strategies $\mathscr{S}^X$ and $\mathscr{S}^Z$ for $G_{\mathscr{C}_{\mathrm{RM2}},M}$. Applying Proposition **??**, we deduce an $O(d\varepsilon)$-homomorphism of the presentation (**??**) of $G(h_{\mathrm{RM2}})$. However, we are interested in constructing an approximate homomorphism of the presentation (**??**), which in addition contains the relations R_k^{COM}. The fact that $x_i \mapsto \widehat{P}^{(W,\mathrm{var},i)}$ satisfies these relations, on average and according to the distribution μ_R, follows from success in the pairwise commutation test executed as part of the Pauli braiding test (Figure **??**). Thus we obtain that $x_i \mapsto \widehat{P}^{(W,\mathrm{var},i)}$ is an $(O(d\varepsilon),\mu_R)$-homomorphism of the presentation (**??**). Applying Theorem **??**, and similarly to the proof of Theorem **??** we obtain a pair of POVMs $\{\tilde{P}_u^W\}_{u\in\mathbb{F}_2^N}$, for $W\in\{X,Z\}$, such that defining $\tilde{P}_b^{W,i}=\sum_{u:u_i=b}\tilde{P}_u^W$ these operators satisfy (**??**), with right-hand side $\delta_2=\delta(O(d\varepsilon))$, with δ the soundness function from Theorem **??**. From here on the proof proceeds exactly as the proof of Theorem **??**. □

5.6 Dimension bounds

The next proposition states a simple consequence of a qubit test, which is that strategies with a high enough success probability must have a large dimension. This consequence is used in [**?**].

Proposition 5.13. *Let $G=(\mathcal{X},\mu,\mathcal{A},D)$ denote a $(k,\delta(\varepsilon))$-qubit test. Then all synchronous strategies $\mathscr{S}$ in $(\mathcal{M},\tau)$ for G that succeed with probability $1-\varepsilon$ must satisfy*

$$\dim(\mathcal{M})\geq\Big(1+O(\sqrt{\delta(\varepsilon)})+\frac{\delta(\varepsilon)}{1-\delta(\varepsilon)}\Big)^{-1}2^k\,.$$

Proof. If $\mathcal{M}$ is infinite-dimensional, then we are done. Suppose instead it were finite-dimensional. Then $\mathcal{M}$ must be (isomorphic to) a direct sum of finite-dimensional matrix algebras. Without loss of generality we assume that $\mathcal{M}=M_d(\mathbb{C})$ with the dimension-normalized trace $\tau=\frac{1}{d}\mathrm{Tr}$.

By the soundness property of qubit tests, the strategy $\mathscr{S}$ is $(\delta(\varepsilon),\tilde{\mu})$-close to a strategy $\mathscr{S}'$ on an algebra $(M_{2^k}(\mathbb{C})\otimes\mathcal{N},\mathrm{tr}_{2^k}\otimes\tau')$ for some tracial algebra $(\mathcal{N},\tau')$ where $\mathrm{tr}_{2^k}=2^{-k}\mathrm{Tr}$. For notational brevity we write $\mathcal{R}=M_{2^k}(\mathbb{C})\otimes\mathcal{N}$ and $\tau^{\mathcal{R}}=\mathrm{tr}_{2^k}\otimes\tau'$. By definition there exists a projection $P\in\mathcal{M}_\infty$ of finite trace and a partial isometry $w\in P\mathcal{M}_\infty 1_{\mathcal{M}}$ satisfying

1. $\mathcal{R}=P\mathcal{M}_\infty P$.
2. $\max\{\tau(1_{\mathcal{M}}-w^*w),\tau^{\mathcal{R}}(P-ww^*)\}\leq\delta(\varepsilon)$.
3. $\tau^{\mathcal{R}}=\tau_\infty/\tau_\infty(P)$.

For $u\in\mathbb{F}_2^k$ let σ_u^Z denote the projection

$$\sigma_u^Z=2^{-k}\sum_{a\in\mathbb{F}_2^k}(-1)^{a\cdot u}\sigma^Z(a)\,.$$

It is easy to verify that $\{\sigma_u^Z\otimes I_{\mathcal{N}}\}_{u\in\mathbb{F}_2^k}$ is a projective measurement in $\mathcal{R}$ and furthermore $\tau^{\mathcal{R}}(\sigma_u^Z\otimes I_{\mathcal{N}})=2^{-k}$. Applying **??** we get that there exists a projective measurement $\{Q_u\}_{u\in\mathbb{F}_2^k}$ on $\mathcal{M}$ such that

$$\sum_u\|Q_u-w^*(\sigma_u^Z\otimes I_{\mathcal{N}})w\|_2^2\leq 56\,\delta(\varepsilon)\,.$$

Applying **??** we get

$$\sum_u\tau\Big(\Big|Q_u-w^*(\sigma_u^Z\otimes I_{\mathcal{N}})w\Big|\Big)\leq O(\sqrt{\delta(\varepsilon)})\,.$$

Then we have

$$\begin{aligned}\sum_u \left|\tau(Q_u) - 2^{-k}\right| &\leq \sum_u \left|\tau(w^*(\sigma_u^Z \otimes I_{\mathcal{N}})w) - 2^{-k}\right| + \tau\Big(\left|Q_u - w^*(\sigma_u^Z \otimes I_{\mathcal{N}})w\right|\Big) \\ &= O(\sqrt{\delta(\varepsilon)}) + \sum_u \left|\tau_\infty(ww^*(\sigma_u^Z \otimes I_{\mathcal{N}})) - 2^{-k}\right| \\ &= O(\sqrt{\delta(\varepsilon)}) + \sum_u \left|\tau_\infty(P(\sigma_u^Z \otimes I_{\mathcal{N}})) - 2^{-k}\right| + \left|\tau_\infty((P - ww^*)(\sigma_u^Z \otimes I_{\mathcal{N}}))\right|\end{aligned}$$

Notice that $\tau_\infty(P(\sigma_u^Z \otimes I_{\mathcal{N}})) = \tau_\infty(\sigma_u^Z \otimes I_{\mathcal{N}}) = 2^{-k}$, and that $ww^* \leq P$ and thus $\tau_\infty((P - ww^*)(\sigma_u^Z \otimes I_{\mathcal{N}}))$ is a nonnegative real number. Therefore the sum in the last line simplifies to

$$\sum_u \tau_\infty((P - ww^*)(\sigma_u^Z \otimes I_{\mathcal{N}})) = \tau_\infty((P - ww^*)P) = \tau_\infty(P - ww^*) \leq \tau_\infty(P) \cdot \delta(\varepsilon)\,.$$

On the other hand the proof of **??** shows that $\tau_\infty(P) \leq \frac{1}{1-\delta(\varepsilon)}$, and thus

$$\sum_u \left|\tau(Q_u) - 2^{-k}\right| \leq O(\sqrt{\delta(\varepsilon)}) + \frac{\delta(\varepsilon)}{1 - \delta(\varepsilon)}\,.$$

By averaging, there exists a $u \in \mathbb{F}_2^k$ such that

$$\tau(Q_u) = \frac{1}{d}\mathrm{Tr}(Q_u) \leq \Big(1 + O(\sqrt{\delta(\varepsilon)}) + \frac{\delta(\varepsilon)}{1 - \delta(\varepsilon)}\Big)2^{-k}\,.$$

Rearranging, this implies that d, the dimension of $\mathcal{M}$, satisfies

$$d \geq \Big(1 + O(\sqrt{\delta(\varepsilon)}) + \frac{\delta(\varepsilon)}{1 - \delta(\varepsilon)}\Big)^{-1} 2^k$$

as desired. □

A Lower bounds on the modulus of stability of (??)

In this Appendix we prove Lemma **??**, and deduce some corollaries from it. Before diving into the proof, we first sketch the idea behind it. We provide a method for translating every graph $([k], E \subseteq \binom{[k]}{2})$ into a collection of order 2 unitaries $A_1, ..., A_k$ (actually, permutations) that satisfy that

$$\forall ij \in \binom{[k]}{2} : \quad \|A_iA_j - A_jA_i\|_\tau^2 = \begin{cases} 2 & ij \in E, \\ 0 & ij \notin E. \end{cases} \tag{42}$$

Therefore, these unitaries correspond to an $(|E|/\binom{k}{2}, \mu_R)$-approximate representation of (**??**). For any pair $ij \in E$ we have $\|A_iA_j - A_jA_i\|_\tau^2 = 2$, and thus the unitaries A_i and A_j would need to be changed by some constant amount to make them commute. Thus, to fix the A_i's into a genuine representation of $\mathbb{Z}_2^k$, we would need to move them by at least a constant times the proportion of the largest matching which embeds in E. Hence, by choosing a graph which is already a matching with c edges, we deduce the Lemma.

We begin by describing our construction. Given a graph $([k], E)$, the permutations $\{A_i\}_{i=1}^k$ act on the set $\mathbb{F}_2^{[k]\cup E}$, namely bit strings indexed by the vertices and edges of the graph. Now, each A_i flips the i^{th} bit.

Furthermore, it conditionally flips the ij^{th} bit for $ij \in E$ where $i < j$, if the j^{th} bit is 1. Namely, the A_i's act as a NOT gate on the i^{th} bit composed with CNOT gates on all the ij^{th} bits, where $i < j$ and $ij \in E$. Now, all of these permutations are of order 2. Equation (**??**) is deduced by noting that $A_iA_j = A_jA_i$ when ij is not an edge, and that $A_iA_jA_iA_j$ has no fixed points when ij is an edge — this is because $A_iA_jA_iA_j$ must flip the ij^{th} bit. The von Neumann algebra $\mathcal{M}$ containing the A_i's is the one acting on $\mathbb{C}^{\mathbb{F}_2^{[k]\cup E}}$ with its standard normalized trace $\tau^{\mathcal{M}}(X) = \frac{1}{2^{k+|E|}}\mathrm{Tr}(X)$.

Remark A.1. *A version of this construction (viewed in a different way) was used by Kozlov–Meshulam [**?**] to upper bound the Cheeger constant of the k-dimensional hypercube — see Section 4.1 therein.*

Claim A.2. *The map $x_i \mapsto A_i \in \mathcal{M}$ is a $(|E|/\binom{k}{2}, \mu_R)$-approximate representation of $\mathbb{Z}_2^k$ with respect to the presentation* (**??**).

Proof. With probability $\frac{1}{2}$, μ_R samples an involution relation, which is always satisfied by the A_i's. Furthermore,

$$\mathop{\mathbb{E}}_{i\neq j\in[k]}\left[\|A_iA_jA_iA_j - I_{\mathcal{M}}\|_\tau^2\right] = \frac{1}{\binom{k}{2}}\sum_{ij\in E}\|A_iA_jA_iA_j - I_{\mathcal{M}}\|_\tau^2 = 2|E|/\binom{k}{2}.$$

By combining these two observations, we deduce the claim. □

Claim A.3. *Assume the largest matching in E contains c edges. Then, for every collection $\{B_i\}_{i=1}^k$ of order 2 unitaries which pairwise commute in $\mathcal{N} = P\mathcal{M}_\infty P$, and every partial isometry $w = PUI_{\mathcal{M}} \in PU(\mathcal{M}_\infty)I_{\mathcal{M}}$, we have*

$$\mathop{\mathbb{E}}_{i\in[k]}\|A_i - w^*B_iw\|_\tau^2 \geq \frac{c}{16k}$$

or

$$\tau^{\mathcal{M}}(I_{\mathcal{M}} - w^*w) \geq \frac{c}{16k}.$$

In particular, there is no genuine representation of $\mathbb{Z}_2^k$ that is $(\frac{c}{16k}, \mu_S)$-close to the A_i's.

Proof. Recall that given our von Neumann algebra $\mathcal{M}$, the algebra $\mathcal{M}_\infty$ acts on the Hilbert space $\mathbb{C}^{\mathbb{F}_2^{[k]\cup E}} \otimes \mathbb{C}^{\mathbb{Z}}$. Let $\{e_v \otimes e_t \mid v \in \mathbb{F}_2^{[k]\cup E}, t \in \mathbb{Z}\}$ be the standard basis of this Hilbert space. Let $\{B_i\}_{i=1}^k$ be order 2 unitaries which pairwise commute in $\mathcal{N} = P\mathcal{M}_\infty P$, and assume there is a partial isometry $w = PUI_{\mathcal{M}} \in PU(\mathcal{M}_\infty)I_{\mathcal{M}}$ and $0 < \varepsilon \leq 3 - 2\sqrt{2} \approx 0.17$ such that

$$\mathop{\mathbb{E}}_{i\in[k]}\|A_i - w^*B_iw\|_\tau^2\,, \qquad \tau^{\mathcal{M}}(I_{\mathcal{M}} - w^*w) \leq \varepsilon.$$

This is a slightly weaker condition than for the B_i's to be (ε, μ_S)-close to the A_i's, as in Definition **??**. Furthermore, we can assume without loss of generality that $U = I_\infty$, otherwise we replace $\mathcal{N}$ by $U^*\mathcal{N}U$ and P by U^*PU. Thus, we are given that

$$\mathop{\mathbb{E}}_{i\in[k]}\|A_i - I_{\mathcal{M}}B_iI_{\mathcal{M}}\|_\tau^2\,, \qquad \tau^{\mathcal{M}}(I_{\mathcal{M}} - I_{\mathcal{M}}PI_{\mathcal{M}}) \leq \varepsilon,$$

and

$$\|I_{\mathcal{M}} - I_{\mathcal{M}}PI_{\mathcal{M}}\|_\tau^2 = \tau^{\mathcal{M}}(I_{\mathcal{M}}\underbrace{-2I_{\mathcal{M}}PI_{\mathcal{M}} + I_{\mathcal{M}}P^2I_{\mathcal{M}}}_{=-I_{\mathcal{M}}PI_{\mathcal{M}}}) \leq \varepsilon.$$

Let $ij \in E$. By (**??**), we have

$$\begin{aligned}\sqrt{2} &= \|A_iA_jA_iA_j - I_{\mathcal{M}}\|_\tau \\ &\le \|A_iA_jA_iA_j - I_{\mathcal{M}}PI_{\mathcal{M}}\|_\tau + \|I_{\mathcal{M}}PI_{\mathcal{M}} - I_{\mathcal{M}}\|_\tau \\ &\le \|A_iA_jA_iA_j - I_{\mathcal{M}}B_iB_jB_iB_jI_{\mathcal{M}}\|_\tau + \sqrt{\varepsilon} \\ &\le \|I_{\mathcal{M}}(A_i - B_i)A_jA_iA_j\|_\tau + \|I_{\mathcal{M}}B_i(A_j - B_j)A_iA_j\|_\tau \\ &+ \|I_{\mathcal{M}}B_iB_j(A_i - B_i)A_j\|_\tau + \|I_{\mathcal{M}}B_iB_jB_i(A_j - B_j)I_{\mathcal{M}}\|_\tau + \sqrt{\varepsilon} \\ &= (\heartsuit) + \sqrt{\varepsilon}.\end{aligned}$$

Since the A's are unitaries in $\mathcal{M}$, and by abusing notation and denoting $\|X\|_\tau = \tau_\infty(X^*X)$, we have

$$\begin{aligned}(\heartsuit) &= \|I_{\mathcal{M}}(A_i - B_i)I_{\mathcal{M}}\|_\tau + \|I_{\mathcal{M}}B_i(A_j - B_j)I_{\mathcal{M}}\|_\tau \\ &+ \|I_{\mathcal{M}}B_iB_j(A_i - B_i)I_{\mathcal{M}}\|_\tau + \|I_{\mathcal{M}}B_iB_jB_i(A_j - B_j)I_{\mathcal{M}}\|_\tau \\ &\le \|(A_i - B_i)I_{\mathcal{M}}\|_\tau + \|(A_j - B_j)I_{\mathcal{M}}\|_\tau + \|(A_i - B_i)I_{\mathcal{M}}\|_\tau + \|(A_j - B_j)I_{\mathcal{M}}\|_\tau \\ &= (\spadesuit).\end{aligned}$$

But,

$$\|(A_i - B_i)I_{\mathcal{M}}\|_\tau^2 = \|I_{\mathcal{M}}(A_i - B_i)I_{\mathcal{M}}\|_\tau^2 + \|(I_\infty - I_{\mathcal{M}})(A_i - B_i)I_{\mathcal{M}}\|_\tau^2,$$

and since $(I_\infty - I_{\mathcal{M}})A_i = 0$ and $I_{\mathcal{M}}A_i = A_iI_{\mathcal{M}}$, we have

$$\|(I_\infty - I_{\mathcal{M}})(A_i - B_i)I_{\mathcal{M}}\|_\tau^2 = \|(I_\infty - I_{\mathcal{M}})B_iI_{\mathcal{M}}\|_\tau^2 = \|(I_\infty - I_{\mathcal{M}})B_iA_iI_{\mathcal{M}}\|_\tau^2.$$

Here the second equality is because by definition,

$$\|(I_\infty - I_{\mathcal{M}})B_iI_{\mathcal{M}}\|_\tau^2 = \sum_{\substack{v' \in \mathbb{F}_2^{[k]\cup E} \\ j \neq 1}} \sum_{v \in \mathbb{F}_2^{[k]\cup E}} |(e_{v'} \otimes e_j)^* B_1 e_v \otimes e_1|^2 = (\diamondsuit)\ ,$$

but since A_i permutes $\{e_v \otimes e_1\}_{v \in \mathbb{F}_2^{[k]\cup E}}$, we have

$$(\diamondsuit) = \sum_{\substack{v' \in \mathbb{F}_2^{[k]\cup E} \\ j \neq 1}} \sum_{v \in \mathbb{F}_2^{[k]\cup E}} |(e_{v'} \otimes e_j)^* B_iA_ie_v \otimes e_1|^2 = \|(I_\infty - I_{\mathcal{M}})B_iA_iI_{\mathcal{M}}\|_\tau^2\ .$$

Now, for every $v \in \mathbb{F}_2^{[k]\cup E}$, we have

$$1 - |(e_v \otimes e_1)^* B_iA_ie_v \otimes e_1|^2 \le |1 - (e_v \otimes e_1)^* B_iA_ie_v \otimes e_1|^2 \le \|I_{\mathcal{M}}(I_\infty - B_iA_i)e_v \otimes e_1\|_2^2.$$

On the other hand, since B_iA_i is a contraction,

$$\begin{aligned}1 - |(e_v \otimes e_1)^* B_iA_ie_v \otimes e_1|^2 &\ge \|B_iA_ie_v \otimes e_1\|_2^2 - |(e_v \otimes e_1)^* B_iA_ie_v \otimes e_1|^2 \\ &= \sum_{(v',j') \neq (v,1)} |(e_{v'} \otimes e_j)^* B_iA_ie_v \otimes e_1|^2 \\ &\ge \|(I_\infty - I_{\mathcal{M}})B_iA_ie_v \otimes e_1\|_2^2\ .\end{aligned}$$

Therefore, by averaging the combined inequalities over $v \in \mathbb{F}_2^{[k]\cup E}$, we get

$$\|(I_\infty - I_\mathcal{M})B_iA_iI_\mathcal{M}\|_\tau^2 \le \|I_\mathcal{M}(I_\infty - B_iA_i)I_\mathcal{M}\|_\tau^2 = \|A_i - I_\mathcal{M}B_iI_\mathcal{M}\|_\tau^2 .$$

Plugging all of this back to (♠), we get

$$\begin{aligned}(\sqrt{2}-\sqrt{\varepsilon})^2 \le (\spadesuit)^2 &\\ &\le 8\|(A_i - B_i)I_\mathcal{M}\|_\tau^2 + 8\|(A_j - B_j)I_\mathcal{M}\|_\tau^2\\ &\le 16\|I_\mathcal{M}(A_i - B_i)I_\mathcal{M}\|_\tau^2 + 16\|I_\mathcal{M}(A_j - B_j)I_\mathcal{M}\|_\tau^2.\end{aligned}$$

and since we assumed $\varepsilon < 3 - 2\sqrt{2}$, we have $\sqrt{2} - \sqrt{\varepsilon} > 1$ and

$$\|I_\mathcal{M}(A_i - B_i)I_\mathcal{M}\|_\tau^2 + \|I_\mathcal{M}(A_j - B_j)I_\mathcal{M}\|_\tau^2 > 1/16.$$

Now, let $i_1j_1, ..., i_cj_c$ be the edges of a maximal matching in E. Then,

$$k\varepsilon \ge \sum_{i\in[k]} \|I_\mathcal{M}(A_i - B_i)I_\mathcal{M}\|_\tau^2 \ge \sum_{t=1}^{c} \|I_\mathcal{M}(A_{i_t} - B_{i_t})I_\mathcal{M}\|_\tau^2 + \|I_\mathcal{M}(A_{j_t} - B_{j_t})I_\mathcal{M}\|_\tau^2 > c/16.$$

This finishes the proof. □

By combining Claims **??** and **??** applied to a graph which is a mathcing with c edges, we deduce Lemma **??**.

A.1 The L^∞ analogue

As discussed in the introduction (see Remark **??**), it is more common in stability literature to use a L^∞ analogue of Definition **??**, where the notion of almost-homomorphism and closeness are both measured by taking a supremum over relations and generators respectively, as opposed to averaging according to distributions μ_R, μ_S. Let us recall the exact definition. We say that a homomorphism $\rho\colon \mathcal{F}(S) \to \mathcal{U}(\mathcal{M})$ is an (ε, ∞)-approximate representation if

$$\forall r \in R : \quad \|\rho(r) - I_\mathcal{M}\|_\tau^2 \le \varepsilon.$$

Furthermore, homomorphisms $\rho\colon \mathcal{F}(S) \to \mathcal{U}(\mathcal{M}), \varphi\colon \mathcal{F}(S) \to \mathcal{U}(\mathcal{N})$ are (δ, ∞)-close if there exists an isometry $w \in P\mathcal{M}_\infty I_\mathcal{M}$ such that

$$\forall s \in S\colon \quad \|\rho(s) - w^*\varphi(s)w\|_\tau^2 \le \delta.$$

The goal of this subsection of the appendix is to provide a somewhat general procedure to convert lower bounds on the modulus of stability with respect to μ_S, μ_R into a lower bound on the L^∞ modulus of stability.

Let $\langle S\colon R\rangle$ be a presentation of a group Γ, and let μ_S, μ_R be fully supported distributions over the generators and relations respectively. Let $\sigma \in \mathrm{Sym}(S)$ be a permutation of the generators. Then, σ extends (by the universal property of the free group) to an automorphism of $\mathcal{F}(S)$ which we still denote by σ as well. The *automorphism group* of the presentation $\langle S\colon R\rangle$ is the subgroup of permutations in $\mathrm{Sym}(S)$ that preserve R. Namely,

$$\Phi = \mathrm{Aut}(\langle S\colon R\rangle) = \{\sigma \in \mathrm{Sym}(S) \mid R = \sigma(R)\}.$$

Assume $R = \bigsqcup R_i$ is the decomposition of R into orbits of Φ, and assume μ_R is uniform over orbits. Let $\rho\colon \mathcal{F}(S) \to \mathcal{U}(\mathcal{M})$ be an (ε, μ_R)-approximate representation of Γ. Define $\rho'\colon \mathcal{F}(S) \to \mathcal{U}(\bigoplus_{\alpha\in\Phi} \mathcal{M})$ as follows:

$$\forall s \in S: \quad \rho'(s) = \bigoplus_{\alpha\in\Phi} \rho(\alpha(s)).$$

We denote by $\mathcal{M}_\Phi = \bigoplus_{\alpha\in\Phi} \mathcal{M} \subseteq \mathcal{M}_\infty$ and by $\mathcal{M}_\alpha$ the copy of $\mathcal{M}$ at the $\alpha \in \Phi$ position.Note that $\mathcal{M}_\Phi$ embeds in $\mathcal{M}_\infty$, e.g. in the coordinates $1, ..., |\Phi|$, and inherits a trace from it by defining

$$\tau^{\mathcal{M}_\Phi}(\bigoplus A_\alpha) = \frac{\tau_\infty(\bigoplus A_\alpha)}{\tau_\infty(I_{\mathcal{M}_\Phi})} = \frac{1}{|\Phi|} \sum_{\alpha\in\Phi} \tau^{\mathcal{M}}(A_\alpha) = \mathop{\mathbb{E}}_{\alpha\in\Phi} \tau^{\mathcal{M}}(A_\alpha).$$

To be consistent, whenever we use $\|X\|_\tau^2$ it means

$$\tau_\infty(X^*X) = \tau_{\mathcal{M}}(X^*X) = |\Phi| \cdot \tau_{\mathcal{M}_\Phi}(X^*X).$$

Let $w_i = \mu_R(R_i)$, namely the probability that μ_R samples a relation from the orbit R_i. Since we assumed μ_R is uniform over orbits, and since Φ acts transitively on each orbit, we know that

$$\forall r \in R_i: \quad \mathop{\mathbb{E}}_{\alpha\in\Phi} f(\alpha(r)) = \mathop{\mathbb{E}}_{r'\in R_i} f(r)$$

for any function $f\colon \mathcal{F}(S) \to \mathbb{C}$. Therefore,

$$\begin{aligned}
\forall r \in R_i: \quad \frac{w_i}{|\Phi|}\|\rho'(r) - I_{\mathcal{M}_\Phi}\|_\tau^2 &= w_i \mathop{\mathbb{E}}_{\alpha\in\Phi} \|\rho(\alpha(r)) - I_{\mathcal{M}}\|_\tau^2 \\
&= w_i \mathop{\mathbb{E}}_{r'\in R_i} \|\rho(r') - I_{\mathcal{M}}\|_\tau^2 \\
&\le \sum_j w_j \mathop{\mathbb{E}}_{r'\in R_j} \|\rho(r') - I_{\mathcal{M}}\|_\tau^2 \\
&= \mathop{\mathbb{E}}_{r\sim\mu_R} \|\rho(r) - I_{\mathcal{M}}\|_\tau^2 \le \varepsilon.
\end{aligned}$$

Hence,

$$\forall r \in R: \ \frac{1}{|\Phi|}\|\rho'(r) - I_{\mathcal{M}_\Phi}\|_\tau^2 \le \max(1/w_i) \cdot \varepsilon,$$

which in turn means that ρ' is a $(\max_i\{1/w_i\} \cdot \varepsilon, \infty)$-approximate representation — since

$$\frac{1}{|\Phi|}\|\cdot\|_\tau^2 = \|\cdot\|_{\tau^{\mathcal{M}_\Phi}}^2,$$

which is the relevant parameter to consider.

Corollary A.4. *Any (ε, μ_R)-approximate representation of $\langle S\colon R\rangle$, where μ_R is uniform over orbits of $\Phi = Aut(\langle S\colon R\rangle)$, can be transformed into a $(\max_i\{1/w_i\} \cdot \varepsilon, \infty)$-approximate representation.*

Remark A.5. *By applying this construction on the example from Claim **??**, the resulting ρ' is a $\left(2|E|/\binom{k}{2}, \infty\right)$-approximate representation of (**??**). This is because $\Phi = Sym(S)$, and there are two orbits — the commutation relations and the involutions — where μ_R is supprted equally on each of them. Namely, $w_{commutation} = w_{involution} = 1/2$.*

Recall that our goal is to translate L^1 lower bounds into L^∞ lower bounds. Namely, we would like to deduce that, if every genuine representation of Γ is (δ, μ_S)-far from ρ, then every genuine representation of Γ is (δ', ∞)-far from ρ', for $\delta' = C\delta$. Since μ_S was assumed to be fully supported, the L^∞ distance is lower bounded by the L^1-distance, and being (δ', μ_S)-far from ρ' implies being (δ', ∞)-far from it. Thus, we can forget about the L^∞ notion of distance and lower bound our usual notion of distance.

To that end, assume that every genuine representation of Γ is (δ, μ_S)-far from ρ. Namely, for every genuine representation $\varphi\colon \Gamma \to \mathcal{U}(\mathcal{N})$, where $\mathcal{N} = P\mathcal{M}_\infty P$, and every isometry $w \in P\mathcal{M}_\infty I_{\mathcal{M}}$, we have

$$\max\left\{ \mathop{\mathbb{E}}_{s\sim\mu_S} \|\rho(s) - w^*\varphi(s)w\|_\tau^2 \,,\, \tau^{\mathcal{M}}(I_{\mathcal{M}} - w^*w) \right\} \geq \delta.$$

Let $\varphi'\colon \Gamma \to \mathcal{U}(\mathcal{N})$ be a genuine representation of Γ which is (δ', μ_S)-close to ρ', i.e. it satisfies

$$\frac{1}{|\Phi|} \mathop{\mathbb{E}}_{s\sim\mu_S} \|\rho'(s) - w^*\varphi'(s)w\|_\tau^2 \,,\, \tau^{\mathcal{M}_\Phi}(I_{\mathcal{M}_\Phi} - w^*w) \leq \delta',$$

where again $\|X\|_\tau^2 = \tau_\infty(X^*X)$. As before, we can assume $w = PI_{\mathcal{M}_\Phi}$. Then,

$$\begin{aligned}\mathop{\mathbb{E}}_{\alpha\in\Phi}\mathop{\mathbb{E}}_{s\sim\mu_S} \|\rho(\alpha(s)) - I_{\mathcal{M}_\alpha}\varphi'(s)I_{\mathcal{M}_\alpha}\|_\tau^2 &= \frac{1}{|\Phi|}\mathop{\mathbb{E}}_{s\sim\mu_S}\sum_{\alpha\in\Phi} \|I_{\mathcal{M}_\alpha}\rho'(s)I_{\mathcal{M}_\alpha} - I_{\mathcal{M}_\alpha}\varphi'(s)I_{\mathcal{M}_\alpha}\|_\tau^2 \\ &\leq \frac{1}{|\Phi|}\mathop{\mathbb{E}}_{s\sim\mu_S} \|\rho'(s) - I_{\mathcal{M}_\Phi}\varphi'(s)I_{\mathcal{M}_\Phi}\|_\tau^2 \leq \delta'.\end{aligned}$$

In particular, by Markov's inequality, for at least two thirds of the $\alpha \in \Phi$ we have

$$\mathop{\mathbb{E}}_{s\sim\mu_S} \|\rho(\alpha(s)) - I_{\mathcal{M}_\alpha}\varphi'(s)I_{\mathcal{M}_\alpha}\|_\tau^2 \leq 3\delta'.$$

Similarly,

$$\begin{aligned}\delta' &\geq \tau^{\mathcal{M}_\Phi}(I_{\mathcal{M}_\Phi} - I_{\mathcal{M}_\Phi}PI_{\mathcal{M}_\Phi}) \\ &= \mathop{\mathbb{E}}_{\alpha\in\Phi} \tau_\infty(I_{\mathcal{M}_\alpha} - I_{\mathcal{M}_\alpha}PI_{\mathcal{M}_\alpha})\,,\end{aligned}$$

and for at least two thirds of the $\alpha \in \Phi$ we have

$$\tau^{\mathcal{M}_\alpha}(I_{\mathcal{M}_\alpha} - I_{\mathcal{M}_\alpha}PI_{\mathcal{M}_\alpha}) \leq 3\delta'.$$

Hence, $\delta' \geq \frac{\delta}{3}$, and we deduce that every genuine representation of $\Gamma \cong \langle S\colon R\rangle$ is at least $(\delta/3, \infty)$-away from ρ'.

Corollary A.6. *The transformation $\rho \mapsto \rho'$ we described translates a (ε, μ_R)-approximate representation into a $((\max 1/w_i)\cdot\varepsilon, \infty)$-approximate representaion, and if every genuine representation of Γ is (δ, μ_S)-far from ρ, then every genuine representation is $(\delta/3, \infty)$-far from ρ'.*

Remark A.7. *By applying this corollary to the construction from the beginning of the appendix, we conclude that the L^∞ modulus of stability of the presentation* (**??**) *is $\Omega(k\varepsilon)$.*

References

[AFRV19] Rotem Arnon-Friedman, Renato Renner, and Thomas Vidick. Simple and tight device-independent security proofs. *SIAM Journal on Computing*, 48(1):181–225, 2019.

[ALM+98] Sanjeev Arora, Carsten Lund, Rajeev Motwani, Madhu Sudan, and Mario Szegedy. Proof verification and the hardness of approximation problems. *J. ACM*, 45(3):501–555, may 1998.

[BC22] Oren Becker and Michael Chapman. Stability of approximate group actions: uniform and probabilistic. *Journal of the European Mathematical Society*, 2022.

[BCLV23] Lewis Bowen, Michael Chapman, Alex Lubotzky, and Thomas Vidick. Subgroup tests and the aldous–lyons conjecture. *preprint*, 2023.

[BFL91] László Babai, Lance Fortnow, and Carsten Lund. Non-deterministic exponential time has two-prover interactive protocols. *Computational complexity*, 1:3–40, 1991.

[BGK+97] László Babai, Albert J Goodman, William M Kantor, Eugene M Luks, and Péter P Pálfy. Short presentations for finite groups. *Journal of Algebra*, 194(1):79–112, 1997.

[BL20] Oren Becker and Alexander Lubotzky. Group stability and property (t). *Journal of Functional Analysis*, 278(1):108298, 2020.

[CGJV19] Andrea Coladangelo, Alex B Grilo, Stacey Jeffery, and Thomas Vidick. Verifier-on-a-leash: new schemes for verifiable delegated quantum computation, with quasilinear resources. In *Annual International Conference on the Theory and Applications of Cryptographic Techniques*, pages 247–277. Springer, 2019.

[CHTW04] Richard Cleve, Peter Hoyer, Benjamin Toner, and John Watrous. Consequences and limits of nonlocal strategies. In *Proceedings. 19th IEEE Annual Conference on Computational Complexity, 2004.*, pages 236–249. IEEE, 2004.

[CL23a] Michael Chapman and Alex Lubotzky. Stability of homomorphisms, coverings and cocycles I: Equivalence. *preprint*, 2023.

[CL23b] Michael Chapman and Alex Lubotzky. Stability of homomorphisms, coverings and cocycles II: Examples, Applications and Open problems. *preprint*, 2023.

[CLS17] Richard Cleve, Li Liu, and William Slofstra. Perfect commuting-operator strategies for linear system games. *Journal of Mathematical Physics*, 58(1), 2017.

[CM14] Richard Cleve and Rajat Mittal. Characterization of binary constraint system games. In *Automata, Languages, and Programming: 41st International Colloquium, ICALP 2014, Copenhagen, Denmark, July 8-11, 2014, Proceedings, Part I 41*, pages 320–331. Springer, 2014.

[CRSV17] Rui Chao, Ben W Reichardt, Chris Sutherland, and Thomas Vidick. Overlapping qubits. In *8th Innovations in Theoretical Computer Science Conference (ITCS 2017)*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2017.

[DCGLT20] MARCUS DE CHIFFRE, LEV GLEBSKY, ALEXANDER LUBOTZKY, and ANDREAS THOM. Stability, cohomology vanishing, and nonapproximable groups. *Forum of Mathematics, Sigma*, 8:e18, 2020.

[DCOT19] Marcus De Chiffre, Narutaka Ozawa, and Andreas Thom. Operator algebraic approach to inverse and stability theorems for amenable groups. *Mathematika*, 65(1):98–118, 2019.

[DEL+22] Irit Dinur, Shai Evra, Ron Livne, Alexander Lubotzky, and Shahar Mozes. Locally testable codes with constant rate, distance, and locality. In *Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2022, page 357–374, New York, NY, USA, 2022. Association for Computing Machinery.

[dlS21] Mikael de la Salle. Orthogonalization of positive operator valued measures. *arXiv preprint arXiv:2103.14126*, 2021.

[dlS22] Mikael de la Salle. Spectral gap and stability for groups and non-local games. *arXiv preprint arXiv:2204.07084*, 2022.

[GH17] William Timothy Gowers and Omid Hatami. Inverse and stability theorems for approximate representations of finite groups. *Sbornik: Mathematics*, 208(12):1784, 2017.

[GKKL08] Robert Guralnick, Willim Kantor, Martin Kassabov, and Alex Lubotzky. Presentations of finite simple groups: a quantitative approach. *Journal of the American Mathematical Society*, 21(3):711–774, 2008.

[Gle10] Lev Glebsky. Almost commuting matrices with respect to normalized hilbert-schmidt norm. *arXiv preprint arXiv:1002.3082*, 2010.

[GR09] Lev Glebsky and Luis Manuel Rivera. Almost solutions of equations in permutations. *Taiwanese Journal of Mathematics*, 13(2A):493–500, 2009.

[Hal76] PR Halmos. Some unsolved problems of unknown depth about operators on hilbert space. *Proceedings of the Royal Society of Edinburgh Section A: Mathematics*, 76(1):67–76, 1976.

[HS18] Don Hadwin and Tatiana Shulman. Stability of group relations under small Hilbert-Schmidt perturbations. *Journal of Functional Analysis*, 275(4):761–792, 2018.

[Ioa20] Adrian Ioana. Stability for product groups and property (τ). *Journal of Functional Analysis*, 279(9):108729, 2020.

[JNV+20a] Zhengfeng Ji, Anand Natarajan, Thomas Vidick, John Wright, and Henry Yuen. MIP*= RE. *arXiv preprint arXiv:2001.04383*, 2020.

[JNV+20b] Zhengfeng Ji, Anand Natarajan, Thomas Vidick, John Wright, and Henry Yuen. Quantum soundness of the classical low individual degree test. *arXiv preprint arXiv:2009.12982*, 2020.

[JNV+21] Zhengfeng Ji, Anand Natarajan, Thomas Vidick, John Wright, and Henry Yuen. MIP* =RE. *Communications of the ACM*, 64(11):131–138, 2021.

[JNV+22] Zhengfeng Ji, Anand Natarajan, Thomas Vidick, John Wright, and Henry Yuen. Quantum soundness of testing tensor codes. In *2021 IEEE 62nd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 586–597. IEEE, 2022.

[Kaz82] David Kazhdan. On ε-representations. *Israel Journal of Mathematics*, 43:315–323, 1982.

[KM19] Dmitry N Kozlov and Roy Meshulam. Quantitative aspects of acyclicity. *Research in the Mathematical Sciences*, 6:1–32, 2019.

[KPS18] Se-Jin Kim, Vern Paulsen, and Christopher Schafhauser. A synchronous game for binary constraint systems. *Journal of Mathematical Physics*, 59(3), 2018.

[KV11] Julia Kempe and Thomas Vidick. Parallel repetition of entangled games. In *Proceedings of the forty-third annual ACM symposium on Theory of computing*, pages 353–362, 2011.

[NV17] Anand Natarajan and Thomas Vidick. A quantum linearity test for robustly verifying entanglement. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing*, pages 1003–1015, 2017.

[NV18] Anand Natarajan and Thomas Vidick. Low-degree testing for quantum states, and a quantum entangled games pcp for qma. In *2018 IEEE 59th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 731–742. IEEE, 2018.

[NZ23] Anand Natarajan and Tina Zhang. Bounding the quantum value of compiled nonlocal games: from CHSH to BQP verification. *arXiv preprint arXiv:2303.01545*, 2023.

[PK22] Pavel Panteleev and Gleb Kalachev. Asymptotically good quantum and locally testable classical ldpc codes. In *Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2022, page 375–388, New York, NY, USA, 2022. Association for Computing Machinery.

[PRSS22] Connor Paddock, Vincent Russo, Turner Silverthorne, and William Slofstra. Arkhipov's theorem, graph minors, and linear system nonlocal games. *arXiv preprint arXiv:2205.04645*, 2022.

[RS96] Ronitt Rubinfeld and Madhu Sudan. Robust characterizations of polynomials with applications to program testing. *SIAM Journal on Computing*, 25(2):252–271, 1996.

[Sch80] Jacob Schwartz. Fast probabilistic algorithms for verification of polynomial identities. *Journal of the ACM*, 27(4):701–717, 1980.

[Slo19] William Slofstra. The set of quantum correlations is not closed. In *Forum of Mathematics, Pi*, volume 7, page e1. Cambridge University Press, 2019.

[SV18] William Slofstra and Thomas Vidick. Entanglement in non-local games and the hyperlinear profile of groups. In *Annales Henri Poincaré*, volume 19, pages 2979–3005. Springer, 2018.

[Ula60] Stanislaw Ulam. A collection of mathematical problems. *Interscience Tracts in Pure and Applied Mathematics, no. 8, Interscience Publishers, New York-London*, 1960.

[Vid22] Thomas Vidick. MIP* =RE, a negative resolution to Connes' embedding problem and Tsirelson's problem. 2022. Notes prepared for the ICM'22, available at http://users.cms.caltech.edu/~vidick/notes/ICM.pdf.

[VN42] John Von Neumann. Approximative properties of matrices of high finite order. *Portugaliae mathematica*, 3(1):1–62, 1942.

[Voi83] Dan Voiculescu. Asymptotically commuting finite rank unitary operators without commuting approximants. *Acta Sci. Math.(Szeged)*, 45(1-4):429–431, 1983.

[VV19] Umesh Vazirani and Thomas Vidick. Fully device independent quantum key distribution. *Communications of the ACM*, 62(4):133–133, 2019.

[Zip79] Richard Zippel. Probabilistic algorithms for sparse polynomials. In *Symbolic and Algebraic Computation*, pages 216–226, 1979.